New tools to study the consequences of micronucleation and micronucleus rupture

Lucian DiPeso

A dissertation
submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee:
Emily M. Hatch, Chair
Toshio Tsukiyama
Douglas M. Fowler

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington

Abstract

New tools to study the consequences of micronucleation and micronucleus rupture

Lucian DiPeso

Chair of the Supervisory Committee:
Emily M. Hatch
Biochemistry

Micronuclei are aberrant nuclear compartments that trap a portion of a cell's chromatin in a distinct organelle separate from the nucleus and are drivers of inflammation, DNA damage, chromosome instability, and chromothripsis. Many of the consequences of micronucleus formation stem from micronucleus rupture: the sudden loss of micronucleus compartmentalization, resulting in mislocalization of nuclear factors and the exposure of chromatin to the cytosol for the remainder of interphase. Micronuclei form primarily from segregation errors during mitosis, errors that also give rise to other, non-exclusive phenotypes, including aneuploidy and chromatin bridges. The stochastic formation of micronuclei and phenotypic overlap confounds the use of population-level assays or hypothesis discovery, requiring labor-intensive techniques to visually identify and follow micronucleated cells individually. In this work, I present a novel technique for automatically identifying and isolating micronucleated cells generally and cells with ruptured micronuclei specifically using a *de novo* neural net combined with Visual Cell Sorting. As a proof of concept, I compare the early transcriptomic responses to micronucleation and micronucleus rupture with previously published responses to aneuploidy, revealing micronucleus rupture to be a potential driver of the aneuploidy response. In addition, I present an unfinished tool using the bacterial enzyme deoxyadenosine methyltransferase (Dam) for the population-level identification of

chromatin, rather than cells, that were trapped in a ruptured micronucleus. This tool, dubbed Dam Tracker, depended on the sequestration of Dam in the cytosol away from chromatin in intact nuclear compartments, which is made difficult during mitosis when nuclear envelopes are disassembled. Should this obstacle be overcome, this tool could open new avenues of research on the long-term consequences of micronucleus rupture to chromatin.

Introduction

Micronuclei (MN) are aberrant nuclear compartments that trap some portion of a cell's chromatin in a spatially distinct organelle separate from the nucleus. They are produced principally via segregation errors or the presence of unrepaired DNA breaks during mitosis.[1] MN have been observed since the late 19th century, but in recent years, they were discovered to be key drivers of chromothripsis, chromosome instability (CIN), inflammation, and even embryonic mosaicism.[2–7] The realization that MN are not merely symptoms but rather drivers of oncogenesis and CIN has fueled intense research in the past two decades. However, this research has been hampered by an inability to produce homogeneous populations of micronucleated cells, requiring the use of labor- and time-intensive protocols to manually identify micronucleated cells and track them over time.

To accelerate research on the causes and consequences of MN formation and MN rupture, this work sought to develop new research tools to overcome some of these limitations—in particular, to enable the investigation of intergenerational consequences of MN rupture.

Micronucleus biogenesis

There are several reported processes that can produce MN, broadly categorized as follows:[1]

1. Spatial mislocalization of chromatin during mitosis
2. Chromatin bridge resolution
3. Defects during mitotic exit
4. Nuclear budding

Spatial mislocalization of chromatin during mitosis

The most common mechanism for MN biogenesis is chromatin mislocalization. Human cells undergo open mitosis: the nuclear envelope begins disassembly during prometaphase to facilitate proper spindle assembly and chromosome segregation

and then reassembled at the end of mitosis.[8] Nuclear envelope reassembly begins with endoplasmic reticulum (ER) recruitment to chromatin surfaces.[9] If a chromosome or chromatin fragment is too far from the bulk chromatin mass for recruited membrane to form a single, contiguous nuclear envelope, multiple nuclear compartments will form: the nucleus and one or more MN. Spatial mislocalization can occur through a variety of mechanisms, including merotelic attachments, spindle assembly checkpoint (SAC) failures, the loss of centromeres, and DNA fragmentation.[1]

Chromatin bridge resolution

Chromatin bridges form when the same chromosome is captured by both opposing spindle poles—the result of dicentric chromosomes produced by telomere fusion or circularization of sister chromatids following double-strand breaks—and pulled into both daughter cells' nuclei, forming a bridge between the two cells.[10] If the bridge breaks before mitosis completes, DNA fragments can be too far from the rest of the cell's chromatin during nuclear envelope reassembly, generating MN.[11,12]

Mitotic defects

MN can form if nuclear envelope reassembly is defective—if membrane recruited to individual chromosomes fails to properly fuse or cytokinesis fails to cleave all daughter cells apart following multipolar mitosis.[12,13]

Nuclear budding

There are reports of nuclear budding producing MN, where some chromatin is actively extruded from a nucleus during interphase. This has been observed following nuclear envelope rupture, in cells with double-minutes, and in some other cases, often in conjunction with observed defects in nuclear lamina.[12,14] However, it is unclear if this process occurs beyond the cell lines in which it has been observed or outside of cell culture conditions. In addition, these compartments form not during nuclear envelope reassembly but during interphase, raising the question of how

similarly they would behave to MN formed from mitotic errors or whether they should even be considered the same type of structure.

While multiple routes of MN formation have been observed, the most common cause (and the most likely source *in vivo*) remains mislocalized chromatin—chiefly from lagging chromosomes or chromatin fragmentation.

Micronucleus structure and function

Once formed, MN resemble nuclei in general structure and function—they possess a double-membrane envelope punctuated with nuclear pore complexes (NPCs) and undergirded with a nuclear lamina. They perform transcription and DNA replication. However, they are flawed nuclear compartments. Imaging reveals that many MN have defects in their nuclear lamina, including what appear to be holes in the lamina meshwork.[15,16] While they usually contain NPCs, they are at a much lower density than what is found in the nucleus, in some cases resulting in little or no nuclear import.[16,17] Transcription is hindered and DNA replication is often delayed or may stall out entirely, leading to replication fork collapse.[15,17,18]

Like the nucleus, MN suffer from nuclear envelope rupture, a process where the nucleus loses compartmentalization during interphase resulting in free intermixing of nuclear and cytoplasmic components.[15,19] In the nucleus, this process is usually transient, with the nuclear envelope rapidly repaired.[19,20] In contrast, MN are seldom capable of repairing their nuclear envelope. This appears to be in part due to misregulation of the ESCRTIII machinery necessary to seal ruptured membrane.[21–24] This long-term loss of MN compartmentalization leads to micronuclear envelope collapse, where ER membrane invades the nucleoplasmic space, and is linked to many of the consequences of MN formation.[15]

Consequences of micronucleus rupture

When micronuclear envelopes collapse, all nuclear processes cease and chromatin is rapidly condensed, stripped of acetylation marks, and damaged.[5,7] For the rest of interphase, any contained chromatin will effectively be lost to the cell—functionally, the cell will be aneuploid. In addition, rupture exposes DNA to the cytoplasm, potentially activating the proinflammatory cGAS/STING pathway or attack by cytoplasmic endonucleases.[5,25–28] However, despite cGAS's canonical role in sequence-independent recognition of cytosolic DNA, there is growing evidence that not all ruptured MN recruit cGAS. MacDonald, *et al.* report that cGAS binding is dependent on histone methylation states, while Mohr, *et al.* report that the ER-associated endonuclease TREX1 inhibits cGAS recruitment to ruptured MN.[27,29] Recent work on cGAS localization has identified a large pool that resides in the nucleus, but is inhibited by some unidentified histone tethering mechanism.[30] cGAS, then, may already reside in some intact MN in an inactivated state and block further cGAS binding following MN rupture, or that the same mechanisms that inhibit cGAS activation in the nucleus may continue to function in a subset of ruptured MN.

Following rupture, collapse of replication forks, premature chromatin condensation, and attack by cytoplasmic nucleases may contribute to the DNA damage frequently seen in MN.[18,27,28,31] DNA damage may be so severe following rupture that it results in a process known as chromothripsis, where a chromosome is "pulverized" into dozens of fragments; following mitosis, if these pieces are incorporated into a daughter cell's nucleus, they will be religated together randomly by non-homologous end-joining (NHEJ), leading to inversions, translocations, deletions, kataegis, and the formation of circular double-minutes.[7,32] Even absent extreme forms of DNA damage, changes to chromatin in ruptured MN persist into daughter cells. If rupture occurs before DNA replication has finished, at least one daughter cell will necessarily be aneuploid.[17] Following cell division and reincorporation into a daughter cell's nucleus, DNA that had been in a ruptured MN sees depressed transcription, loss of euchromatin marks, and increased markers of DNA damage.[33]

MN can arise because of existing CIN or DNA damage; however, their propensity to rupture and the consequences thereof results in more DNA damage, increased rates of aneuploidy, persistent changes in gene expression, increased CIN, and an increased propensity for MN formation. This feed-forward loop can rapidly drive evolution and oncogenesis rather than be a follow-on development of malignant lesions, revealed by cancer genotyping identifying chromothripsis as an originating event in many cancers.[34]

Micronucleus model systems

Research on MN generally rely on protocols to disrupt mitosis and induce micronucleation in model cell lines. These methods include microtubule poisons like nocodazole, hyper- or hypothermia to destabilize kinetochore attachments, or inhibitors of a key kinase of the SAC—Mps1 inhibitors—to rush mitosis before chromosomes have a chance to properly align.[35–38] In addition to producing MN, these methods will also result in chromatin bridge formation, aneuploidy, and abnormal nuclear shapes. To limit study to micronucleation requires identifying micronucleated cells by eye to avoid confusing the consequences of micronucleation and these other possible phenotypes. Indeed, as this work finds, the use of Mps1 inhibitors to study aneuploidy without considering these overlapping phenotypes may have confounded the consequences of MN rupture with that of aneuploidy.

To overcome the need for labor- and time-intensive protocols, other models have been developed. Crasta, *et al.* used a cell line with an artificial chromosome whose kinetochore could be selectively inactivated, allowing them to induce micronucleation in a bulk population with the added benefit of knowing the contents of the MN, without otherwise disrupting mitosis.[39] In a similar vein, Ly, *et al.* relied on unique differences in the Y-chromosome centromere to selectively inactivate the Y-chromosome kinetochore, again allowing them to induce micronucleation of a specific chromosome in a bulk population.[32] However, these methods limit research to the effects of micronucleation on a specific chromosome (either a wholly artificial chromosome or a unique chromosome) in specific cell lines.

An older, simpler technique is to limit the study of MN to erythrocytes, an enucleated cell type. Micronucleated erythrocytes can be sorted out by gating for erythrocytes with a positive Hoechst signal, since erythrocyte enucleation fails to remove MN.[40,41] However, studying a very unique, terminally differentiated cell type without any obvious link to cancer development *in vivo* limits the usefulness of this technique.

Goals of this study

This dissertation summarizes my work to develop tools to study the consequences of micronucleation without complex model systems or manual identification of micronucleated cells. There are two projects which I pursued:

- A protocol to identify and isolate *en masse* micronucleated cells, and further distinguish between cells with and without ruptured MN, through the use of artificial intelligence to automate the identification of MN
- An enzymatic approach to covalently tag chromatin in ruptured MN, which can then be used to pull down tagged chromatin in bulk or track the daughters of cells that had undergone MN rupture

Each project will be presented separately, with its own figures, results, and methods sections.

I. Using machine vision to identify and isolate micronucleated cells

In a collaboration with the Fowler Lab at the University of Washington, we sought to develop a general protocol to identify and isolate micronucleated cells, and to further distinguish cells with ruptured MN versus those with intact MN. We opted to develop a neural net that could segment MN from micrographs taken with only a 20x widefield objective. We coupled this neural net with Visual Cell Sorting (VCS), a technique developed by the Fowler Lab to isolate a specific population of cells identified by a visual phenotype using FACS. [42] Working off progress made by Sriram Pendyala, an MD/PhD candidate in the Fowler Lab, and Heather Huang, a former lab

technician in the Hatch Lab, I improved and validated the neural net's ability to identify and isolate micronucleated cells. To demonstrate the capability of this system, I used it to profile the early transcriptomic response to micronucleation and MN rupture in an aneuploid cell population. I then developed a more generally applicable neural net, released on the Hatch Lab's GitHub repository, to aid in future research on the effects of MN in other conditions and cell populations.[43]

Results

Machine vision identifies micronucleated RPE-1 cells within a mixed population

We developed an automated pipeline for identifying micronucleated cells in a mixed population of hTERT RPE-1 (RPE-1) cells. RPE-1 cells are a near-diploid, non-transformed human cell line not prone to micronucleation. In addition, the RPE-1 transcriptomic response to aneuploidy is already well studied, providing a convenient background to examine the specific effects of micronucleation and MN rupture.[44–46] Treatment with a low dose of the Mps1 inhibitor BAY-1217389 (Mps1i) causes whole chromosome missegregation and micronucleation in approximately 50% of cells (**Fig 1-S1a**) with a modal MN number of 1 per cell in MN+ cells (**Fig 1-S1b**). To facilitate identification of MN and the use of VCS, we transduced RPE-1 cells with H2B fused to emiRFP703—a far-red fluorophore—and a 3-fold repeat of Dendra2, a UV-convertible fluorophore, tagged with a nuclear localization signal (NLS) (**Fig 1-1a**).[15,47] This cell line will hereafter be referred to as RFP703/Dendra cells. To maximize the efficiency of the VCS protocol, we imaged RFP703/Dendra cells on a widefield fluorescent microscope at 20x (**Fig 1-1b**) and developed an image analysis pipeline that privileges processing speed and positive predictive value (PPV) over recall. In this pipeline, nuclei are segmented using the H2B signal with a pre-existing neural network trained to recognize nuclei (see Methods).[48] Image crops centered at each nucleus segment are generated and processed with Sobel edge detection, added as a second channel. Each 2-channel crop is used as input into a custom neural network that segments MN within each crop. These MN segments are reassembled onto the full, uncropped micrograph and assigned to "parent" nuclei by proximity.

Manual analysis of MN association with nuclei demonstrated that proximity is sufficient to correctly associate 97% of MN (**Fig 1-S1c**). Nuclei with associated MN are classed as micronucleus+ (MN+) and those without as micronucleus- (MN-). Comparison of MN masks with manually identified MN demonstrated a recall value of approximately 70% and a PPV of 89% with a statistically significant reduction in recall among ruptured MN, determined manually by loss of NLS-3xDendra2 (**Fig 1-1c**). When the whole field is considered, including regions that are not captured by the crops or are excluded due to proximity to the field edge, MN recall drops to 50% (**Fig 1-1d**). Recall of MN- and MN+ cells is 86% and 65%, respectively, and PPV is 73% and 93%, respectively (**Fig 1-1e**). In summary, we can automatically identify the majority of micronucleated and non-micronucleated cells in a low-resolution image with very high specificity.

Dendra2 intensity distinguishes intact micronuclei from ruptured micronuclei

Similar to previously used nuclear integrity markers, NLS-3xDendra2 is present in intact MN and absent in ruptured MN (**Fig 1-2a**).[15] To automatically identify cells with ruptured MN, the maximum Dendra2 intensity is measured in the nucleus and corresponding MN segments and the ratio taken. Defining a threshold at 0.16 for MN rupture correctly classifies approximately 90% of MN (**Fig 1-2b**). This information can be integrated into our pipeline to further classify MN+ cells into those with no ruptured MN (rupture- cells) from those with 1 or more ruptured MN (rupture+ cells) (**Fig 1-S2a**), correctly identifying 60% of rupture- cells and 70% of rupture+ cells with a PPV near 75% in both cases (**Fig 1-2c**). The difference in recall is likely due to the increased probability of a multi-micronucleated cell being correctly classified as MN+ and having at least one ruptured MN (**Figs 1-S2b, 1-S2c**).

Visual Cell Sorting isolates multiple populations of photoconverted cells

VCS was previously shown to be able to specifically label and isolate multiple populations of cells by targeting each with a UV pulse of a different duration, which induces differential proportions of UV-converted Dendra2.[42] To confirm that we could

replicate previously demonstrated specificity with our cell line, we plated a 1:1 mixture of cells treated with CellTrace far red or with DMSO loading control (**Figs 1-3a, 1-3b**). We then segmented nuclei using the NLS-3xDendra2 channel and the same nuclear segmenter used in our MN classification pipeline. The masks generated were used to target a UV laser to individual nuclei. We thresholded on the nuclear far-red intensity, targeting CellTrace- cells with 200 ms UV pulses and CellTrace+ cells with 800 ms UV pulses. By thresholding on the Dendra2 red-emission/green-emission ratios during FACS, we could specifically separate a mixed population of cells either stained with CellTrace or left unstained (**Figs 1-3c, 1-3d**).

Cell yield from VCS recapitulates aneuploidy transcriptomic signatures

To confirm that our machine vision/VCS pipeline was suitable for generating input populations for RNAseq and to assess the effect of different UV pulses on cells treated with Mps1i, we photoconverted asynch-ronous RFP703/Dendra cells treated with either Mps1i or DMSO loading control, with cells randomly assigned to either 800 ms or 200 ms of UV (**Fig 1-4a**). Given the duration of the complete experiment, following photocon-version, we reimaged the first and last 5 positions to confirm that photocon-verted cells retained differential levels of Dendra2 ratio (**Fig 1-S3a**). Further, to prevent cells from going through mitosis between the start and end of the experiment—and therefore diluting photoconverted Dendra2 signal or changing their MN status—we added the reversible Cdk1 inhibitor RO-3306 to inhibit mitosis.[49] This experimental design allowed us to look at differential gene expression between Mps1i and DMSO treatments and between UV pulse lengths. Conversion of 1,500–2,000 fields allowed us to capture 13k cells after FACS for each photoconverted population in each condition. Principle components analysis (PCA) revealed that, as expected, cells clustered first by treatment group and then by replicate (**Fig 1-4b**), confirming that the effects of Mps1 inhibition can be differentiated from loading control but also the need to control for batch effects. Differential expression analysis identified 2,200 differentially expressed genes (DEGs), 63 of which had absolute fold-changes >1.5, suggesting a mod-erately strong response to Mps1 inhibition vs DMSO treatment (**Fig 1-4c**). Comparing DMSO-treated cells exposed to 800 ms vs 200 ms of

UV radiation showed no difference in populations, confirming that these data were the result of Mps1i treatment and not the VCS pipeline itself (**Figs 1-S4a, 1-Sb**). To determine whether the VCS pipeline preserved the previously reported aneuploidy response, we compared our results to RNAseq data from two prior studies that induced aneuploidy in RPE-1 cells using either a different Mps1 inhibitor or by releasing cells from an 8-hour nocodazole block.[38,44,45] Both studies identified a characteristic downregulation of many cell-cycle and metabolic pathways alongside upregulation of inflammatory pathways. To directly compare our data, we first reanalyzed these prior studies' raw counts using the same DESeq2 analysis pipeline we used for our own data (see Methods) (**Figs 1-S4c–f**).

Gene-set enrichment analysis (GSEA) using GO terms identified substantial overlap between all three data sets, though with many more terms identified in our data set than the other two (**Fig 1-4d**). GSEA identifies those genes that are most important for driving differential regulation in a pathway, termed *leading edge genes*. All three data sets shared many leading edge genes. GSEA using Hallmark pathways identified a similar downregulation of cell-cycle pathways and upregulation of inflammatory pathways across the data sets (**Fig 1-4e**). Overall, we recapitulated a similar aneuploidy response as had been previously published. Our data set also resulted in many more terms and pathways being identified, likely a result of greater paired-end coverage compared with the shallower, single-end reads used in the earlier studies (**Fig 1-S4g**).

In summary, these data confirm that the VCS pipeline can reproduce data consistent with the previously published transcriptomic signature of aneuploid RPE-1 cells and validates this aneuploidy signature.

Micronucleus rupture, but not micronucleation, induces a limited general transcriptional response

To determine whether the formation of MN induced a transcriptional response, we treated RFP703/Dendra cells with Mps1i and isolated MN+ and MN- cells using our machine-vision pipeline. Prior to processing samples for RNAseq, we performed

several quality control checks: manually confirming in a subset of images that the classifier was performing as expected; replating an aliquot of sorted cells to check sorting accuracy between MN+ and MN- cells (**Fig 1-S5a**); and checking the quality of extracted RNA by TapeStation analysis (**Fig 1-S5b**). Running PCA on these data showed results clustered largely by replicate, rather than micronucleation status. This was confirmed by MA plot, which revealed nearly all DEGs to have absolute fold-changes less than 1.5. Only 2 genes were above this cutoff—the immediate-early gene EGR1, a transcription factor regulating cellular stress and growth responses, and A20, a ubiquitin-editing enzyme regulating inflammatory responses (**Figs 1-5b, 1-5c**).[50–53] These data suggest that micronucleation alone produces a small initial transcriptional response in aneuploid cells.

MN continue to rupture throughout interphase. [15,16] Over the 4–5 hours necessary for photoconversion and isolation, we expected some intact MN to rupture, decreasing the purity of the rupture- population (and conversely increasing the purity of the rupture+ population). To ensure minimal and predictable MN rupture during analysis and conversion of cells, we synchronized the cells in G1 using a Cdk4/6 inhibitor prior to release into Mps1i.[54] This causes a larger but more predictable change in rupture frequency compared with asynchronous cells (**Fig 1-S5c**). In addition, a single technical replicate was performed per experiment to minimize the time between photoconversion and trypsinization (**Fig 1-5d**). Using this protocol, we determined the maximum change in rupture frequency from the start of an experiment vs the end, performed by re-imaging the first few fields of view at the end of the experiment. We found a significant increase in rupture frequency and used this change to estimate a reduction in rupture- PPV with a concomitant increase in rupture+ PPV (**Fig 1-5e**). These data suggest that RNAseq would still capture differential expression between these two populations, despite a loss in purity of the rupture- population.

To determine whether MN rupture induced a general transcriptional signature, we performed RNAseq analysis on isolated cells enriched for either intact-only MN (rupture-) or at least one ruptured MN (rupture+). PCA showed clustering based on

condition, indicating a significant difference between the two cell populations. Differential expression analysis confirmed a transcriptomic change with 106 DEGs, 14 of which had absolute fold changes greater than 1.5 (**Figs 1-5f, 1-5g**). Both EGR1 and A20 are upregulated in cells with ruptured MN compared to cells with intact MN.

Aneuploidy is slightly enriched in cells with ruptured MN

To assess whether the transcriptional changes we detected in cells with ruptured MN were the result of micronucleation or aneuploidy, we analyzed single chromosome aneuploidies by FISH (**Fig 1-5h**). Asynchronous RPE-1 cells were treated with Mps1i for 24 h, fixed, and labeled with probes against chromosomes 1, 11, or 18. Quantification of the number of foci per cell demonstrated that aneuploidy was correlated with micronucleation for chromosomes 11 and 18 (**Fig 1-5i**). The contribution of micronucleation status to aneuploidy frequency was analyzed using a generalized estimating equation (GEE) with MN status, chromosome identity, and their interaction modeled as predictors against the frequency of single-chromosome aneuploidy. Micronucleation was found to be statistically significant, although with differing rates depending on the chromosome. While chromosome 1 was more likely to missegregate than the other two chromosomes (**Fig 1-S5d**), when missegregated, it was less likely to correlate with the formation of MN or multiple micronucleation events (**Fig 1-S5e**). The accompaniment of multiple MN and micronucleus formation with missegregation suggests missegregation of chromosomes 11 and 18 may be indicative of more severe mitotic defects, increasing the probability of aneuploidy and micronucleation. We performed the same aneuploidy analysis comparing cells with intact vs ruptured MN and found a strong correlation with chromosome identity, as before, but that rupture alone was not a significant predictor of aneuploidy (**Fig 1-5j**). However, chromosome identity and rupture together were statistically significant, confirming an increase in the rates of aneuploidy among cells with missegregation in chromosomes 11 or 18. A correlation between MN rupture and aneuploidy may be explained by the increased frequency of multiple micronucleation events in cells with missegregated chromosomes 11 or 18 and an increase in the probability of at least one ruptured MN as the number of MN increase (**Figs 1-S5e, 1-S2b**). Chromosomes in

a ruptured MN are no longer available to perform transcription or DNA replication. We reanalyzed the same FISH data, discounting any chromosomes within a ruptured MN to determine rates of "functional aneuploidy," finding nearly identical results (**Figs 1-S5f, 1-S5g**). Together, these data suggest that the transcriptional effects of aneuploidy are likely to be strong when comparing MN+ with MN- cells, but that the contribution of aneuploidy to the transcriptional response in rupture+ cells is less clear, given the weaker link between rupture and aneuploidy and the stronger transcriptional response produced.

ATF3 levels are increased among rupture+ cells

To validate our RNAseq results, we examined the expression of ATF3 in cells with and without ruptured MN by immunofluorescence. ATF3 is an immediate-early gene and transcription factor induced by a variety of stressors, and is transcriptionally upregulated in rupture+ cells (**Fig 1-5g**).[55,56]

As a positive control, we treated cells with doxorubicin, a known activator of ATF3 expression (**Figs 1-6a, 1-S6a**).[57] Asynchronous RPE-1 cells expressing 2xRFP-NLS were incubated with Mps1i or DMSO loading control for 24 hours, including 1 hour of treatment with doxorubicin in positive controls. Cells were then fixed and labeled with antibodies against γH2AX and ATF3, then manually identified as MN-, MN+ & rupture-, or MN+ & rupture+. [15,58–60] ATF3 and γH2AX intensity were quantified in automatically segmented nuclei. ATF3 intensity showed a strong increase in cells treated with doxorubicin and a correlation with γH2AX intensity (**Figs 1-S6b, 1-S6c**). Mps1i treatment induced an increase in ATF3 expression in a subset of cells (**Fig 1-S6a**). Further classification of Mps1i-treated cells into MN+/- and rupture+/- populations identified a small increase in cells with high ATF3 among MN- and rupture+ cells (**Fig 1-6b**). The increase in ATF3-high cells in MN- and rupture+ cells is consistent with our RNAseq results, which found ATF3 to be a DEG when comparing Mps1i treatment against DMSO and comparing rupture+ with rupture- cells, but not between MN+ and MN- cells. (**Figs 1-4c, 1-5c, 1-5g**).

A potential confounding factor is that increased aneuploidy is associated with both an increase in micronucleation and DNA replication defects, DNA damage, and ATF3 activation. To account for this, we constructed a GEE to model the effects of rupture status and γH2AX intensity on ATF3 intensities among rupture+ and rupture- cells (**Fig 1-6c**). This model found that ATF3 intensity is predicted to increase as γH2AX intensity increases, but that this effect is slightly stronger in MN- cells, weak in rupture- cells, and pronounced among rupture+ cells. These data suggest that ATF3 expression is upregulated in response to DNA damage and this response is heightened in cells with ruptured MN, consistent with our RNAseq results.

Micronucleus rupture is a significant driver of transcriptomic changes in Mps1i-treated cells

To determine what pathways were affected by MN rupture, we compared changes to the Hallmark pathway ontology across our three analyses: Mps1i+/-, MN+/-, and rupture+/- conditions. There is broad overlap both in which pathways are differentially regulated and the direction and magnitude of those changes (**Fig 1-7a**). Many pathways share the same leading edge genes. Given the relatively small number of DEGs—of which some are regulatory transcription factors—we took the unique set of leading edge genes found in any differentially regulated pathway from each analysis, in order to determine if there was a core set of genes driving changes across many pathways. While there is some overlap in the leading edge genes from the Mps1i+/- experiment and the MN+/- experiment, the rupture+/- leading edge genes share significant overlap between both (**Fig 1-7b**). This suggests that rupture may be playing an important role in driving the changes in pathway regulation observed in both Mps1i-treated and micronucleated cell populations.

To determine the degree to which micronucleation and micronucleus rupture may be determining the GSEA results seen with Mps1i treatment, we plotted the cumulative proportion of differentially regulated pathways represented by a subset of

leading edge genes (**Figs 1-7c**–**f**, in black). Along the x-axis, we ranked all leading edge genes by the number of pathways they belong to; on the y-axis are the cumulative proportion of differentially regulated pathways represented. All differentially regulated pathways in the Mps1i+/- data set are fully represented by just the first 6% of leading edge genes (**Fig 1-7c**, dotted line). This top 6% of leading edge genes is also more likely to have a higher fold-change compared to the entire set (**Fig 1-S7a**). Together, this suggests a core set of genes most responsible for differential pathway regulation among Mps1i-treated cells.

To determine if these core genes were enriched in MN+/- leading edge genes, we plotted the contribution provided by those Mps1i+/- leading edge genes that were also leading edge genes in the MN+/- data set (**Fig 1-7c**, in pink). These genes are 3 times more likely to appear within the core set of genes (**Fig 1-7c**, odds ratio). The contribution is also somewhat greater than what would be expected by chance alone (**Fig 1-7c**, gray 66% and 95% confidence intervals). However, when we determine the contribution of Mps1i+/- leading edge genes also present in the MN+/- data set, but not in the rupture+/- data set, the contribution of these genes to pathway enrichment in Mps1i+ cells drops to what would be expected by chance alone (**Fig 1-7d**). Leading edge genes found in both the Mps1i+/- and rupture+/- data sets are represented in 23 of the 24 differentially regulated pathways and are also significantly more likely to be present among the core Mps1i+/- leading edge genes (**Fig 1-7e**). Conversely, Mps1i+/- leading edge genes not found in the rupture+/- data set are significantly more likely to be found outside of the core set, individually contribute to fewer pathways, and have a lower RNAseq fold change (**Figs 1-7f, 1-S7a**). These data suggest that the genes most important for driving the changes identified by GSEA in Mps1i-treated cells are also leading edge genes seen in rupture+ cells, despite both rupture+ and rupture- cells having been treated with Mps1i. These trends also hold when considering the much larger set of GO terms (**Figs 1-S7b**–**e**) and when comparing the contribution of rupture to the transcriptional response to micronucleation (**Figs 1-S8a**–**c**). Overrepresentation analysis of the Mps1i+/- Hallmark pathways identifies the specific pathways whose genes include a greater proportion of rupture+/- leading

edge genes than would be expected by chance alone (**1-Fig 7a**, in bold).[61,62] These rupture-overrepresented pathways include inflammatory, stress and damage response, and cell proliferation pathways. In summary, these data suggest that cells with ruptured micronuclei, despite only representing ~25% of the total population of Mps1i-treated cells (**Fig 1-S8d**), are disproportionately driving the transcriptomic response to Mps1i treatment, specifically those pathways related to stress, inflammation, and proliferation, suggesting that at least a portion of the aneuploidy response previously characterized may be a response to events correlated with MN rupture specifically.

A redesigned neural net improves overall performance in a wider range of conditions

The neural net used in the above experiments prioritized speed and PPV in specific imaging conditions. However, other experiments would benefit from an increase in recall or good performance in other imaging conditions. To generalize neural net performance to other imaging conditions, we redesigned the neural net architecture to use a simplified UNet model with no transfer learning employed. One difficulty in segmenting MN is the imbalanced nature of the data set—there are many more pixels that are not MN in an image than are. To improve training, we switched from a cross-entropy loss function to a focal loss function, which is designed to deal with highly imbalanced data sets.[63] A second challenge with micronucleus segmentation is the similarity between nucleus-associated and micronucleus-associated pixels. We modified the training data and the neural net to segment both nuclei and MN. While nucleus segmentation underperforms existing nucleus segmenters, providing this information to the micronucleus neural net improved its ability to distinguish the two. We also more than doubled the size of the cropped region to give the neural net greater contextual information, and then trained it on both RPE-1 and U2OS cells expressing H2B-emiRFP703, to broaden the neural net's ability to discern MN in different cell types with different nuclear shapes. The original neural net relied on crops centered on already segmented nuclei to reduce processing speed, at the cost of coverage. To improve coverage and predictions, at the cost of processing time, we

oversample the entire field and predictions are averaged. Final MN segments are compared with nucleus segments generated by an existing, purpose-built nucleus-segmenting neural net to trim down any MN segments that overlap with predicted nuclei (**Fig 1-8a**).[64]

When tested against RFP703/Dendra cells, this new neural net has similar recall and PPV as the neural net used for VCS, though for the whole field rather than just cropped regions. However, it maintains a similar recall and PPV under different imaging conditions and with different cell lines, with a generally improved mean intersection-over-union (mIOU) (**Fig 1-8b**). This neural net is being offered as a Python package installable with the Python package manager pip, described on the Hatch Lab GitHub repository.[43]

Materials and methods

Plasmid construction

pLVX-EF1a-NLS-3xDendra2-blast was created by PCR of NLS-3xDendra2 from pLenti-CMV-Dendra2x3-P2A-H2B-miRFP (a gift from Doug Fowler) with forward primer (with prepended NotI digest site) 5'-***ggaaggg*****gcggccgc**caagtttgtacaaaaaagttggcaccATGG-3' and reverse primer (with prepended EcoRI digest site and stop codon) 5'-***ggaaggg*****gaattcTTA**GGAAAAATTCGTTGCGCCGCTCCC-3', followed by insertion into pLVX-EF1a-blast by restriction digest cloning (NEB NotI-HF, #R3189; EcoRI-HF, #R3101). Bold signifies regions that were prepended with primer extension; bold-italic mark sites added to facilitate restriction enzyme activity and are cleaved before insertion into the final plasmid.

pLVX-EF1a-H2B-emiRFP703-neo was created by PCR of H2B-emiRFP703 from pH2B-emiRP703 (a gift from Vladislav Verkhusha, AddGene #136567) with forward primer (with prepended NotI digest site) 5'-***gaaaggg*****gcggccgc**cATGCCAGAGCCAGCGAAG-3' and reverse primer (with prepended MluI site) 5'-

***gaaaggg*gaattc**cTTAGCTCTCAAGCGCGGTGATC-3', followed by insertion into pLVX-EF1a-neo by restriction digest cloning (NEB MluI-HF, #R3198).

All primers manufactured by IDT and resuspended to 100 μM stock solution in 1xTE (0.5 M EDTA [Sigma, #E5134] and 1 M Tris Base [Fisher, #BP152] in H_2O, adjusted to pH 8.0) and diluted to a 10 μM working solution in ultrapure H_2O (Genesee Scientific, #18-195).

Lentivirus production

For pLVXE-based plasmids, HEK293T cells were plated in 10 cm tissue-culture–treated dishes (Fisher Scientific, #664160) at 70% confluency in DMEM (Gibco, #11965092) supplemented with 10% FBS (Sigma, #F0926, lot: 18C539) and 1% pen/strep (Gibco, #15140122). The next day, we replaced with fresh media and transfected cells with a solution of 3 μg psPAX2, 1.5 μg pVSV-G (gifts from Didier Trono, AddGene #12260 and #12259, respectively), and 2 μg of pLVX-EF1a-NLS-3xDendra2-blast, pLVX-EF1a-H2B-emiRFP703-neo, or pLenti-CMV-Dendra2x3-P2A-H2B-miRFP after a 5 minute incubation in OptiMEM (Gibco, #31985070) with 40 μg linear PEI (polsciences, #23966-2).

The following day, we refreshed media with media supplemented with a sterile 25 mM HEPES (Sigma, #H3375) solution in diH_2O.

The following day, we removed 5 mL of supernatant and filtered it through 0.45 μm PES syringe filters (Genesee Scientific, #25-246) and stored the filtrate at 4 °C. The remaining media was removed and replaced with fresh, HEPES-supplemented media.

The next day, we repeated this procedure, adding the filtrate to the previous day's. We then added 2 U/mL DNaseI (Fisher, #FEREN0523) and $MgCl_2$ to a 2 mM final concentration. After incubating at room temperature for 30 min, we moved the

filtrate to 4 °C for another 4 hours, then aliquoted virus and snap froze in liquid N_2. Virus aliquots were thereafter stored at -80 °C.

Cell lines

The hTERT RPE-1 NLS-3xDendra2/H2B-emiRFP703 and U2OS NLS-3xDendra2/H2B-emiRFP703 cell lines were produced through serial transduction of both plasmids. Immediately after splitting hTERT RPE-1 or U2OS cells to 40% confluency, a polybrene (Sigma, #H9268) solution in PBS (Gibco, #14190) was added to a final concentration of 6 μg/mL, followed by 1 mL of virus filtrate added dropwise. Cells were then split twice before selection was added: blasticidin (Invivogen, #ant-bl-1) to a final concentration of 10 μg/mL and Geneticin/G418 (Gibco, #11811031) to a final concentration of 500 μg/mL active G418. Cells were then sorted on by FACS on a BD Aria II for the top 20% brightest Dendra2 and emiRFP703 expressing cells.

hTERT RPE-1 NLS-3xDendra2-P2A-H2B-miRFP703 cells were created through a single transduction as described above, without selection applied, and sorted by Dendra2 and miRFP703 intensity.

Tissue culture

hTERT RPE-1 cell lines were kept in DMEM/F12 (Gibco, #11330057) supplemented with 10% FBS, 1% pen/strep, and 0.01 mg/mL hygromycin B (Sigma, #H3274) and stored in a cell incubator (HeraCell 150i) with a 5% CO_2 atmosphere at 37 °C.

RPE-1 NLS-3xDendra2/H2B-emiRFP703 were supplemented with 10 μg/mL blasticidin and 500 μg/mL G418.

U2OS NLS-3xDendra2/H2B-emiRFP703 cells were kept in DMEM supplemented with 10% FBS and 1% pen/strep, and stored in a cell incubator with a 10% CO2 atmosphere at 37 °C and likewise supplemented with blasticidin and G418.

Micronucleus segmentation and cell classifier

VCS neural net

The neural net used for the experiments in this paper was created using the FastAI 1.0 library in Python. It is a U-Net with Torchvision's ResNet18 pre-trained model as its base architecture.[65] Additional training was performed specifically for micronucleus recognition using ~2,000 48x48 crops of individual cells as training data, a further 164 for validation, and 177 for testing.

Training data was generated using the hTERT RPE-1 NLS-3xDendra2-P2A-H2B-miRFP703 cell line. Training data was generated by treating cells with either 0.5 μM reversine or DMSO loading control, then imaging the next day on a Leica DMi8 at the Fowler Lab (Fowler Lab microscope) using a 20x widefield objective in the NIR and GFP channels. Following nucleus segmentation using the Deep Retina neural net,[48] 48x48 px crops of the H2B channel were generated centered at each nucleus. We then annotated each crop for the presence of micronuclei using a custom program.

The classifier takes as input a 20x image with the H2B-emiRFP703 and NLS-3xDendra2 channels. It uses the Deep Retina neural net to segment nuclei, discarding any touching the edge of the image. It then generates 48x48 px crops centered at each nucleus. Each crop is processed with Sobel edge detection. The VCS neural net architecture is based on Torchvision's ResNet18, which expects a 3-channel image. To accommodate this architecture, we transform each crop into a 3-channel image: the first 2 channels are duplicates of the H2B channel and the final channel the results of Sobel edge detection.

Crops are processed by the VCS neural net. Segmented micronuclei are localized on the original, full 20x image, then assigned to the closest nucleus segment. Any micronucleus segments whose nearest segmented nucleus is more than 40 px away are discarded.

Once micronuclei are assigned, the classifier takes the maximum Dendra2 intensity from each micronucleus segment and its associated nucleus segment to generate micronucleus/nucleus Dendra2 ratios. Micronuclei with a ratio below 0.16 are classified as being ruptured micronuclei. This threshold was chosen by manually identifying micronuclei as ruptured or intact from experiments isolating MN+ and MN- cells, then using the JRip classifier in Weka 3.8.6 to identify an optimal threshold to separate intact and ruptured micronuclei by this Dendra2 ratio.[66,67]

Nuclear segments are classified as MN+ or MN- based on the presence or absence of an associated micronucleus segment. MN+ segments are then further classified into those with only intact micronuclei (rupture-) or those with at least one ruptured micronucleus (rupture+).

Images taken and the classification results from each experiment were saved. After an experiment finished, we compared classifier predictions against ground truth by manual inspection. Micronuclei with attached chromatin bridges were ignored to ensure that micronuclei were discrete nuclear compartments.

Individual micronuclei were segmented by hand using PixelStudio on an Apple iPad. Recall was calculated as the proportion of all micronuclei that overlapped with a predicted segment. Positive predictive value was calculated as the proportion of all predicted segments that overlapped with a micronucleus.

Generalized neural net

The generalized neural net was created using TensorFlow 2.0 also using a U-Net architecture, but without transfer learning. Training was performed on 128x128 px crops generated at time of training from images of hTERT RPE-1 NLS-3xDendra2/H2B-emiRFP703 and U2OS NLS-3xDendra2/H2B-emiRFP703 cells taken using a 20x widefield objective on the Fowler Lab microscope in the NIR

channel. Annotations of nuclei and micronuclei were performed by hand using PixelStudio version 4.5 on an Apple iPad.

Input is expected to be an image of nuclei taken at 20x; images taken at 40x were first scaled down by a factor of 2. A 128x128 px sliding window is used to generate crops, advancing the window by 64 px horizontally and vertically to oversample the image. Crops are expanded into 2-channel images, with the second channel the result of Sobel edge detection. Crops are processed by the neural net to segment micronuclei and the average value taken from oversampled regions.

The same image is used as input for the Stardist nucleus segmenter to identify nuclei. Any regions where micronucleus segments overlap nucleus segments are removed to produce a final set of micronucleus labels.[64]

The updated neural net was validated against images of hTERT RPE-1, U2OS, hTERT HFF, and HeLa cells. Live RPE-1 and U2OS cells expressing NLS-3xDendra2/H2B-emiRFP703 were taken on the Fowler Lab microscope with a 20x widefield objective in phenol-free media. Fixed cells were taken on either a Leica DMi8 (Hatch Lab microscope) or Leica spinning disk microscope with a 40x confocal objective after being fixed in 4% PFA (Electron Microscopy Sciences, #15710) in PBS for 5 min at room temperature with coverslips mounted on glass slides with VectaShield anti-fade medium (Vector Laboratories, #H-1000) and sealed with nail polish (Wet *n* Wild, wildshine base coat).

HeLa cell images were taken by Molly Zych and HFF cell images were taken Emily Hatch, both at the Hatch Lab.

Recall and positive predictive values were calculated using the same protocol used for VCS classifier validation.

VCS CellTrace sorting accuracy validation

We trypsinized hTERT RPE-1 NLS-3xDendra2/H2B-emiRFP703 and hTERT RPE-1 cells with 0.25% trypsin in EDTA (Gibco, #25200-056) after rinsing cells in PBS. After pelleting cells at 300 rcf for 5 min, we resuspended each population of cells in 1 mL PBS. We then made 2 aliquots of 1 million cells for each population. Cells were then re-pelleted and resuspended in 1 mL PBS. To one aliquot from each population, we added 1 μL CellTrace far red (ThermoFisher, #C34572). To the other aliquots, we added 1 μL DMSO (US Biological, #D8050-01).

Aliquots were incubated at 37 °C for 10 min in the dark, gently mixed by pipette, then incubated for a further 10 min. We then added 200 μL PBS to each aliquot, mixing gently by pipette and incubated for another 5 min.

Aliquots were pelleted as before, then resuspended in media. Cells were then plated on a 6-well glass-bottom dish (Cellvis, #P06-1.5H-N), 175k cells/well, as follows: A1: RPE-1, -CellTrace; A2: RPE-1, +CellTrace; B1: Dendra2-expressing, -CellTrace; B2: Dendra2-expressing, +CellTrace; A3: 1:1 mix of Dendra2-expressing +/- CellTrace.

Cells were then transported from the Hatch Lab at the Fred Hutchinson Cancer Center to the Fowler Lab at the University of Washington (~20 min transport) in the dark and placed in a cell incubator (HeraCell Vios 160i) with a 5% CO_2 atmosphere at 37 °C.

The following day, we rinsed cells 3x with PBS then replaced media with phenol-free DMEM/F12 (Gibco, #21041025) supplemented with 10% FBS, 1% pen/strep, 0.01 mg/mL hygromycin B, and 10 μM Cdk1i (RO-3306, Sigma, #SML0569) that had been filtered through a 0.22 μm PES syringe filter (Genesee Scientific, #25-244) to remove residual RO-3306 precipitate. Cells were then placed in a temperature-controlled, CO_2-supplemented microscope incubation chamber.

1 hour following treatment with Cdk1i, the mixture well was imaged by scanning across the well using a 20x widefield objective on the Fowler Lab microscope, imaging in NIR and GFP wavelengths. The NIR channel captures both H2B-emiRFP703 and CellTrace far red, though the latter is significantly brighter. Custom journals transmitted each image to another computer equipped with a CellTrace classifier.

This classifier segmented nuclei with the Deep Retina neural net using the GFP channel as input, then measured the mean NIR intensity in the nucleus. Cells above a given threshold, determined on the day by capturing a test image and measuring the CellTrace intensity of CellTrace+ and CellTrace- cells by hand, were classified as CellTrace+ and those below this threshold were classified as CellTrace-. 1-bit masks of each population were transmitted back to MetaMorph, which used a digital micromirror device to target CellTrace+ cells with 800 ms UV and CellTrace- cells with 200 ms UV.

Following photoconversion of the mixture well, cells were trypsinized and resuspended in media. We then pelleted cells at 300 rcf for 5 min, then resuspended cells in FACS buffer (1 % BSA [Sigma, #A4503, lot: SLCG9729], 0.2 mM EDTA, 25 mM HEPES in PBS, sterile filtered with 0.22 μm PES syringe filter) and placed on ice.

We then sorted cells with a BD Aria III, using Alexa 700 to gate on emiRFP703+ cells and FITC to gate on Dendra2-hi cells. Final sorting was performed by gating on a ratio of PE-blue/FITC. Compensation for PE-blue exciting unphotoconverted Dendra2 was performed using Dendra2-expressing cells without CellTrace, as was compensation for CellTrace far red and emiRFP703 spectral overlap. The sorted populations were then re-analyzed on the same sorter to validate the purity of each sorted population.

Mps1i+/- experiment and VCS validation

We plated 175k hTERT RPE-1 NLS-3xDendra2/H2B-emiRFP703 cells in 3 wells of a glass-bottom 6-well dish. The following day, media was replaced and supplemented with 100 nM Mps1i (BAY-1217389, Fisher, #501872752) or DMSO loading control.

Cells were then transported in the dark from labs at the Fred Hutchinson Cancer Center to the Fowler Lab at the University of Washington, as above. 19 hours later, we rinsed cells 3x with PBS and then replaced media with filtered phenol-free media supplemented with 10 μM Cdk1i. Cells were then placed in a temperature-controlled, CO_2-supplemented microscope incubation chamber.

1 hour following treatment with Cdk1i, two wells were imaged sequentially with a 20x widefield objective in the NIR and GFP channels. As previously described, MetaMorph scanned over each well, and custom journals transmitted each image to another computer equipped with a random classifier. 1-bit masks of nuclei randomly assigned to receive 800 ms or 200 ms UV pulses were transmitted back to MetaMorph, which directed UV pulses at the segmented nuclei.

At the end of each well, the first and last few positions were re-imaged in NIR, GFP, and RFP channels to ensure cells were being properly photoconverted. At the end of the experiment, several positions in the third, un-photoconverted well was also imaged.

Following Dendra2 photoconversion of two wells, cells were trypsinized, resuspended in media, and pelleted. Cells were then resuspended in FACS buffer and placed on ice.

We sorted cells with a BD Aria III, using Alexa 700 to gate on emiRFP703+ cells and FITC to gate on Dendra2-hi cells. Final sorting was performed by gating on a ratio of PE-blue/FITC. Compensation for PE-blue exciting unphotoconverted Dendra2 was performed using the final well not treated with UV.

After collecting ~13k cells in a given condition, sorting was paused and collection tubes swapped out in order to collect further samples of ~13k cells. Samples were then pelleted, media aspirated, and pellets flash frozen in dry ice before being placed at -80 °C for RNA extraction.

To validate the maintenance of a Dendra ratio over the course of the experiment, we thresholded images taken at the end of the experiment using the GFP channel to segment nuclei. We used these segments to measure the mean GFP and RFP channel intensities. Images taken of the last few positions represented cells that had just been photoconverted, while images taken of the first few positions represented cells that had been photoconverted 4 or 8 hours prior. Plotting the ratio of RFP/GFP intensities revealed three distinct populations in all three image sets, representing unphotoconverted cells, cells that received 200 ms UV, and cells that received 800 ms UV.

Micronucleus+/- isolation

This followed the same protocol used for the Mps1i+/- experiment, but with all cells receiving 100 nM Mps1i. As above, two wells were imaged sequentially with a 20x widefield objective in the NIR and GFP channels. 1-bit masks of MN+ nuclei and MN- nuclei were transmitted back to MetaMorph, which directed UV pulses at the segmented nuclei. In one well, MN+ nuclei received 800 ms UV and MN- nuclei 200 mn UV; in the other well, this was reversed. Whether the MN+:800 ms/MN-:200 ms well or the MN+:200 ms/MN-: 800 ms was imaged first was flipped from one experiment to the next to ensure UV exposure could be controlled for during downstream analyses.

Sorting was performed as described above. During one experiment, one aliquot was retained for replating instead of being flash frozen, as described below.

MN+/- sorting accuracy determination

To assess sorting accuracy, one aliquot was transported back to the Hatch Lab on ice in the dark (travel time ~20 min) where cells were pelleted and resuspended in media supplemented with 10 μM Cdk1i to prevent mitosis, then replated into 24-well plates (ThermoFisher, #930186) on lysine-coated coverslips. The following day, we rinsed cells in PBS then fixed in 4% PFA in PBS at room temperature for 5 min. Cells were

permeabilized in PBS-BT (PBS supplemented with 0.4% TritonX-100 [Sigma, #T9284], 3% BSA, and 0.02% sodium azide [Sigma, #S2002]) for 15 min, then incubated in 1 μg/mL DAPI [Invitrogen, #D1306] in PBS for 5 min. After a final rinse in PBS, we mounted coverslips on glass slides with VectaShield and sealed with nail polish.

We then imaged coverslips on the Hatch Lab microscope with a 40x confocal objective, taking z-stacks of multiple positions until ~100 cells of each population were imaged. We then randomized the names of images, to obscure which image came from which population, then counted the number of MN+ cells in each image. Once counting was finished, we re-identified which images belonged to which population.

Rupture frequency pulse-chase

RPE-1 NLS-3xDendra2/H2B-emiRFP703 cells were plated into 2 wells of a 6-well glass-bottom dish. One well was plated at 225k cells/well supplemented with 1 μM Cdk4/6i (PD-0332991, Sigma, #PZ0199) to synchronize cells at the G1/S transition. The other well was plated at 175k cells with DMSO loading control.

24 hours later, we rinsed wells 3x with PBS before replacing media supplemented with 100 nM Mps1i. Cells were transported to the Fowler Lab as described above. 19 hours later, we rinsed cells 3x with PBS and replaced media with phenol-free media supplemented with 10 μM Cdk1i that had been filtered to remove precipitate.

After incubating cells in the microscope incubation chamber for 1 hour, we imaged 5 positions of each well, running the micronucleus classifier as described above. However, we restricted photoconversion to cells that were classified as rupture-, exposing them to 800 ms UV pulses. No cells were treated with 200 ms UV.

After photoconversion, every hour for 11 hours, we imaged a 25-position window centered at the original 5 photoconverted positions in the NIR, GFP, and RFP

channels to track those cells that had been photoconverted. Images were then analyzed to validate by eye the proportion of cells that were rupture-, rupture+, or MN-.

Rupture+/- isolation

Cells were plated as in the micronucleus+/- section, but only 2 wells were plated at 225k cells/well. In addition, cells were originally plated in media supplemented with 1 μM Cdk4/6i (PD-0332991, Sigma, #PZ0199) to synchronize cells.

24 hours later, we rinsed cells 3x with PBS and replaced with media supplemented with 100 nM Mps1i and then transported to the Fowler Lab.

The following day, cells were prepared as described above with Cdk1i, phenol-free media and one well was photoconverted, with 1-bit masks for rupture- and rupture+ cells being returned by the classifier. Rupture- cells received 800 ms UV and rupture+ cells 200 ms. Cells were then sorted as described in the micronucleus+/- isolation section.

RNA isolation and sequencing

We extracted RNA from frozen pellets using ThermoFishers's RNAqueous micro kit (ThermoFisher, #AM1931), starting with immediately resuspending frozen pellets in lysis solution and vortexing vigorously. All other steps were according to the manufacturer's protocol. Residual DNA was removed with the optional DNase I treatment that is included with the kit, resulting in ~20 μL of RNA solution.

To remove residual guanidium salts, RNA was further purified by glycogen precipitation. 20 μL pH 7 TE was added, then 4 μL of 3 M sodium acetate (Fisher, #S210) and 0.4 μL 15 mg/mL GlycoBlue (ThermoFisher, #AM9515) or 0.3 μL 20 mg/mL RNA-grade glycogen (ThermoFisher, #R0551). 133 μL of -20 °C 100% ethanol was then added and the mixture placed at -20 °C overnight. The next day, the sample

was centrifuged at 14,000 rcf for 10 min at 4 °C in a microcentrifuge and then supernatant removed by pipette. The glycogen/RNA pellet was then washed with 500 µL of -20 °C 70% ethanol in ultra-pure H_2O and pelleted again at 21,000 rcf for 10 min at 4 °C. The supernatant was removed by pipette and the tube allowed to dry with the cap open at room temperature. Finally, the RNA was resuspended in ultra-pure H_2O heated to 65 °C. RNA quality and concentration was checked by the Genomics Core at the Fred Hutchinson Cancer Center with an Agilent 4200 Tapestation HighSense RNA assay.

cDNA synthesis and library preparations were performed by the Genomics Core at the Fred Hutchinson Cancer Center using the SMARTv4 for ultra-low RNA input and Nextera XT kits, respectively (Takara). Sequencing was also performed by the Genomics Core using an Illumina NextSeq 2000 sequencing system with paired-end, 50 bp reads.

RNAseq analysis

We quantified transcripts with Salmon to map reads against the UCSC hg38 assembly provided by http://refgenomes.databio.org (digest: `2230c535660fb4774114bfa966a62f823fdb6d21acf138d4`), using bootstrapped abundance estimates and corrections for GC bias.[68]

For comparisons with data from He, *et al.* and Santaguida, *et al.*, the original FASTA files deposited at the Sequence Read Archive were downloaded with NCBI's SRA Toolkit and quantified with Salmon.[44,45] No GC-bias correction was applied, as this is still experimental at time of writing for single-end reads.

Resulting transcript abundances were then processed to find differentially expressed genes (DE genes) with the R package DESeq2 version 3.16 in R 4.2.1, RStudio 2022.07.2 build 576, and Sublime Text build 4143. Briefly, files were imported into DESeq2 with the R package tximeta.[69,70] Estimated transcript counts were summarized to gene-

level and low-abundance genes were filtered by keeping only those genes with estimated counts ≥ 700 in at least 2 samples.

DEGs were identified using a likelihood ratio test comparing the full model with one with the condition of interest dropped and an FDR of 0.05. Log-fold changes were corrected using empirical Bayes adaptive shrinkage.[71]

Gene-set enrichment analysis

GSEA was performed using the R package fgsea version 1.25.1, comparing log-fold changes of genes against either the full *Homo sapiens* Gene Ontology database (GO) version 2022.1 or the Hallmark Gene Sets version 2022.1, both part of the MSigDB resource published by UC San Diego and the Broad Institute.[72–79] Santaguida, *et al.*, provides data from cells that underwent cell-cycle arrest and cells that continued to cycle; our GSEA found no significantly enriched terms from still-cycling cells.

We generated a unique list of leading edge genes for each data set and the number of pathways or terms that leading edge gene was a member of. These genes were then divided into core genes and non-core genes, with the former being the unique list of genes that were members of at least 99% of pathways (for the Hallmark ontology) or 98% of terms (for the GO ontology).

To compare the contributions of different subsets of leading edge genes, we generated an odds ratio by comparing the ratio of core/non-core genes in the subset vs the total leading edge set. To test if the subset's contribution was more or less than what was expected by chance, we generated a confidence interval from a null distribution created by 200 bootstrapped subsets whose overlap with the total leading edge gene list was the same proportion as the subset of interest.

To identify Mps1i+/- pathways whose genes include more rupture+/- leading edge genes than would be expected, we used the R package clusterProfiler version 4.6.2

XCE FISH and analysis

We plated 175k hTERT RPE-1 cells/well in a 6-well dish (ThermoFisher, #130184) onto lysine-coated coverslips. The following day, we replaced media with fresh media supplemented with 100 nM Mps1i. 20 hours following, we rinsed cells in PBS then fixed in -20 ºC 100% methanol for 10 min.

To determine the rupture status of micronuclei, we performed immunofluorescence for acetylated lysine 27 on histone 3 (H3K27-acetyl).[15,16,27] Briefly, we rehydrated cells with a 2x wash in PBS for 5 min each at room temperature, then blocked in PBS-BT for 30 min at room temperature. We then incubated coverslips in polyclonal rb α H3K27-acetyl antibody (Abcam, #ab4729) at 2 μg/mL in PBS for 30 min at room temperature followed by a 3x PBS wash for 5 min each at room temperature. We then incubated coverslips in a secondary gt α rb antibody conjugated to AlexaFluor-488 (Invitrogen, #A11034) at 1 μg/mL in PBS-BT for 30 min at room temperature. After a final 3x PBS wash for 5 min each at room temperature, we re-fixed cells in 4% PFA in PBS for 5 min at room temperature.

We then rinsed coverslips in 2xSSC (20x SSC, Sigma, #S6639 diluted in ultrapure H_2O), then permeabilized cells in a solution of 0.2 M HCl (Sigma, #258148) and 0.7% TritonX-100 in ultrapure H_2O for 15 min at room temperature. After 2, 5-min washes in 2xSSC at room temperature, we incubated coverslips in a 1:1 solution of formamide (Millipore, #4650) and 2xSSC for 1 hour at room temperature. After another 2, 5-min washes in 2xSSC at room temperature, we applied 3 μL of chr 1, 11, or 18 XCE orange probe (MetaSystems, #D-0801-050-OR, D-0818-050-OR, and D-0811-050-OR) directly to the coverslip and mounted it on a glass slide, sealed with rubber cement. We immediately placed the slide on a 74 ºC hot plate for 3 min, then moved the slide to a pre-warmed, humidified incubation chamber at 37 ºC. We incubated each slide at 37 ºC for either 4 hours (chr 18) or 14.5 hours (chrs 1 and 11).

Following incubation, we removed rubber cement and placed the coverslip into a 0.4xSSC solution heated to 74 ºC, and incubated each coverslip at 74 ºC for 5 min. We then performed 2, 5-min washes in 2xSSC supplemented with 0.1% Tween20 (Fisher, #BP337). Finally, we incubated the coverslip in 1 μg DAPI in PBS supplemented with 0.4% TritonX-100 for 5 min at room temperature, rinsed once in PBS, then mounted each coverslip onto a glass slide using VectaShield anti-fade medium, sealed with nail polish.

Images were taken on the Hatch Lab microscope using a 40x confocal objective. For each cell, the number of centromeres in the cell, the number of centromeres in micronuclei, and the number of centromeres in ruptured micronuclei were counted. For determining missegregation frequencies, cells without 2 centromeres for a given chromosome were classified as missegregated. For determining functional aneuploidy, the cells without 2 centromeres in any intact nuclear compartment (ie, centromeres in ruptured micronuclei did not count toward the total) were classified as aneuploid.

Micronucleation status was performed by hand. Rupture status of any given micronucleus was determined by hand through examining the presence or absence of H3K27-acetyl.

In total, 3 replicates of chromosomes 11 and 18 and 2 replicates of chromosome 1 were performed.

Immunofluorescence

For ATF3 validation, we plated 175k RPE-1 2xRFP-NLS cells in 4 wells of a 6-well dish along with lysine-coated coverslips. The following day, we changed the media in all wells, replacing one well's media with that supplemented with 100 nM Mps1i and the others with DMSO loading control. 23 hours later, we replaced media in all wells following a 3x rinse with PBS at room temperature with fresh media supplemented with 2 μg/mL doxorubicin hydrochloride (Fisher, #BP25165) or DMSO loading

control. The resulting treatment conditions were: DMSO ⇒ DMSO, Mps1i ⇒ DMSO, or DMSO ⇒ doxorubicin. Following a 1 hour incubation, we rinsed cells 1x in PBS and fixed them in a 4% PFA solution in PBS at room temperature in the dark for 5 min.

Immunofluorescence proceeded in the dark at room temperature, starting with a 30 min block in PBS-BT. Coverslips were then incubated in mouse-anti-γH2AX (BioLegend, #613401) at a dilution of 1:499 in PBS for 30 min, followed by 3x 5-min washes in PBS. Coverslips were incubated in a secondary antibody, goat-anti-mouse conjugated to Alexa647 (Life Technologies, #A21236) diluted 1:999 in PBS-BT for 30 min. After 3x 5-min PBS washes, we incubated coverslips in rabbit-anti-ATF3 (Cell Signaling Technology, #18665S) at a dilution of 1:399 in PBS for 30 min, followed by 3x 5-min washes in PBS and incubation in goat-anti-rabbit conjugated to Alexa488 (Life Technologies, #A11034) diluted 1:1999 in PBS-BT for 30 min. After a final set of 3x 5-min PBS washes, coverslips were briefly incubated in 1 μg/mL DAPI in PBS followed by a final rinse in PBS and mounting to slides using VectaShield and nail polish.

ATF3 quantification

ATF3-stained coverslips were imaged as z-stacks on the Hatch Lab microscope using a 40x confocal objective. Following flat-field correction, images were subject to rolling-ball background subtraction in FIJI 2.9.0 with a pixel radius of 60 px. Using FIJI, in the middle of each z-stack, we thresholded on the RFP-NLS signal to create ROIs. For each ROI, we identified the middle of the nucleus using the RFP-NLS channel, then measured the mean intensities in both the ATF3 and γH2AX channels using the ROI as a mask. In addition, to determine the number of γH2AX foci, we ran the Find maxima... command in FIJI, with a threshold of 20. Measuring the raw integrated density and dividing by 255 rendered an approximate number of foci.

To aid in the visual comparison of ATF3 intensities across replicates, the values displayed in Fig 6b were normalized by first scaling all ATF3 mean intensities to the minimum and maximum values within that replicate, then translating the values

such that the 90th percentile in the DMSO-treated condition was at 1. These normalized values were only to aid in visual inspection and were not used in any statistical analysis.

Statistical analyses

Shorthand p-values are as follows:

ns: p-value >= 0.05

*: p-value < 0.05

**: p-value < 0.01

***: p-value < 0.001

****: p-value < 0.0001

The correlation between MN rupture status and segmenter recall in Figs 1c–d was determined using a generalized estimating equation (GEE) using a binomial distribution and a logit link function:[80]

(# recalled, # missed) ~ MN status

where *MN status* is whether an MN was ruptured or not. The change in rupture frequency in Fig 5e was likewise analyzed with a GEE using a binomial distribution and a logit link function.

Chromosome missegregation and functional aneuploidy results were tested with a GEE using a binomial distribution and a logit link function:

(# aneuploid, # normal) ~ Status × Chr

Where *Status* is whether the cell was MN+/- (Figs 5i, S5f) or Rupture+/- (Figs 5j, S5g) and *Chr* is chromosome identity. p-values for each individual property was calculated using the drop1 function in R.

The predicted change to classifier PPV in Fig 1-5e was determined by taking the difference in mean rupture frequencies between the beginning and end of the experiments and adjusting the PPV of rupture- and rupture+ populations by reducing

the true positive rate in the rupture- population by this difference and increasing the true positive rate in the rupture+ population by the same:

To determine the effect of rupture on the measured intensity of ATF3 in Fig 1-6c, we constructed a GEE using a Gamma distribution and log link function:

ATF3 intensity ~ Population x γH2AX

with γH2AX values first scaled between 0–1. p-values were calculated as above.

Odds ratios in Figs 1-7 and 1-S7 were tested by Barnard's exact test.

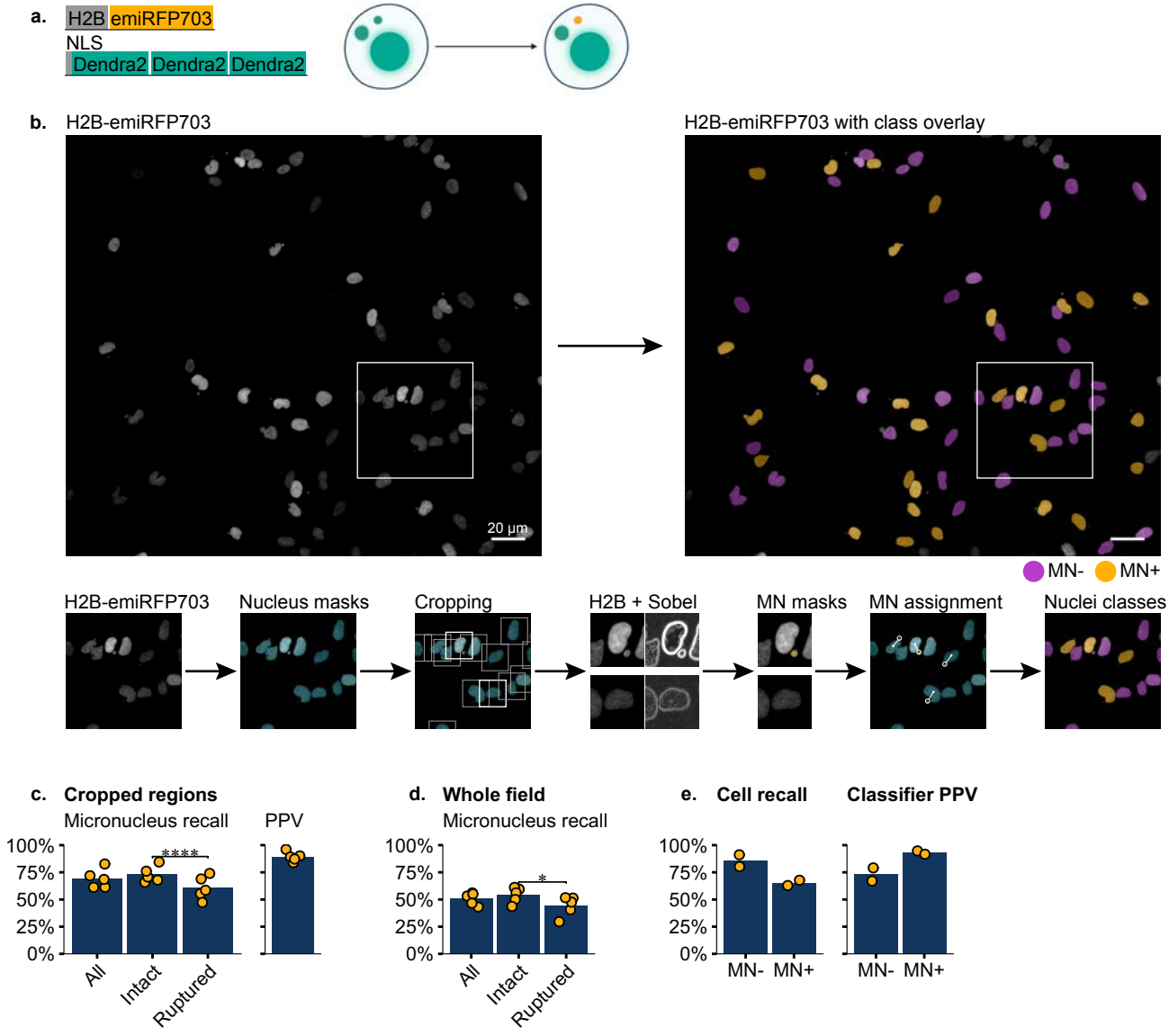

Figure 1: Machine vision identifies micronucleated cells
a) Diagrams of constructs transduced into RFP703/Dendra cells and how these constructs localize in micronucleated cells before and following micronucleus rupture. Graphic created with BioRender.com. **b)** Image of H2B-emiRFP703 in RFP703/Dendra cells at 20x and overlaid with MN+ and MN- classification results, with a visual depiction of the entire classification. **c,d)** Recall and positive predictive value (PPV) of MN within regions analyzed by the neural net and recall for the entire field of view. MN were manually classified as intact or ruptured by Dendra2 intensity. N=5, n=264, 158, 365, 283, 249. **e)** Recall and PPV of MN- and MN+ cells. N=2, n=328, 186.

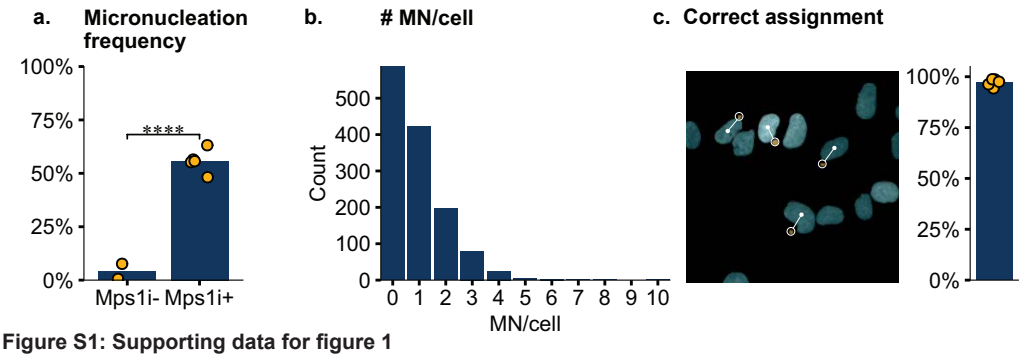

Figure S1: Supporting data for figure 1
a) Micronucleation rates in hTERT RPE-1s treated with either DMSO (Mps1i-) or the Mps1-inhibitor BAY-1217389 (Mps1i+) for 20 hrs. **b)** Histogram of the number of MN/cell in RFP703/Dendra cells after treatment with Mps1i for 20 hrs. **c)** The proportion of MN assigned to the correct nucleus by proximity alone.

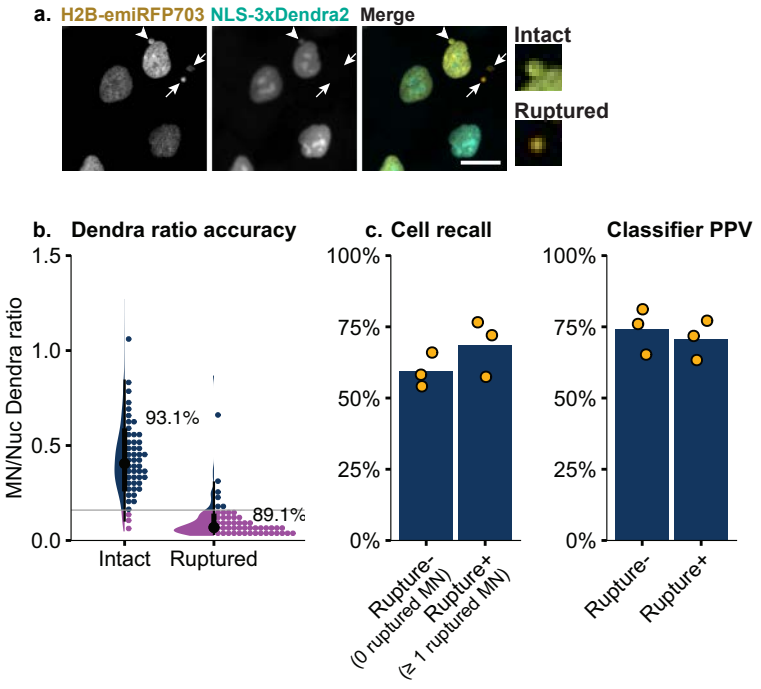

Figure 2: Micronucleated cells can be further classified into rupture+ and rupture- cells

a) RFP703/Dendra cells with ruptured (arrows) and intact (arrowhead) MN showing a loss of NLS-3xDendra2 signal in ruptured MN. **b)** Distributions of MN/nucleus Dendra2 intensity ratios between manually classified intact or ruptured MN. Solid gray line = threshold used to automatically classify intact and ruptured MN. N=3, n=179, 113, 105. **c)** Recall and PPV for rupture-based cell classification. N=3, n=120, 91, 82.

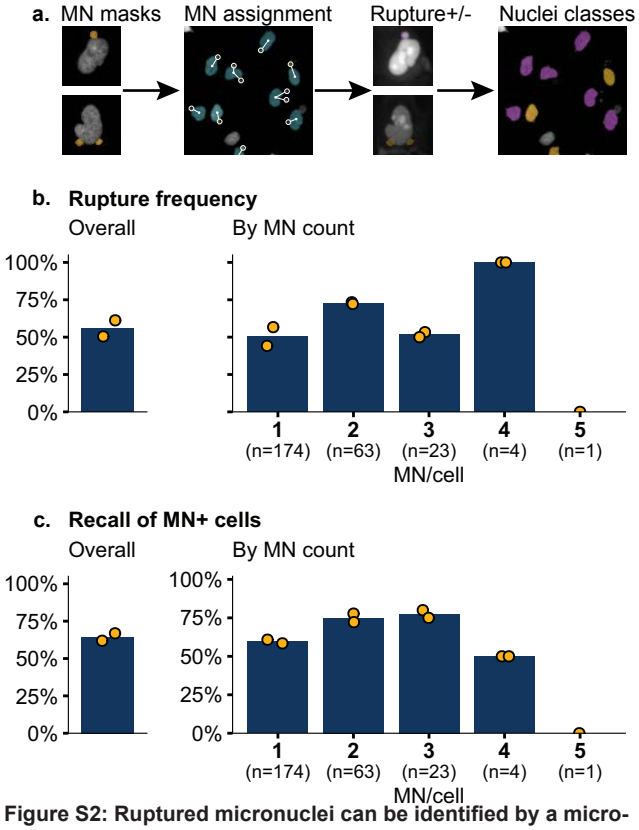

Figure S2: Ruptured micronuclei can be identified by a micronucleus/nucleus Dendra2 ratio

a) Modification to classification pipeline to include MN rupture status. **b,c)** Rupture frequency and recall of MN+ cells by # MN. Cells manually classified. N=2, n=328, 186.

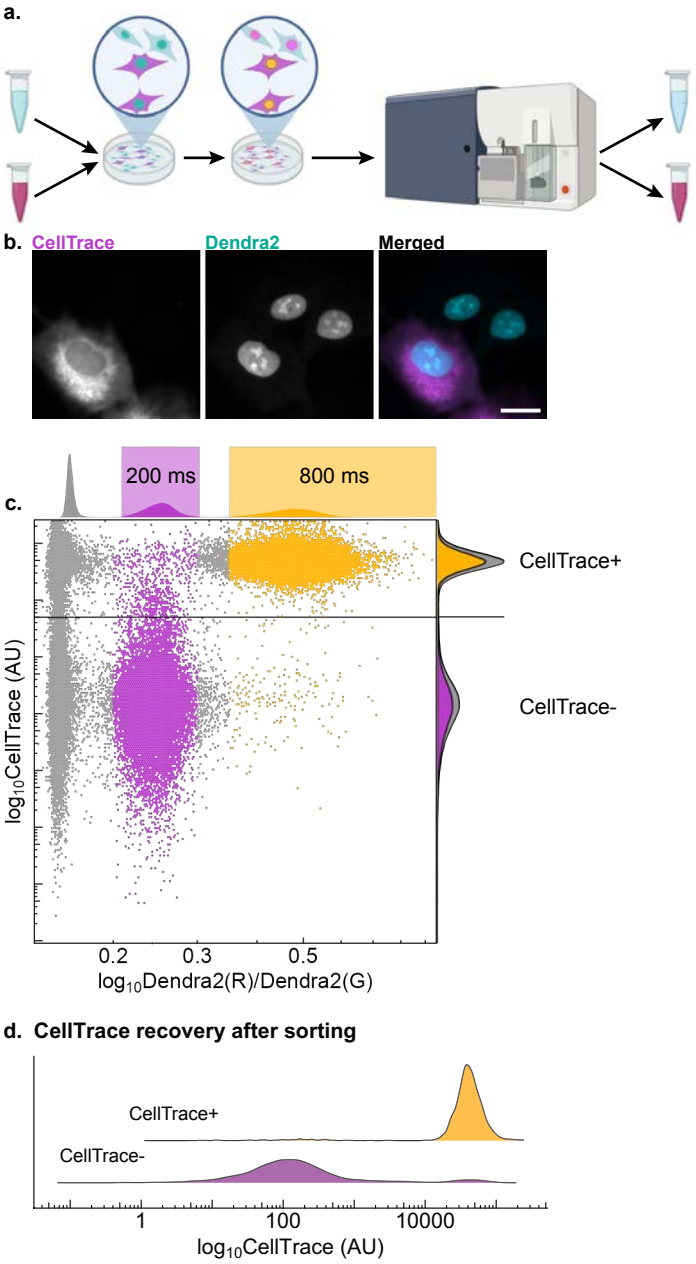

Figure 3: VCS can specifically isolate multiple cell populations from RFP703/Dendra cells
a) Overview of CellTrace mixture experiment: 1:1 mix of CellTrace+:CellTrace- RFP703/Dendra cells were differentially photoconverted and sorted by Dendra ratio to recover the original unmixed populations. Graphic created with BioRender.com. **b)** Representative image of unphotoconverted cells from CellTrace mixture experiment. **c)** FACS plot of cell population after differential photoconversion of CellTrace+ and CellTrace- cells with gates used to sort populations by UV exposure along the top margin and CellTrace distributions on the right margin. **d)** Re-analysis of CellTrace intensity of the two recovered populations after FACS.

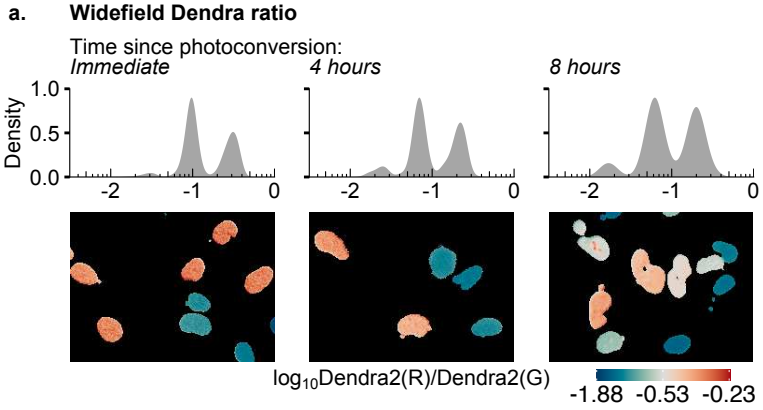

Figure S3: Dendra ratios persist for many hours
a) The Dendra-red/Dendra-green ratios from images taken post-photoconversion, pre-FACS during the experiment in Fig 4 showing 3 discernible populations even after 8 hours. Representative images from each time point are color-coded by Dendra ratio.

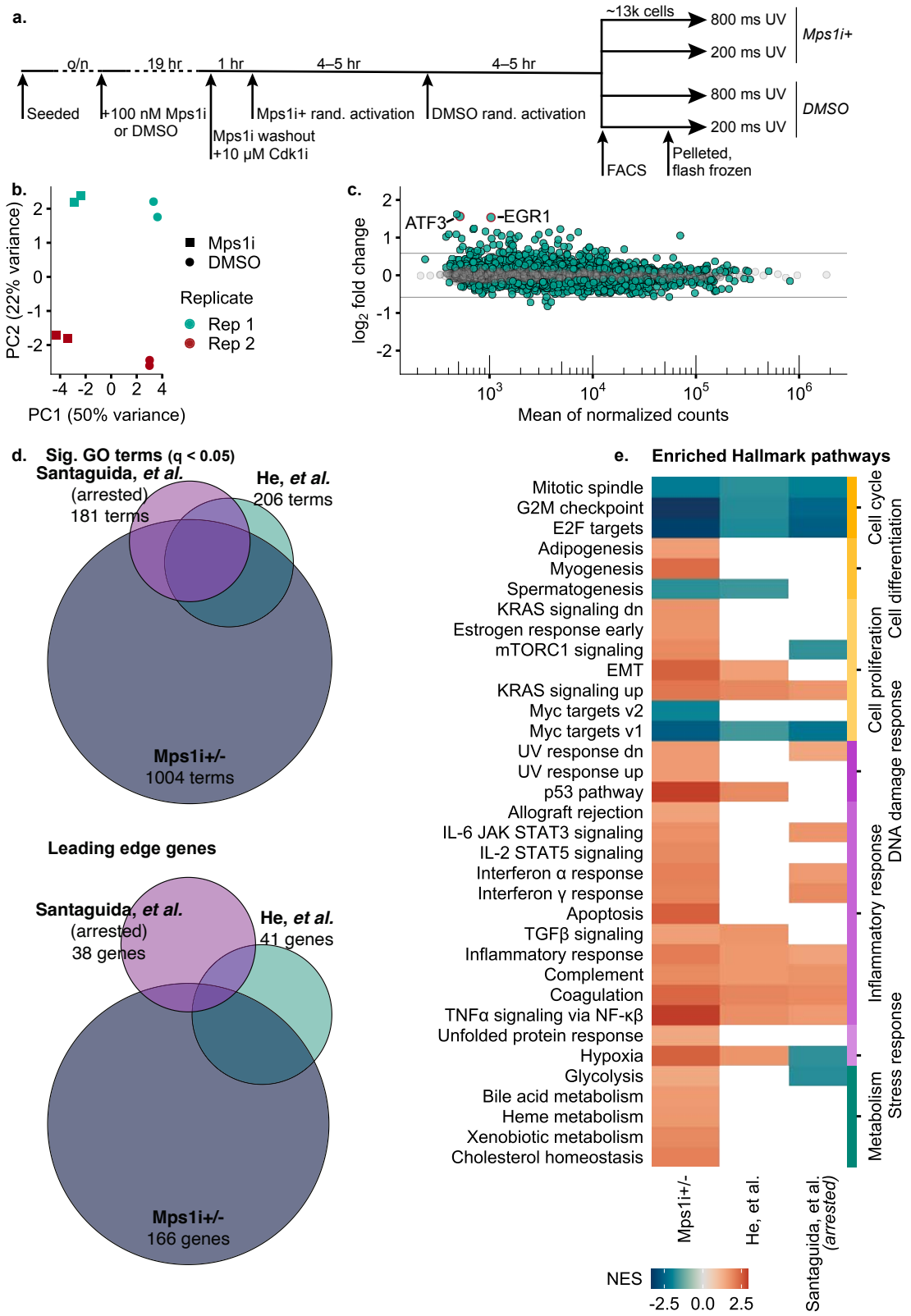

Figure 4: Spindle assembly checkpoint inhibition recapitulates characteristic aneuploidy transcriptional response

a) Timeline of experiment. **b)** PCA plot showing clustering of Mps1i-treated and DMSO-treated cells, along with minor clustering along the replicate axis. Each experimental replicate includes 2 technical replicates. **c)** MA plot. Genes with a differential expression q-value < 0.05 are highlighted in green. Gray lines represent 1.5 fold-change in expression. **d)** Venn diagram of the overlap in GO terms and leading edge genes between our data and the reanalyses of data from Santaguida, *et al.* and He, *et al.* **e)** Heatmap of Hallmark pathway enrichment between the 3 data sets in d). Hallmark pathways were grouped into broad categories on the right.

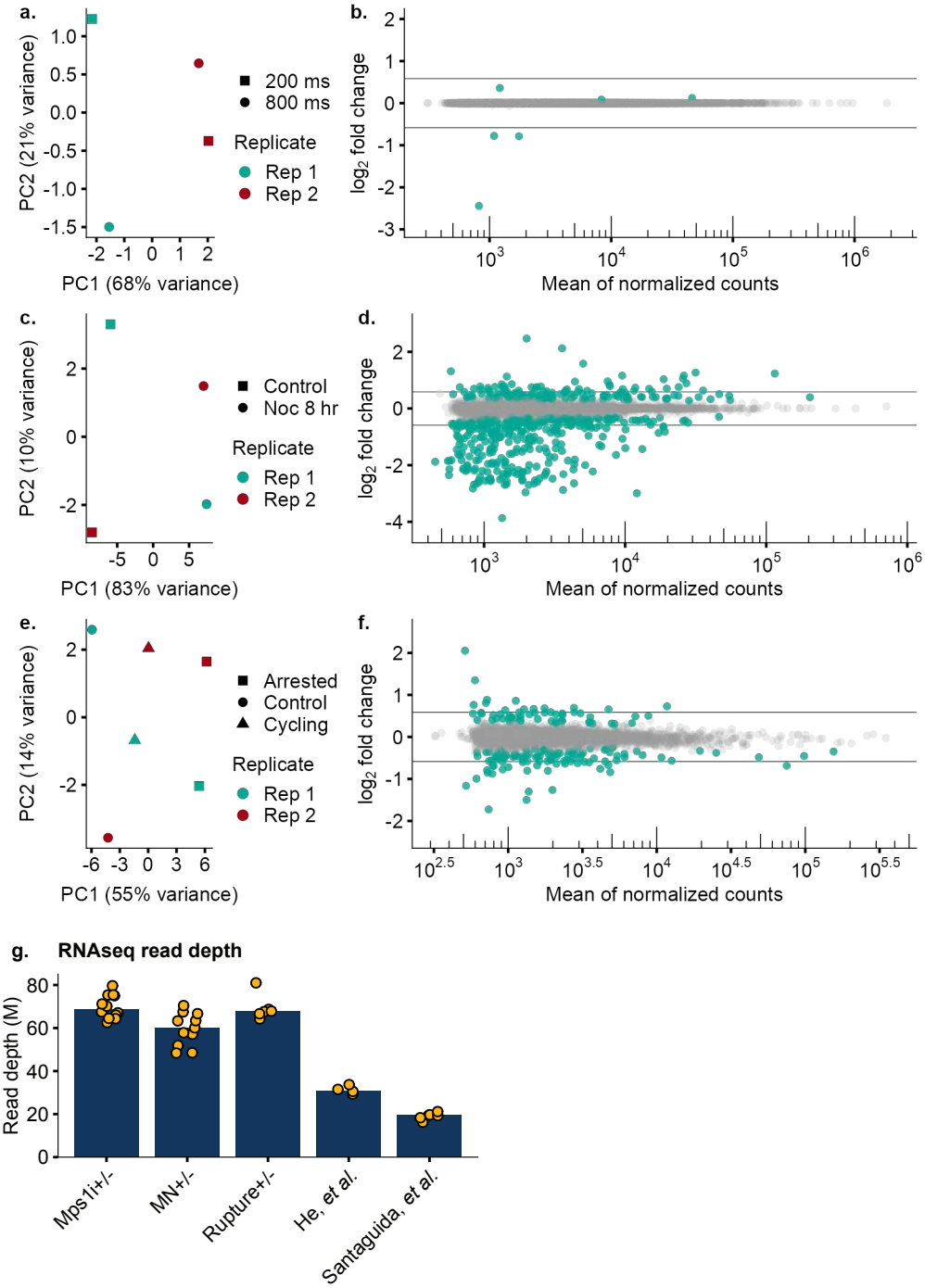

Figure S4: Controls and comparisons of differential expression analysis in Fig 4
a) PCA plot of cells treated with DMSO comparing those exposed to 800 ms UV vs 200 ms. **b)** MA plot of the data in a). Only 6 differentially expressed genes were identified in cells exposed to 800 ms vs 200 ms UV. **c,d)** PCA plot and MA plot of the reanalyzed data from He, *et al.* **e)** PCA plot of the reanalyzed data from Santaguida, *et al.*, recapitulating the three clusters of cells they identified: cells treated with DMSO, cells treated with the Mps1 inhibitor reversine and which were still cycling, and cells treated with reversine but which arrested. **f)** MA plot of the data in e) comparing Mps1i+ cells with control. **g)** RNAseq read depth of our 3 analyses vs the data from Santaguida, *et al.* and He, *et al.* Dots represent each technical replicate. n=12, 11, 6, 4, 6.

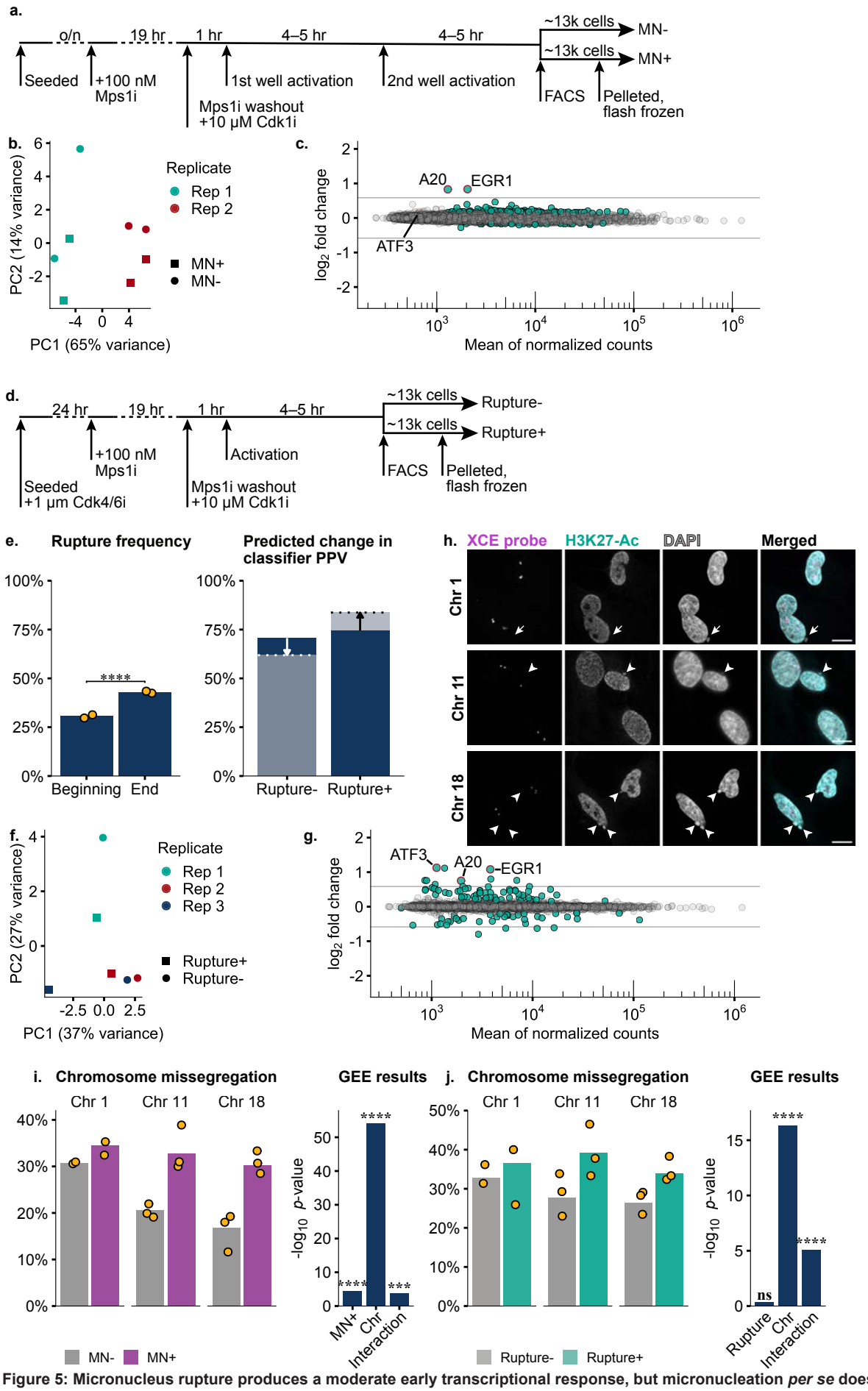

Figure 5: Micronucleus rupture produces a moderate early transcriptional response, but micronucleation *per se* does not

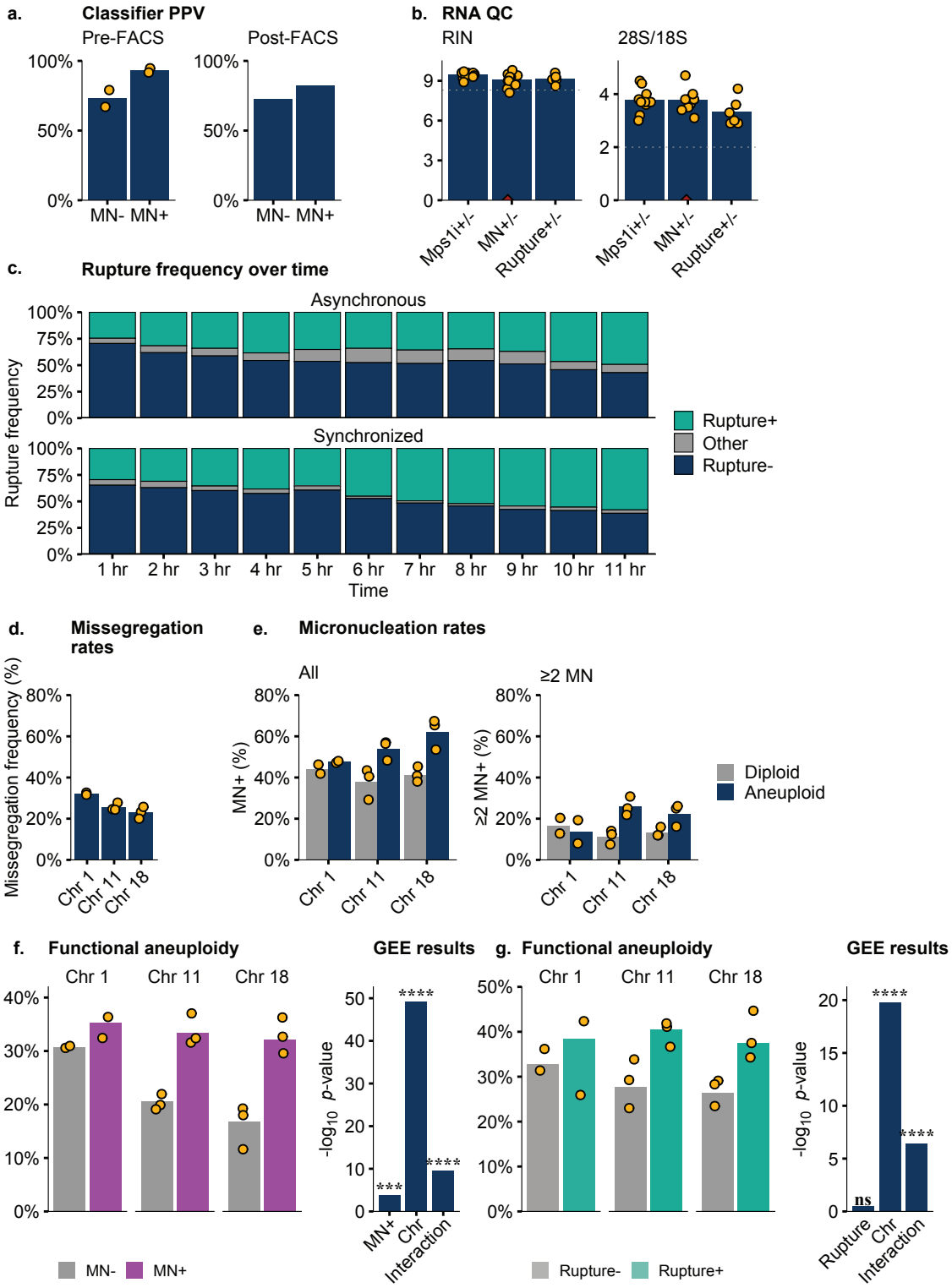

Figure S5: Supporting data related to figure 5

a) Accuracy after sorting cells and replating in Cdk1i to inhibit mitosis. Pre-FACS N=2; post-FACS N=1. **b)** RNA quality control measures from each experiment showing RIN scores and 28S/18S area measurements. Dots represent each technical replicate. n=12, 11, 6. One technical replicate had insufficient RNA concentration to measure RIN or 28S/18S area, represented in red. **c)** Change in rupture frequency over time in asynchronous and synchronized cells treated with Cdk1i. *Other* represents mitotic or MN- cells, or cells without Dendra2 expression. N=1, n=~200 cells. **d,e)** Quantification of data in Fig 5h, showing overall missegregation frequency in chromosomes 1, 11, or 18, and micronucleation rates in cells with and without missegregated chromosomes 1, 11, or 18. N=2,3,3. **f,g)** Same as Figs 5i,j but with a different measure of aneuploidy. Chromosomes in ruptured MN were considered to be functionally lost to the cell, so were ignored when tallying chromosomes/cell.

Figure 5: Micronucleus rupture produces a moderate early transcriptional response, but micronucleation *per se* does not

a) Timeline of experiment to isolate cells with and without MN. **b,c)** PCA and MA plots comparing MN+ vs MN- cells. **d)** Timeline of experiment to isolate MN+ cells with only intact or at least 1 ruptured MN. **e)** Change in MN rupture frequency between the start and end of an experiment and the predicted change in classifier PPV because of previously intact MN rupturing before the end of an experiment. **f,g)** PCA and MA plots of rupture+ cells vs rupture- cells.

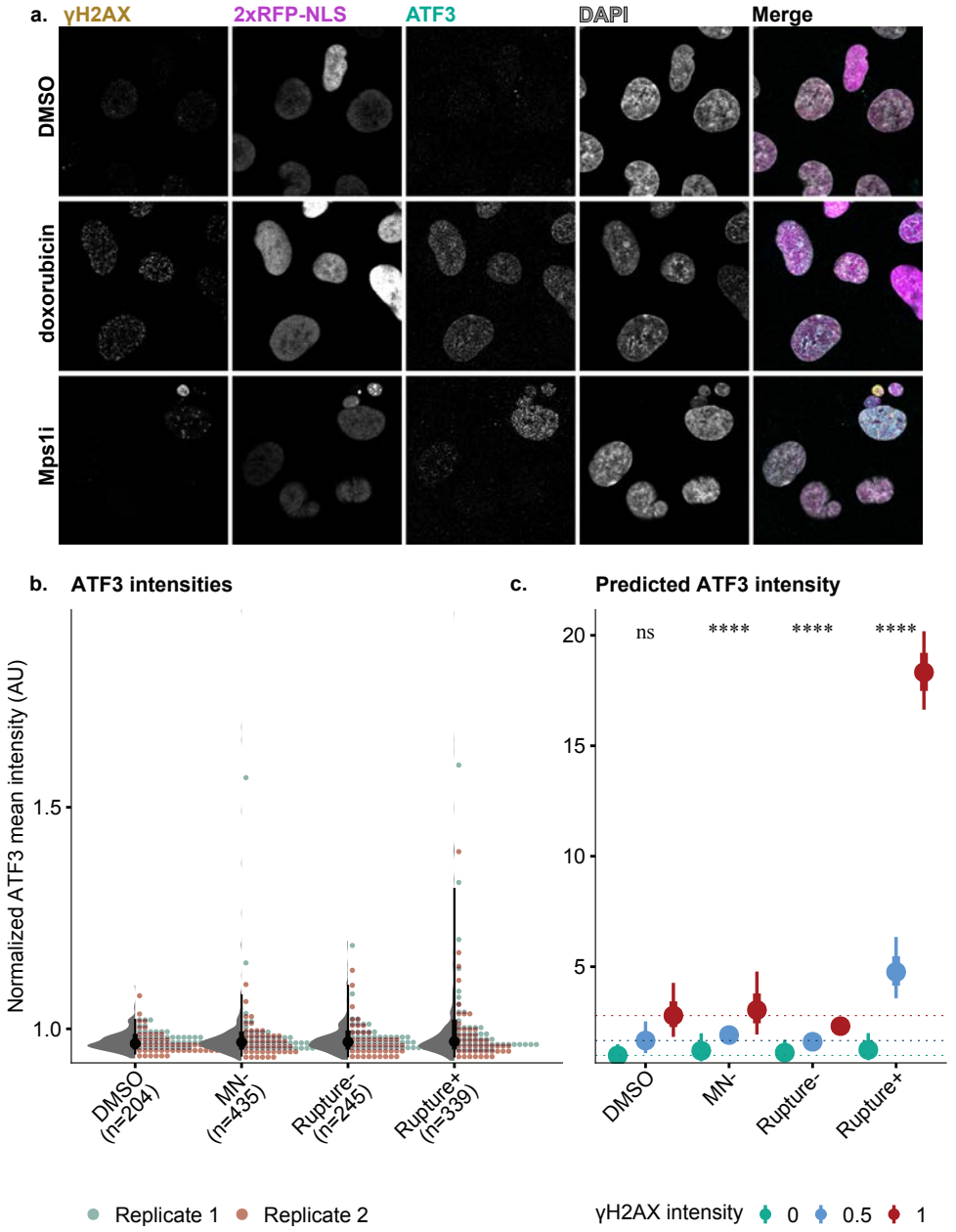

Figure 6: ATF3 is moderately increased in MN- and rupture+ populations
a) Example immunofluorescence images for quantifying γH2AX and ATF3 expression in cells treated with DMSO (neg. control), 2 μg/mL doxorubicin (pos. control), or with Mps1i. **b)** Violin & dot plots of ATF3 intensities from cells treated with DMSO or Mps1i. ATF3 intensity was normalized to max and min ATF3 intensities from all conditions per replicate then translated such that the 90th percentile of DMSO data was 1. **c)** GEE predicted ATF3 values showing interactions between population subgroups and γH2AX intensity, with GEE significance results for each population. Intervals show mean prediction and 95% and 66% CIs. Dotted horizontal lines show the mean prediction from the DMSO group.

Figure 5: Micronucleus rupture produces a moderate early transcriptional response, but micronucleation *per se* does not
h) Example of centromere FISH on chromosomes 1, 11, and 18 and the use of H3 lysine 27 acetylation (H3K27-Ac) to mark intact MN. Ruptured MN are annotated with arrows and intact with arrowheads. **i)** Quantification of the data in h). Chromosome missegregation examines the total number of chromosomes in each cell, regardless of its placement within a MN. Chr 1: N=2, n=429, 158; Chr 11: N=3, n=406, 313, 160; Chr 18: N=3, n=425, 202, 230. **j)** Same as i), but comparing rupture- and rupture+ cells. Chr 1: N=2, n=187, 74; Chr 11: N=3, n=190, 108, 71; Chr 18: N=3, n=186, 102, 101.

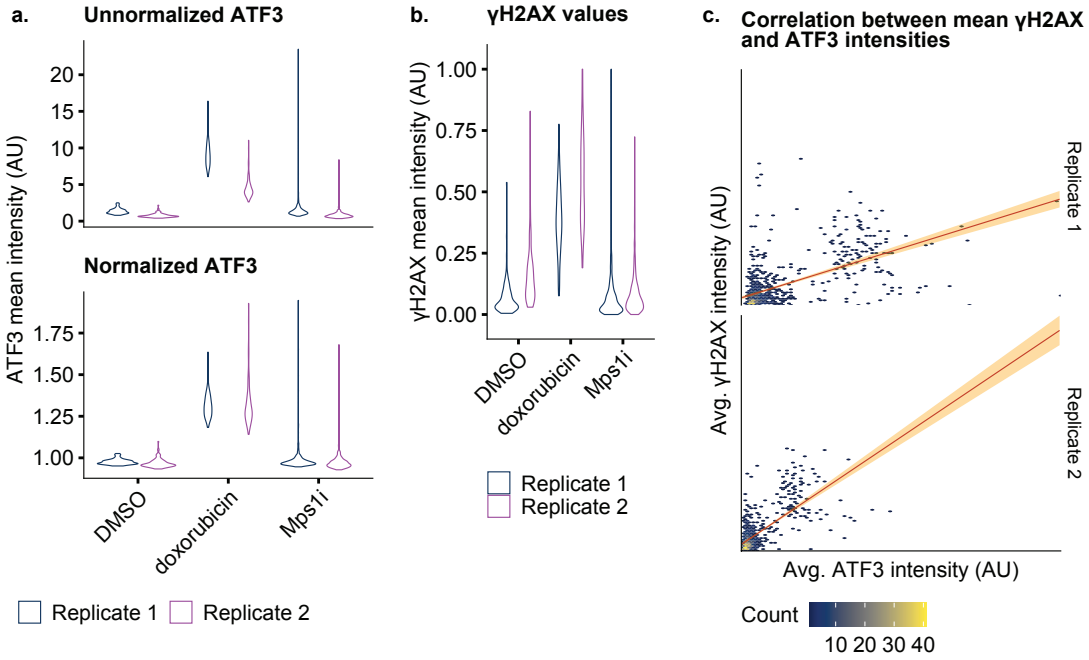

Figure S6: Supporting data for Figure 6

a) Violin plots of ATF3 intensities with and without rescaling or translation applied in all treatment conditions. **b)** Violin plots of γH2AX intensities in all treatment conditions. **c)** Plots of correlation between γH2AX and ATF3 intensities per replicate.

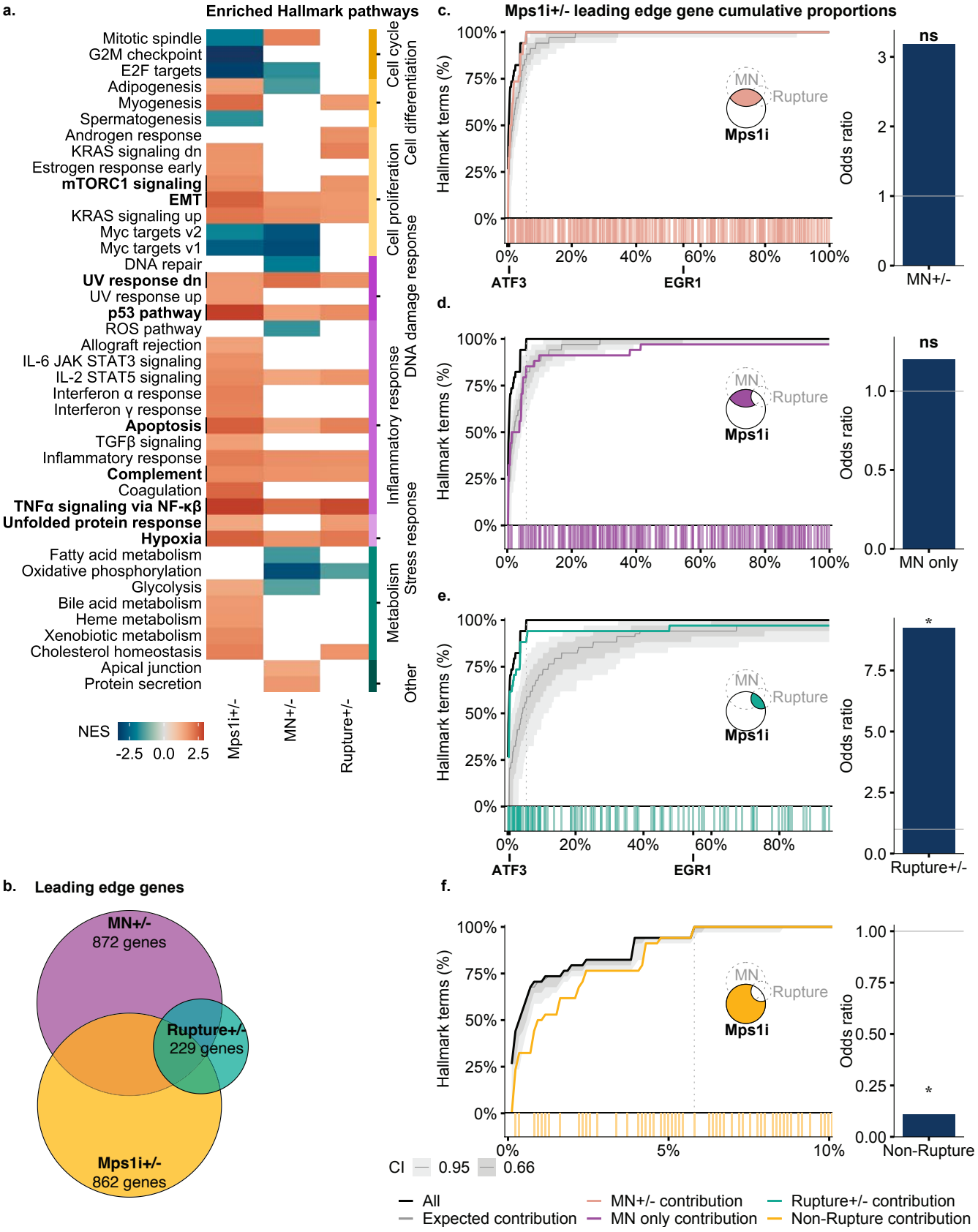

Figure 7: Micronucleus rupture is a strong contributor to the transcriptomic response to Mps1i treatment
a) Heatmap of Hallmark pathways identified in RFP703/Dendra cells treated with DMSO vs Mps1i; Mps1i-treated cells with and without MN; and synchronized, Mps1i-treated, MN+ cells with and without MN rupture. Pathways with an overrepresentation of rupture+/- leading edge genes in the Mps1i+/- condition are marked in bold. **b)** Venn diagram of overlap in leading edge genes between the three data sets. **c)** Analysis of the contribution of MN+/- leading edge genes to the Mps1i+/- transcriptomic response. Genes that are sufficient to represent all Hallmark pathways identified in these experiments (core genes) are to the left of the dotted line, and the placement of shared MN+/- genes displayed along the x-axis. The number of Hallmark pathways represented only by the MN+/- subset are represented in pink, with the expected contribution if these leading edge genes were randomly distributed in gray. An odds ratio comparing the odds of shared genes being among core genes vs not is represented on the right. **d–f)** Same as c, but with different subsets of genes identified from the overlap indicated. f) is zoomed on the first 10% of genes.

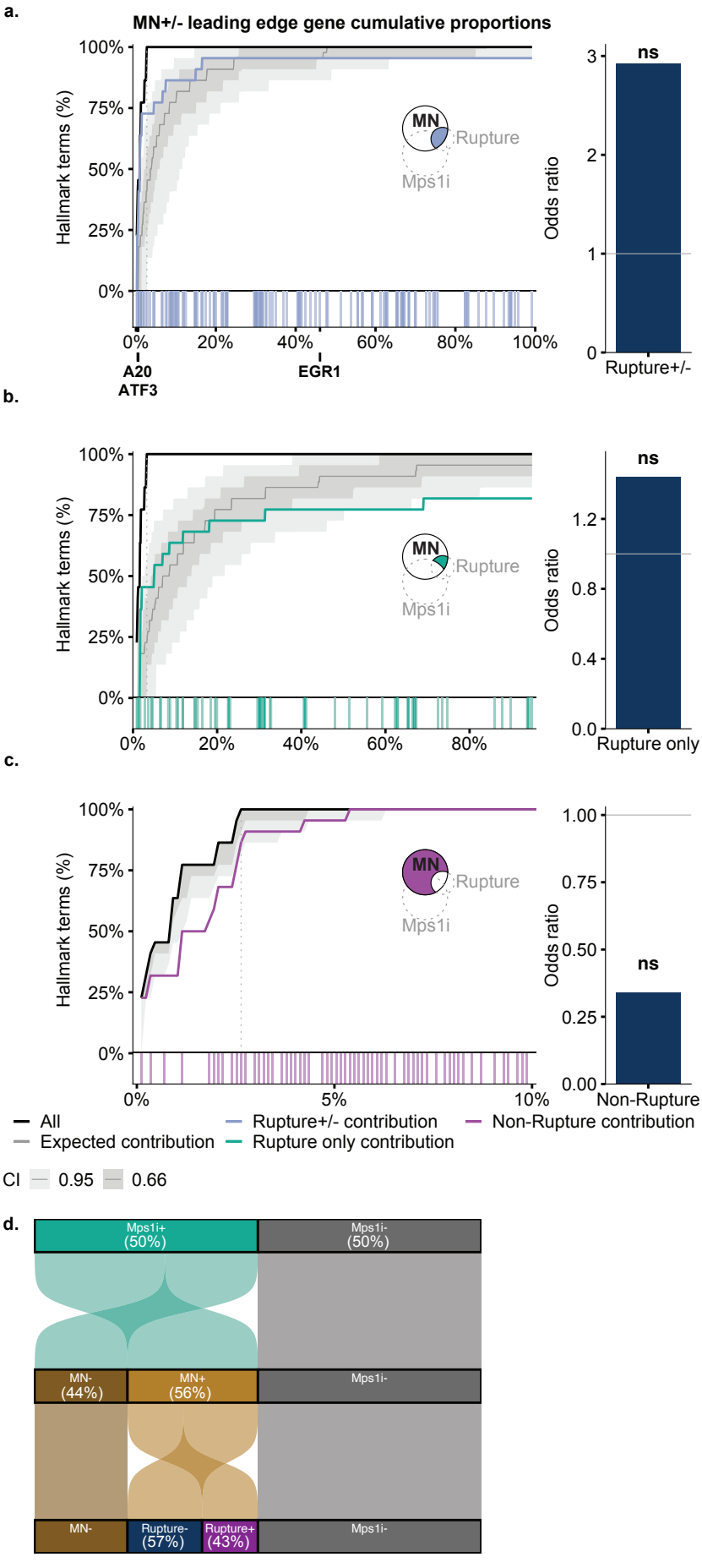

Figure S7: Micronucleus rupture is a contributer to the transcriptomic response to micronucleation
a–c) Same as Figs 7b–f, but examining contributions to the MN+/- transcriptomic response. **d)** Sankey diagram showing the breakdown of populations in the Mps1i+/- treatment condition.

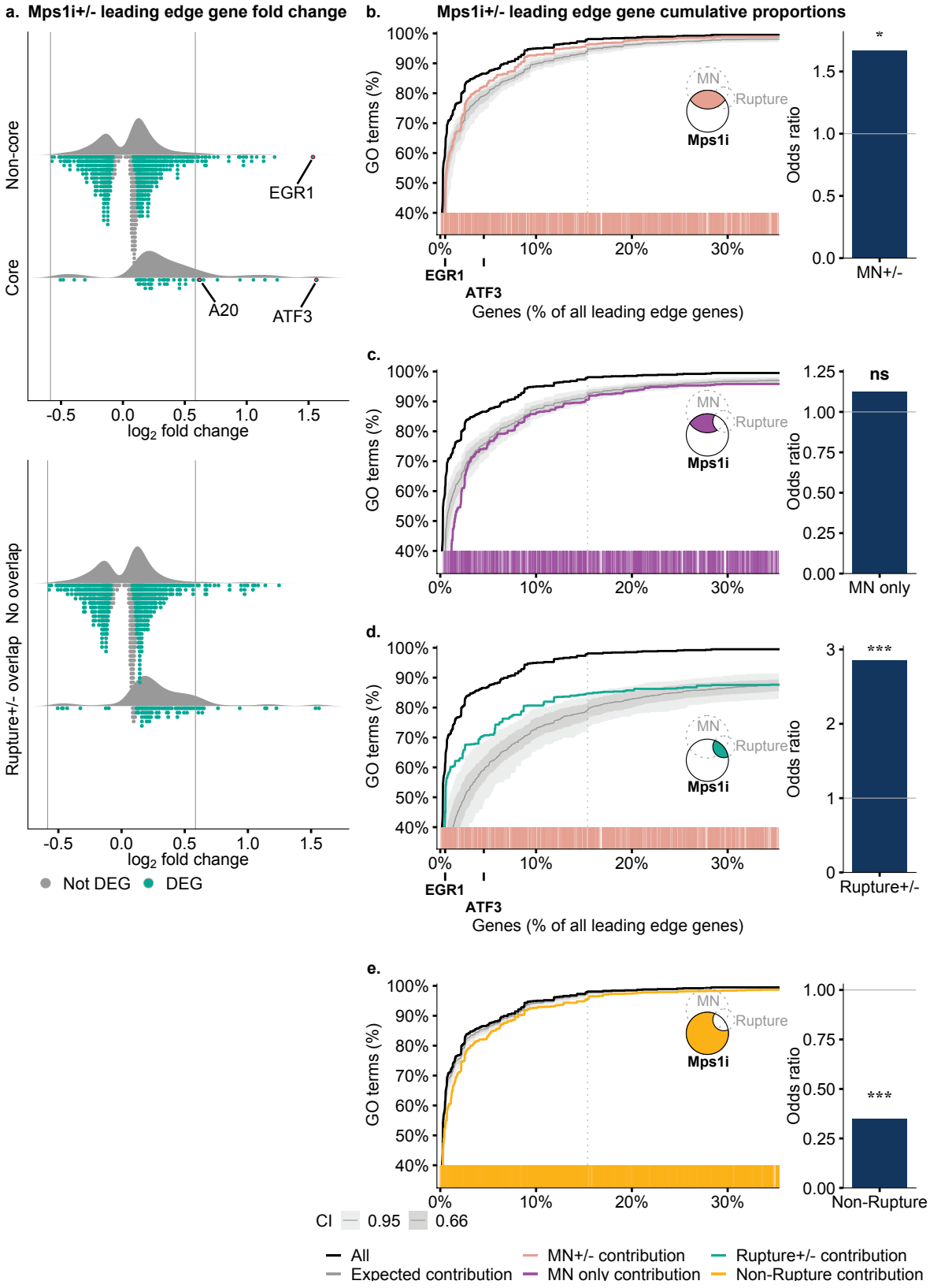

Figure S8: MN rupture predicts changes to GO terms
a) Rain cloud plots of the fold-change distribution of leading edge genes from the transcriptomic response to Mps1i treatment, by core and non-core genes, and by those genes that are leading edge genes in rupture+ cells vs not. DEGs are highlighted in green. Vertical lines mark a 1.5 fold change. **b–e)** Same as Figs 7c–f but looking at the GO ontology.

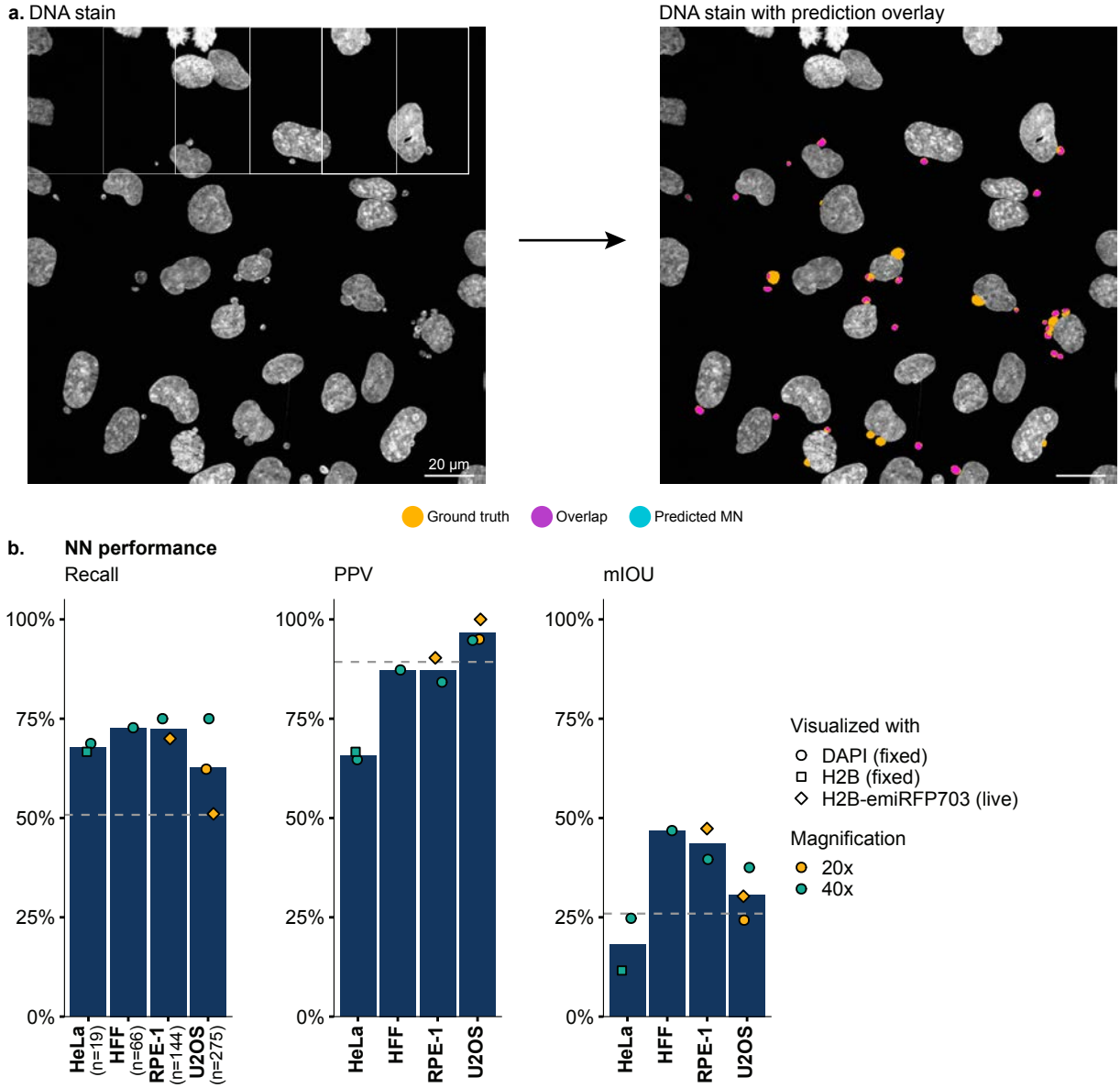

Figure 8: An updated neural net improves micronucleus segmentation for several different cell lines and imaging conditions
a) Overview of the neural net process showing oversampling by sliding window, with overlays showing ground truth, MN predictions, and overlap. **b)** Recall, PPV, and mean intersection over union (mIOU) by cell line. Bars indicate mean performance and dashed line the baseline performance from the VCS neural net.

II. Dam Tracker

While the consequences of MN rupture to cell is of interest, there are also consequences to the chromatin within itself. These changes include deacetylation of histone marks and persistent changes to transcriptional regulation, even into following cell cycles.[33] However, there were few tools to study these changes, particularly in daughter cells. One method to facilitate this study would be to covalently tag DNA within ruptured MN. By semi-permanently tagging DNA, it could be either pulled down in bulk by ChIP and probed directly or followed visually over time. However, this tag would need to be unique and not interrupt other cellular processes. One such candidate would be N6-methyladenosine using an extension of a technique known as DamID. In addition, in conjunction with VCS, it could provide a mechanism to rapidly identify and sort out the daughters of cells that had experienced MN rupture.

Deoxyadenosine methyltransferase (Dam) is a bacterial enzyme that decorates adenosine in double-stranded DNA with a methyl group on the N6 position, generating m6-adenosine (m6A).[81] Targeting GATC motifs, this mark is used for gene regulation, mismatch repair, and regulating DNA replication in *E. coli*.[82] This is a mark that does not occur normally in human cells, or only very rarely (possibly as a result of the occasional incorporation of RNA bases during DNA replication). [83–87] This makes it a useful mechanism to mark human DNA. Dam was initially used to probe chromatin structure before being developed into a protocol known as DamID, where Dam is fused to a transcription factor of interest and methylates nearby GATC sites.[88–90] Antibodies to m6A can then be used to pull down the marked chromatin, providing a map of transcription binding sites.[91]

In 2013, van Steensel's lab produced a Lamin-Dam fusion to mark the dynamics of how heterochromatin is organized at the nucleus periphery.[92] To enable live-cell tracking, they randomly truncated DpnI, a bacterial enzyme that cleaves G^{m6}ATC, until they found a mutant that would bind these sites but which lacked any catalytic

activity. Fusing this truncated protein to GFP gave them a live-cell tracker, which they dubbed m6A-tracer. Later research on the structure of DpnI revealed that it is composed of an N-terminal catalytic domain responsible for cleavage and which participates in m6A recognition and a C-terminal "winged-helix" domain that participates in m6A recognition but has no catalytic activity.[93] The truncation van Steensel's group produced was one that included just the winged-helix domain and a small portion of the catalytic domain.

Dam's ability to covalently mark DNA and the existence of a live-cell tracker made it a viable option for tagging the chromatin in ruptured MN and following it over time. The plan as originally conceived was to use Dam fused to a nuclear exclusion signal (NES) and a bulky fluorophore such that it would be restricted to the cytosol. If a MN formed and ruptured, Dam would gain access to this compartment and methylate contained chromatin. However, for this to specifically mark chromatin in ruptured MN, Dam activity would have to be restricted to interphase: if any Dam remained active during mitosis, mitotic nuclear envelope breakdown would allow Dam to non-specifically methylate all available GATC sites. To solve this problem, the plan was to place Dam under the control of a Tet-inducible promoter and to attach a conditional degron, such that Dam could be transcribed and degraded at will. Cells would be synchronized, then Dam "turned on" at the start of interphase and degraded before cells entered mitosis.

Results

Dam expression can be controlled and tuned with a conditional degron and tet-inducible promoter

I developed a construct with Takara's "destabilization domain" as the conditional degron, based on previous work done by Heather Huang, a former lab technician. This degron is a mutated FKBP12 domain that fails to properly fold, triggering protein degradation.[94] Addition of a small molecule, Shield1, stabilizes the degron and prevents protein degradation. I fused this domain to an NES, Dam, a 4x repeat of the

V5 epitope tag to facilitate detection by immunofluorescence and Western blot, and finally a TagRFP fluorophore (**Fig 2-1a**). An earlier construct used a single V5 epitope tag, but this proved insufficient for detecting protein when expressed at low levels. This construct was placed under the control of a TRE-Tight promoter from Takara, a promoter with repeated binding sites for a Tet-inducible transactivator, and transduced into RPE-1 cells that already expressed the TetOn transactivator (see Methods). The 4xV5 cell line will be hereafter known as RPE-1 Dam-4xV5; the cell line transduced with the single-V5 construct will be identified as RPE-1 Dam-V5.

To determine the kinetics of Dam expression, I performed a time course where RPE-1 Dam-4xV5 cells were treated with 0–64 ng/mL doxycycline +/- 1 μM Shield1. Cell lysate was then sampled immediately following drug addition, then every hour for 6 hours, and once more 20 hrs post-treatment. Without addition of doxycycline or Shield1, no V5 signal appeared even after 20 hrs. Addition of doxycycline, even without Shield1, eventually induced a detectable V5 signal. Addition of Shield1 decreased the time till detectable levels of V5 were present, suggesting translation can overwhelm degradation. With the addition of Shield1, V5 was detectable within 1–2 hrs, even with only 1 ng/mL doxycyline (**Figs 2-1b, 2-1c**).

To test the kinetics of protein degradation, I incubated cells in doxycycline + Shield1 for 4 hours—sampling lysate every hour—then washed out both drugs and continued to sample cell lysate every hour for 5 additional hours. V5 signal was still strong 5 hours after drug washout, even with incubation in low levels of doxycycline (**Fig 2-1d**). To speed protein degradation, I pulsed doxycycline at 4 ng/mL for 15 min in the absence of Shield1 then washed it out in order to fix the number of mRNA transcripts created. However, even after 8 hours, there was detectable V5 signal (**Fig 2-1e**). In summary, these data suggest that Shield1 promotes a faster increase in Dam expression and increases peak expression, but that the primary tunable parameter in Dam expression is doxycycline concentration. In addition, while protein expression is rapidly induced, degradation—even in the absence of Shield1 ligand—is slow.

m6A is detectable following Dam expression through mitosis by dot blot and restriction-digest PCR

To detect the presence of m6A, I extracted genomic DNA from RPE-1 Dam-V5 cells with and without incubation in doxycycline and Shield1 for 24 hours, then blotted DNA onto a nitrocellulose membrane with a dot blot apparatus that I had constructed. m6A visualization by immunofluorescence showed a global increase in m6A levels (**Fig 2-2a**). This was confirmed using restriction-digest PCR, where genomic DNA is subjected to mock digestion or DpnII digestion, which digests only unmethylated GATC sites.[88] In brief, if all GATC sites are methylated, template DNA will be protected from DpnII digestion. Using primers for a locus in the *LMNA* gene flanking multiple GATC sites, I confirmed that Dam was active only in the presence of doxycycline and Shield1 (**Fig 2-2b**).

An improved, mild protocol for visualizing m6A allows for standard immunofluorescence protocols

To better visualize m6A, I sought to use immunofluorescence on fixed cells. However, there is no standard protocol for visualizing m6A by immunofluorescence. Using related protocols for DNA fluorescence in-situ hybridization (FISH) and BrdU staining, I used incubation in high heat and formamide to denature DNA such that anti-m6A could bind. While I was able to detect m6A following a 1 hr incubation in a 50% formamide solution at 74 °C (**Fig 1-3a**), the harsh conditions made co-staining additional epitopes impossible and rendered this protocol insufficient for my needs.

To get around this limitation, I developed a milder m6A staining protocol. DpnI specifically recognizes G^{m6}ATC and creates blunt-end cuts between the adenosine and thymidine residues. I reasoned that this alone may be sufficient for recognition and binding by anti-m6A. After fixation in methanol and permeabilization in a 2% TritonX solution, I incubated cells in 50 U/mL DpnI in restriction enzyme buffer overnight at 37 °C. Later optimizations found this could be reduced to as little as 30 min and that either PFA or methanol fixation was suitable. After washing out enzyme, a standard immunofluorescence protocol was followed. Remarkably, this protocol

enabled the visualization of m6A while permitting co-staining of other epitopes (**Fig 2-3b**). To test the colocalization of m6A and the m6A-tracer developed by Kind, *et al.*, and to compare m6A-tracer with a rationally designed tracer—EGFP-NES fused to just the winged-helix domain of DpnI, hereafter m6A-wH—I plated RPE-1 Dam-4xV5 cells transduced with either m6A-tracer or m6A-wH. The following day, I treated them with Mps1i to induce micronucleation. The next day, half the cells were supplemented with 16 ng/mL doxycycline and 1 μM Shield1 for 8 hours, still in the presence of Mps1i. Following incubation, some coverslips were fixed with methanol and stained for m6A. Methanol fixation can severely diminish intrinsic fluorophore signal, so Dam localization was imaged using anti-V5 bound to Alexa Fluor 568-conjugated 2° antibodies (to overlap with any remaining TagRFP signal) and m6A-tracer and m6A-wH signal was visualized with anti-GFP bound to Alexa Fluor 488-conjugated 2° antibodies. The m6A-tracer showed considerable colocalization with m6A (**Fig 2-3c**), quantified by Li's intensity correlation quotient (ICQ).[95] This analysis showed that nearly all m6A signal overlapped with m6A-tracer. In contrast, the m6A-wH construct was localized to most nuclei absent any corresponding m6A signal (**Fig 2-3d**). This was reflected in a lower ICQ and a sharp cutoff when comparing m6A signal to m6A-wH (**Fig 2-3d**, top graph), indicative of m6A signal only overlapping with a subset of m6A-wH signal. This could be due to mislocalization of m6A-wH, or a true reflection of underlying background Dam activity, only seen because of greater sensitivity from the m6A-wH construct.

m6A-tracer localizes specifically to ruptured micronuclei

Given the higher background of the m6A-wH construct, I decided to only continue with m6A-tracer. To test the specificity of the m6A-tracer construct for ruptured MN, the remaining coverslips used to test for m6A colocalization were fixed with 4% PFA and stained for acetylated lysines on histone 3 using a pan-acetyl antibody that recognizes acetylated lysine 9, 14, 18, 23, or 27 (H3K•-acetyl). Loss of histone acetylation is indicative of MN rupture and can be used to distinguish intact and ruptured MN.[15,16,27] Because of a generally weak TagRFP signal, I also stained for V5

with Alexa Fluor 568-conjugated 2° antibodies to visualize Dam (**Fig 2-4**). Cells were asynchronous and may have undergone mitosis while Dam was active, which should result in an m6A-tracer+ nucleus. Excluding V5- cells, indicating no Dam expression, or cells with m6A-tracer+ nuclei, in a manual tally, 79% of ruptured MN and only 21% of intact MN were m6A-tracer+ (**Fig 2-4**), suggesting some specificity for ruptured MN. Acetylation status is a good but imperfect proxy for MN rupture status—deacetylation takes some amount of time, resulting in recently ruptured MN still bearing acetylation marks at the time of fixation. This suggests that the true proportion of m6a-tracer+ intact MN may be lower.

Future directions

The goal of this project was to develop a construct that could selectively mark the chromatin in ruptured MN with a covalent tag, m6A. This required the construct only be active during interphase. The inability to return Dam to baseline expression within 8 hours, even with very short incubations in doxycycline and without Shield1, made this impractical without the use of cell-cycle drugs that could extend the length of interphase. I wanted to avoid this, if possible, so that I wouldn't introduce confounding artifacts as we studied the long-term effects of ruptured MN, particularly in future cell cycles.

Papathanasiou, *et al.*, created a similar construct—Dam fused to multiple degrons—that they were able to successfully use.[33] This confirms the overall strategy. Papthanasiou, *et al.*, overcame the problem of open mitosis by using a different conditional degron system, AID, and multiple repeats of this conditional degron.

Tobacco vein mosaic virus protease (TVMVP) is a protease that recognizes a relatively long motif and is specific, making it a useful tool for engineering biological circuits in mammalian cells.[96–98] Proteolysis is a much faster process than proteasomal degradation, and this raises the possibility of an alternative method for ending Dam activity prior to mitosis.

Such a system would rely on a TVMVP fused to a nuclear localization signal (NLS) and a fluorophore, both to track TVMVP expression and to distinguish ruptured and intact micronuclei, and a modified Dam enzyme that could be cleaved by TVMVP. During MN rupture, some TVMVP would be released into the cytoplasm (potentially cleaving some Dam molecules) before being imported into another intact MN or, more likely, the nucleus. During prometaphase, as the nuclear envelope broke down, TVMVP would be released *en masse* into the cytoplasm and cleave Dam. If cleavage was fast enough, and Dam activity slow enough, this could occur before non-specific methylation. If needed, there are known Dam mutations to reduce Dam activity and processivity.[99]

To produce a cleavable Dam enzyme, I modified a split Dam system created by Hass, *et al.*[100] I joined the two halves of this construct with a linker region containing the TVMVP recognition motif. Expressing this construct in the presence of constitutively active TVMVP resulted in cleavage of a subset of Dam, raising the possibility that this system could be tuned to disable Dam at the onset of mitosis (**Fig 2-5**).

The benefits of a proteolysis-regulated system would be cell-intrinsic regulation: since mitosis itself would initiate Dam degradation, there would be no need to use synchronized cells. This could simplify study design and potentially avoid confounding artifacts introduced by cell synchronization.

Materials and Methods

Plasmid construction

pLVX-Tight-DD-NES-Dam-V5-TagRFP was created through several intermediate steps. First, pLVX-Tight-DD-V5-TagRFP was created by PCR of DD-V5-TagRFP from pQCXIB-CMV-TO-DD-V5-TagRFP-NES-Dam-HA, a plasmid created by a lab technician, Heather Huang, with forward primer (with prepended BamHI digest site) 5'-***gaaagg*****ggatcc**ATGGGAGTGCAGGTGGAAAC-3' and reverse primer (with prepended EcoRI digest site and stop codon) 5'-***gaaagg*****gaattc**TTAGTGCCCCAGTTTGCTAGGG-3',

followed by insertion into pLVX-Tight by restriction digest cloning (NEB BamHI-HF, #R3136;; EcoRI-HF, #R3101). NES-Dam was extracted by PCR from pCMV-DD-V5-C1-TagRFP-Luc-NES-Dam-ns-HA, a plasmid created by Dr. Hatch, with forward primer (with prepended AgeI digest site) 5'-***gaaagg*****ACCGGTT**CTACAGCTACCGCCGCTGG-3' and reverse primer (with prepended AgeI digest site) 5'-***gaaagg*****ACCGGTGG**TTTTTTCGCGGGTGAAACGACTC-3', followed by insertion into pLVX-Tight-DD-V5-TagRFP between the destruction domain (DD) and the V5 epitope tag (NEB AgeI-HF, #R3552).

Because Dam has an internal BamHI digest site, I then inserted two new digest sites in front of the construct by PCR extension, extracting DD-NES-Dam-V5-TagRFP using forward primer (with inserted XbaI and NotI digest sites after BamHI) 5'-***ggaagg***ggatcc**tctagagcggccgc**ATGAAGAAAAATCGCGCTTTTTTG-3' and reverse primer 5'-***gaaagg***gaattcTTAGTGCCCCAGTTTGCTAGGG-3'. This was re-cloned into the pLVX-Tight backbone by restriction digest cloning (NEB NotI-HF, #R3189; EcoRI-HF). Digestion of the insert produced two products, because of the BamHI cut site, so both were ligated together with the vector and Sanger sequencing by Genewiz used to verify proper orientation after ligation.

pLVX-Tight-DD-NES-Dam-4xV5-TagRFP was created through several intermediate steps. To facilitate insertion of a 3xV5 gBlock from IDT, I used PCR extension and site-directed mutagenesis to insert a multiple digest site (MDS) after the V5 epitope tag. Using forward primer 5'-***ggaagg***gcggccgcATGGGAGTGCAGGTGGAAAC-3' and reverse primer (with prepended ClaI, XbaI, HindIII, NheI sites) 5'-**CGTAGAATCGAGACCGAGGAGAG**GCTAGCAAGCTTTCTAGAATCGAT-3', I extracted DD-NES-Dam-V5-MDS. Using forward primer (with prepended MDS) 5'-**AGCGAGCTGATTAAGGAGAACATG**ATCGATTCTAGAAAGCTTGCTAGC-3' and reverse primer 5'-***gaaagg***gaattcTTAGTGCCCCAGTTTGCTAGGG-3', I extracted MDS-TagRFP. To create a single DD-NES-Dam-V5-MDS-TagRFP construct, I used splicing by overlap extension PCR (SOE) to join both fragments together, using the most distal primers.[102] This

construct was cloned into the pLVX-Tight vector plasmid by restriction digest cloning. I ordered a 3xV5 repeat gBlock from IDT flanked by HindIII and NheI digest sites. To prevent recombination during cloning and errors during gBlock synthesis, I used synonymous codons in each V5 repeat to make each repeat as dissimilar as possible, with the full gBlock sequence being: 5'-ATCGATGGCAAACCCATTCCCAATCCCTTATTAGGCTTAGACAGCACTGGAAAGCCAATACCAAACCCAT TGTTGGGATTGGATAGTACCGGGAAACCGATCCCGAATCCGCTTCTTGGGCTTGACTCCACAGCTAGC-3'. This gBlock was inserted into pLVX-Tight-DD-NES-Dam-V5-MDS-TagRFP by restriction digest cloning (NEB HindIII, #R0104; NheI-HF, #R3131).

Cleavable Dam was created by regenerating the DD-NES-Dam-4xV5-TagRFP construct, but with split Dam linked by a TVMVP-site-containing linker replacing Dam and emiRFP703 replacing TagRFP. First, DD-NES-Dam-4xV5 was isolated from pLVX-Tight-DD-NES-Dam-4xV5-TagRFP by restriction digest (NEB XbaI, #R0145, NEB NheI-HF). emiRFP703 was isolated from pLVX-EF1a-H2B-emiRFP703-neo by PCR using forward primer (with prepended NheI digest site) 5'-***ggaaggg*****gctagc**GCCACCATGGCGGAAGGC-3' and reverse primer (with prepended EcoRI digest site) 5'-***ggaaggg*****gaattc**CTTAGCTCTCAAGCGCGGTGATC-3'. The two components were ligated together by restriction digest cloning to produce DD-NES-Dam-4xV5-emiRFP703. DD-NES-$_{N}$Dam was extracted from DD-NES-Dam-4xV5-emiRFP703 by PCR using forward primer 5'-***ggaaggg*****gcggccgc**ATGGGAGTGCAGGTGGAAAC-3' (with prepended NotI digest site) and reverse primer 5'-***ggaaggg*****GGGCCC**GTACTCATCAGTACGCATCTTCACAATG-3' (with prepended ApaI digest site). $_{C}$Dam was isolated from pBI-Tet-Notch1dE-N-Dam / RBPj-C-Dam-Ert2 (a gift from the Kopan Lab) by PCR using forward primer 5'-***ggaaggg*****GGCGCGCCG**TCTAGTCCAGTGTGGTGGAATTCTGC-3' (with prepended AscI digest site) and reverse primer 5'-***ggaaggg*****ACCGGTG**GCTTCGAGATCGCCATCCCGG-3' (with prepended AgeI digest site). 4xV5-emiRFP703 was isolated from DD-NES-Dam-4xV5-emiRFP703 by restriction digest (NEB AgeI, EcoRI-HF). A linker containing GGGGS repeats and the TVMVP cut site ETVRFQS with flanking ApaI and AscI digest sites was ordered from IDT. All components were ligated together as follows: *NotI*-DD-NES-

$_N$Dam-*ApaI* + *ApaI*-linker-*AscI* + *AscI*-CDam-*AgeI* + *AgeI*-4xV5-emiRFP703-*EcoRI* and ligated into the pLVX-Tight backbone.

pLVX-EF1a-EGFP-NES-m6A-tracer was created by Dr. Hatch by subcloning EGFP-NES-m6A-tracer from a plasmid provided by the Hetzer lab into a pLVX-EF1a vector.

pLVX-EF1a-EGFP-NES-wH was created by first ordering a gBlock containing the full DpnI sequence from IDT, flanked by NotI and EcoRI digest sites. This was then subcloned into a pLVX-EF1a vector by restriction digest cloning to create pLVX-EF1a-DpnI. EGFP-NES was extracted by PCR from pLVX-EF1a-EGFP-NES-m6A-tracer using forward primer (with prepended NotI digest site) 5'-***gaaaggcggcc***ATGGTGAGCAAGGGCGA-3' and reverse primer (with prepended SalI digest site) 5'-***ggaagg*GTCGAC**CAGAGTCAGTCTTTCCAGCGG-3'. The winged-helix domain of DpnI was extracted by PCR with forward primer (with prepended SalI digest site) 5'-***ggaagg*GTCGAC**AGCTCTCTGTCATCAAGAGGTTG-3' and reverse 5'-***ggaagg***gaattcTTATCATAATTTCCGATACTTTCCTCTACC-3'. The construct was then created by restriction digest cloning (NEB SalI-HF, #R3138).

All primers were manufactured by IDT and resuspended to 100 μM stock solution in 1xTE and diluted to a 10 μM working solution in ultrapure H_2O. All plasmid sequences were checked with Sanger sequencing by Genewiz.

Lentivirus production

HEK293T cells were plated in 10-cm tissue-culture–treated dishes at 70% confluency in DMEM supplemented with 10% FBS and 1% pen/strep. The next day, I replaced with fresh media and transfected cells with a solution of 3 μg psPAX2, 1.5 μg pVSV-G, and 2 μg of construct plasmid after a 5 minute incubation in OptiMEM with 40 μg linear PEI.

The following day, I refreshed with media supplemented with a sterile 25 mM HEPES solution in diH_2O. For pLVX-Tight plasmids, this media used Tet-free FBS.

The following day, I removed 5 mL of supernatant and filtered it through 0.45 μm PES syringe filters and stored the filtrate at 4 °C. The remaining media was removed and replaced with fresh, HEPES-supplemented media.

The next day, I harvested an additional 5 mL and filtered it, adding the filtrate to the previous day's. I then added 20 μg DNaseI (Fisher, #FEREN0523) and $MgCl_2$ to a 2 mM final concentration. After incubating at room temperature for 30 min, I moved the filtrate to 4 °C for another 4 hours, then aliquoted virus and snap froze in liquid N_2. Virus aliquots were thereafter stored at -80 °C.

Competent virus for the m6A-tracer and m6A-wH constructs could not be produced from plasmids harvested in regular *E. coli* lines, such as Stbl3 or TOP10. All plasmids needed to be produced in Dam-incompetent bacteria, to prevent the construct from coating the plasmids. I first transformed the plasmids, including helper plasmids, into One Shot INV110 chemically competent cells (ThermoFisher, #C717103), then harvested these m6A-free plasmids and used them for virus production.

Cell lines

hTERT RPE-1 TO/DD-NES-Dam-V5-TagRFP and hTERT RPE-1 TO/DD-NES-Dam-4xV5-TagRFP cell lines were produced through transduction of the Dam plasmid into a lab stock of hTERT RPE-1 cells with the TetOn transactivator already transduced (RPE-1 TO cells). Addition of the m6a-tracer or m6A-wH constructs was likewise achieved by transduction into the hTERT RPE-1 TO/DD-NES-Dam-4xV5-TagRFP cell line. U2OS TO cells were produced by transduction of pLVX-EF1a-TetOn (a gift from the Hetzer lab) into U2OS cells.

Transduction was performed by splitting cells to 40% confluency into a 10-cm dish, and immediately adding a polybrene solution in PBS to a final concentration of 6

μg/mL, followed by virus filtrate added dropwise. For Dam plasmids, 20 μL of virus filtrate was used; for all others, 1 mL. Cells were then split twice before selection was added: blasticidin to a final concentration of 10 μg/mL, Geneticin/G418 to a final concentration of 500 μg/mL active G418, and/or puromycin to a final concentration of 20 μg/mL. TetOn carries a neomycin resistance gene; the Dam constructs, blasticidin resistance; and the m6A-tracer and -wH constructs, puromycin resistance.

Little virus was used for transducing Dam constructs to ensure single-copy insertions. To facilitate growth of the few RPE-1 cells successfully transduced, I split the remaining cells following 2 days of selection down to a 12-well plate and supplemented media with conditioned media—media from a confluent plate of hTERT RPE-1 cells filtered through a 0.22 μm filter (Genesee Scientific, #25-244)—until cell numbers increased.

Tissue culture

hTERT RPE-1 cell lines were kept in DMEM/F12 supplemented with 10% FBS, 1% pen/strep, and 0.01 mg/mL hygromycin B and stored in a cell incubator with a 5% CO_2 atmosphere at 37 ºC. Cells with transduced Dam constructs were supplemented with 10 μg/mL blasticidin and 500 μg/mL G418. Cells with m6A-tracer or -wH constructs were additionally supplemented with 20 μg/mL puromycin.

U2OS TO cells were kept in DMEM supplemented with 10% FBS and 1% pen/strep, and stored in a cell incubator with a 10% CO2 atmosphere at 37 ºC and likewise supplemented with G418.

Dam expression time course

I plated 150k RPE-1 Dam-4xV5 cells/well into 8, 12-well dishes (ThermoFisher, #130185). Each plate was used for a single time point. The following day, I replaced the media with media supplemented with 0, 1, 2, 4, 8, or 64 ng/mL doxycycline (Clontech, #631311) +/- 1 μM Shield1 (Takara, #632189). I immediately washed the

first plate with PBS and lysed the cells with 100 μL/well of 1x sample loading buffer (Invitrogen, NuPAGE LDS sample buffer, #NP0007) supplemented with 6.25% 2-mercaptoethanol (BME [Sigma, #M3148]). Lysate was prepared for Western blotting. I repeated this for each plate at 1, 2, 3, 4, 5, 6, and 20 hrs post-addition of doxycycline.

Dam expression/degradation time course

I plated 50k RPE-1 Dam-4xV5/m6A-tracer cells into 12-well dishes. 2 days later, I replaced the media with media supplemented with 1 or 4 ng/mL doxycycline and 1 μM Shield1. Immediately, I washed one well with PBS then lysed in 1x SLB+BME and prepped it for Western blotting. Every hour thereafter, I sampled another well in a similar manner. After 4 hours, I washed the remaining wells 3x with PBS, then replaced with plain media. I continued to sample every hour until 9 hours post-drug addition.

Dam pulse-doxycycline time course

I plated 50k RPE-1 Dam-4xV5/m6A-tracer cells into 12-well dishes. 2 days later, I replaced media with 4 ng/mL doxycycline without Shield1 for 15 min, then washed the wells 3x with PBS and replaced with fresh media. Immediately, I lysed one well with 1xSLB+BME. At 1, 2, 4, 5, 6, 7, 8 hrs post addition of doxycycline, I collected cell lysate in the same manner.

Western blotting

After vigorous scraping with a cell scraper (Corning, #29442), I transferred cell lysates to 1.7 mL microcentrifuge tubes (Fisher, #C-3260-1) and heated them to 95 °C for 10 min. 12 μL of each lysate was loaded onto 4–20% SDS-PAGE gradient gels (Bio-Rad, #4561093) and run at 100 V. Any remaining lysate was stored at -20 °C. Once the dye front reached the bottom, protein was transferred to 0.2 μm nitrocellulose membranes (BioRad, #1620112) at 30 V overnight at 4 °C. I blocked blots in a solution of 5% powdered skim milk in TBST (20 mM Tris [Tris Base with added HCl, Fisher, #BP152-1], 150 mM NaCl [Fisher, #BP358-212], 0.1% Tween-20 in H_2O, pH7.4) for 1 hr

at room temperature. Following, I incubated blots in primary antibody: rabbit-anti-V5 (Novus Biologicals, #NB600-381) diluted to 200 ng/mL in TBST, and mouse-anti-GAPDH (GeneTex, #GTX627408) diluted to 400 ng/mL in TBST. After a 1 hr incubation at room temperature and 2, 15-min washes in TBST, I incubate blots in secondary antibodies: donkey-anti-rabbit conjugated to Alexa 790 (ThermoFisher, #A11374) and donkey-anti-mouse conjugated to Alexa 680 (ThermoFisher, #A10038), both diluted to 200 ng/mL in 5% milk in TBST. Blots were incubated in the dark for 45 min at room temperature, washed twice for 15 min in TBST at room temperature, then imaged on a near-infrared blot imager (LI-COR Odyssey).

DNA dot blot

For m6A- samples, I harvested the genomic DNA from hTERT RPE-1 TO cells using a miniprep kit according to manufacturer's instructions (Zymogen Quick-DNA miniprep kit, #D3025). For m6A+ cells, I harvested the genomic DNA of RPE-1 Dam-V5 cells treated with 62.5–125 ng/mL doxycycline for 24 hours. m6A status was checked by restriction-digest PCR (see below). 100 ng DNA from each condition was brought to a final volume of 200 µL in 6x SSC (made in-house with NaCl and NaCitrate: Millipore, #567446) and incubated at 95 °C for 10 min. Another 200 µL of 20x SSC was then added and then solutions placed on ice. I incubated a 0.2 µm nitrocellulose membrane and a piece of Whatman filter paper (Fisher, #09925144) in 20x SSC for 10 min. The membrane and Whatman filter paper were placed inside of a dot-blot manifold I created with laser-cut acrylic and acrylic cement, with a hole drilled to place a vacuum tube. Vacuum was turned on and 400 µL 20x SSC was added to each well and allowed to filter through the membrane and Whatman paper. I then added DNA samples 50–100 µL at a time to each well, with unused wells filled with 20x SSC. Once all sample was adsorbed to the membrane, I saturated another piece of Whatman paper in a denaturation solution of 1.5 M NaCl and 0.5 M NaOH (Sigma, #221465) in diH_2O, then placed the blot DNA-side up on top of the saturated filter paper and let sit for 10 min, covered. I then moved the blot to another piece of Whatman paper saturated with a neutralization solution of 1 M NaCl and 0.5 M Tris, pH 7.0, in diH_2O for another 10 min, covered. I then incubated the blot on a second

piece of filter paper saturated with neutralization solution for 5 min, covered. The blot was then moved to a dry piece of Whatman paper and allowed to dry, covered. I then placed the blot between two pieces of dry Whatman paper and baked under vacuum for 2 hrs at 80 °C. Following baking, I blocked the blot in PBS-BT for 30 min at room temperature. I then incubated the blot in primary antibody: rabbit-anti-m6A (Synaptic Systems, #202 003) diluted to 500 ng/mL in PBS-BT for 30 min at room temperature. Following 2 washes in PBS-BT for 15 min each, I incubated in donkey-anti-rabbit antibody conjugated to Alexa Fluor 790, diluted to 200 ng/mL in PBS-BT, for 30 min in the dark at room temperature. Following 2 additional 15-min washes in PBS-BT, the blot was then imaged on a near-infrared blot imager.

Restriction-digest PCR

After incubating RPE-1 TO cells or RPE-1 Dam-V5 cells in 0–500 ng/mL doxycycline, I extracted genomic DNA. I then performed a mock digestion or digested with DpnII (NEB, #R0543) for 4 hours at 37 °C. I then performed a touchdown PCR for *LMNA* using forward primer 5'- GAAAGGGGATCCCCTGCCTACCTGACCCTCTC-3' and reverse primer 5'-GAAAGGGAATTCCCTCTCCCAGAGACCACCAC-3'. Primers are not added until after a 2 min incubation at 95 °C, then the reaction is cycled 16 times with an annealing temperature starting at 70 °C and decreasing 1 °C/cycle. Finally, 30 cycles are performed by a standard PCR protocol with an annealing temperature of 66 °C. An extension time of 1 min 47 s was used for all cycles.

m6A immunofluorescence with formamide denaturation

I plated 300k RPE-1 Dam-4xV5 cells into a 6-well dish onto lysine-coated coverslips +/- 500 ng/mL doxycycline. 2 days following, I washed cells in PBS then fixed in 100% MeOH at -20 °C for 10 min. I washed the cells 2 times in PBS-BT for 5 min at room temperature, then incubated cells in 1 μg/mL RNase A (Sigma, #10109142001) in PBS-BT for 1 hr at 37 °C to remove any m6A-decorated RNA. Following 2 washes in 2xSSC for 5 min at room temperature, I then incubated cells in 0.2 M HCl and 2% TritonX-100 in H_2O for 10 min at room temperature, followed by 2 more washes in 2xSSC. I

then incubated coverslips in a 1:1 solution of 2xSSC and formamide for 1 hr at 74 ºC. Immediately following incubation, I washed with ice-cold PBS, then incubated in primary antibody overnight at 4 ºC: rabbit-anti-m6A that had been diluted in PBS to 2 μg/mL and chilled on ice. The next day, I washed coverslips 3 times in PBS-BT, then incubated in goat-anti-rabbit conjugated to Alexa Fluor 488 diluted to 1 μg/mL in PBS-BT for 30 min at room temperature in the dark. After washing 3 times in PBS-BT for 5 min at room temperature, I incubated coverslips in 1 μg/mL DAPI then rinsed once more in PBS before mounting on glass slides with VectaShield anti-fade medium and sealed with nail polish. Z-stacks were imaged on a Leica point scanning microscope (Leica SPE) with a 40x objective. Images were taken on the same day under the same imaging conditions. Individual slices are shown.

m6A immunofluorescence with DpnI digestion

I plated 450k RPE-1 Dam-4xV5/m6A-tracer cells into 2 wells of a 6-well dish onto lysine-coated coverslips, +/- 16 ng/mL doxycycline and 1 μM Shield1. The following day, I washed cells in PBS then fixed in -20 ºC MeOH for 10 min, then washed again in PBS. I RNase-treated and blocked coverslips in 2 μg/mL RNase A in PBS-BT for 30 min at 37 ºC, then permeabilized cells in 2% TritonX-100 in PBS for 10 min at room temperature. I then incubated the coverslips in mouse-anti-H3K27-acetyl diluted to 2 μg/mL (Active Motif, #39685) in PBS for 30 min at room temperature. Following 3 more washes in PBS-BT, I incubated coverslips in goat-anti-mouse conjugated to Alexa Fluor 568 diluted to 1 μg/mL for 30 min in the dark. While Dam carries a TagRFP fluorophore, methanol fixation quenches fluorophore excitation sufficiently to distinguish acetylated histone over any Dam signal. I then refixed coverslips in 4% PFA (Electron Microscopy Sciences, #15710) in PBS for 5 min at room temperature, then washed in PBS-BT. I incubated cells overnight at 37 ºC in 50 U DpnI in 1x CutSmart buffer (NEB, #B7204). The following day, I washed 3 times in PBS-BT for 5 min at room temperature, then performed immunofluorescence for m6A using rabbit-anti-m6A diluted to 2 μg/mL and donkey-anti-rabbit conjugated to Alexa Fluor 647 diluted to a concentration of 4 μg/mL. Finally, I rinsed cells in 1 μg/mL DAPI and

mounted coverslips on glass slides with VectaShield sealed with nail polish. Cells were imaged on a Leica SPE at 40x. A single slice is displayed.

m6A and m6A-tracer/m6A-wH colocalization

I plated 150k RPE-1 Dam-4xV5 cells/well transduced with either m6A-tracer or m6A-wH into a 6-well dish onto lysine-coated coverslips. The following day, I replaced media with media supplemented with 100 nM Mps1i. 20 hours later, I replaced with fresh media supplemented with 100 nM Mps1i +/- 16 ng/mL doxycycline and 1 μM Shield1. 8 hours later, I fixed coverslips in -20 °C 100% MeOH for 10 min, then rehydrated in PBS. Coverslips were then blocked and RNase-treated in 2 μg/mL RNase A in PBS-BT for 30 min at 37 °C, then washed twice with PBS-BT for 5 min each at room temperature. To recover quenched m6A-tracer and -wH fluorophore signal, I stained for GFP using chicken-anti-GFP (Fisher, # CGFP-45ALY) diluted to 1 μg/mL in PBS followed by goat-anti-chicken conjugated to Alexa Fluor 488 diluted to 2 μg/mL in PBS-BT. I then refixed coverslips in 4% PFA in PBS at room temperature, washed twice with PBS-BT for 5 min each at room temperature, and permeabilized cells in a 2% TritonX-100 solution in PBS at room temperature for 10 min. I then incubated cells in 50 U/mL DpnI in 1x CutSmart buffer overnight at 37 °C. After incubation, I performed immunofluorescence using rabbit-anti-m6A with Alexa Fluor 647-conjugated 2° antibodies as described above, before washing once with 1 μg/mL DAPI followed by PBS, then mounting in VectaShield onto glass slides sealed with nail polish. Cells were imaged at 40x on a Leica SPE on the same day under identical imaging conditions. Individual slices are shown.

Colocalization beween m6A and m6A-tracer or m6A-wH was performed following Li's intensity correlation quotient using FIJI 2.9.0/1.53t and the JACoP plugin 2.0.[103] Colocalization was performed on full z-stacks comparing m6A and m6A-tracer/-wH channels for the images shown in **Figs 2-3c** and **2-3d**.

m6A-tracer and ruptured MN colocalization

I plated 150k RPE-1 Dam-4xV5/m6A-tracer cells into a 6-well dish onto lysine-coated coverslips. The following day, I replaced media with media supplemented with 100 nM Mps1i. 20 hours later, I replaced with fresh media supplemented with 100 nM Mps1i +/- 16 ng/mL doxycycline and 1 μM Shield1. 8 hours later, I fixed coverslips in 4% PFA in PBS for 5 min at room temperature. I then performed immunofluorescence as described above using rabbit-anti-H3K(9/14/18/23/27)-acetyl (Abcam, #ab47915) diluted 1:500 in PBS with donkey-anti-rabbit conjugated to Alexa Fluor 647 2° antibody diluted to 4 μg/mL in PBS-BT and mouse-anti-V5 (Fisher, #377500) diluted to 1 μg/mL in PBS with donkey-anti-mouse conjugated to Alexa Fluor 564 2° antibody diluted to 2 μg/mL in PBS-BT. After a rinse in 1 μg/mL DAPI followed by a rinse in PBS, I mounted coverslips with VectaShield onto glass slides sealed with nail polish. Images were taken on a 40x objective on a Leica SPE. Single slices shown. Cells were manually classified as V5+/- and nuclei classified as m6A-tracer+/-. MN were likewise categorized manually as V5+/-, m6A-tracer+/-, and H3-acetyl+/-. For the analysis shown in **Fig 2-4**, only MN from V5+ cells with m6A-tracer- nuclei were included, to exclude cells without Dam expression and cells that had undergone mitosis while Dam was active.

U2OS splitable Dam transfection

I plated U2OS cells transduced with the TetOn transactivator in a 12-well plate. The following day, each well was transfected with one of the following Dam constructs: mock (empty vector with blasticidin resistance), DD-NES-Dam-4xV5-TagRFP (Dam-4xV5), or a split Dam construct linked with a TVMVP recognition motif, DD-NES-SpDam-4xV5-emiRFP703 (SpDam-4xV5). Additionally, each well was transfected with either empty vector (with puromycin resistance) or TVMVP. Transfection was done using the standard Lipofectamine 2000 transfection protocol (Invitrogen, #11668019) in antibiotic-free media. 4 hours after transfection, cells were washed 2x in PBS and then incubated in tet-free media with penicillin and streptomycin. 23 hours following, media was replaced with fresh tet-free media supplemented with 500 ng/mL doxycycline, 10 μg/mL blasticidin, and 10 μg/mL puromycin. The following

day, cells were washed with PBS, cell lysate collected, and Western blots performed as described.

Statistical analyses

Shorthand p-values are as follows:

ns: p-value >= 0.05

*: p-value < 0.05

**: p-value < 0.01

***: p-value < 0.001

****: p-value < 0.0001

Comparison of m6A-tracer+/- MN was performed by Barnard's exact test in R using the Barnard package version 1.8.

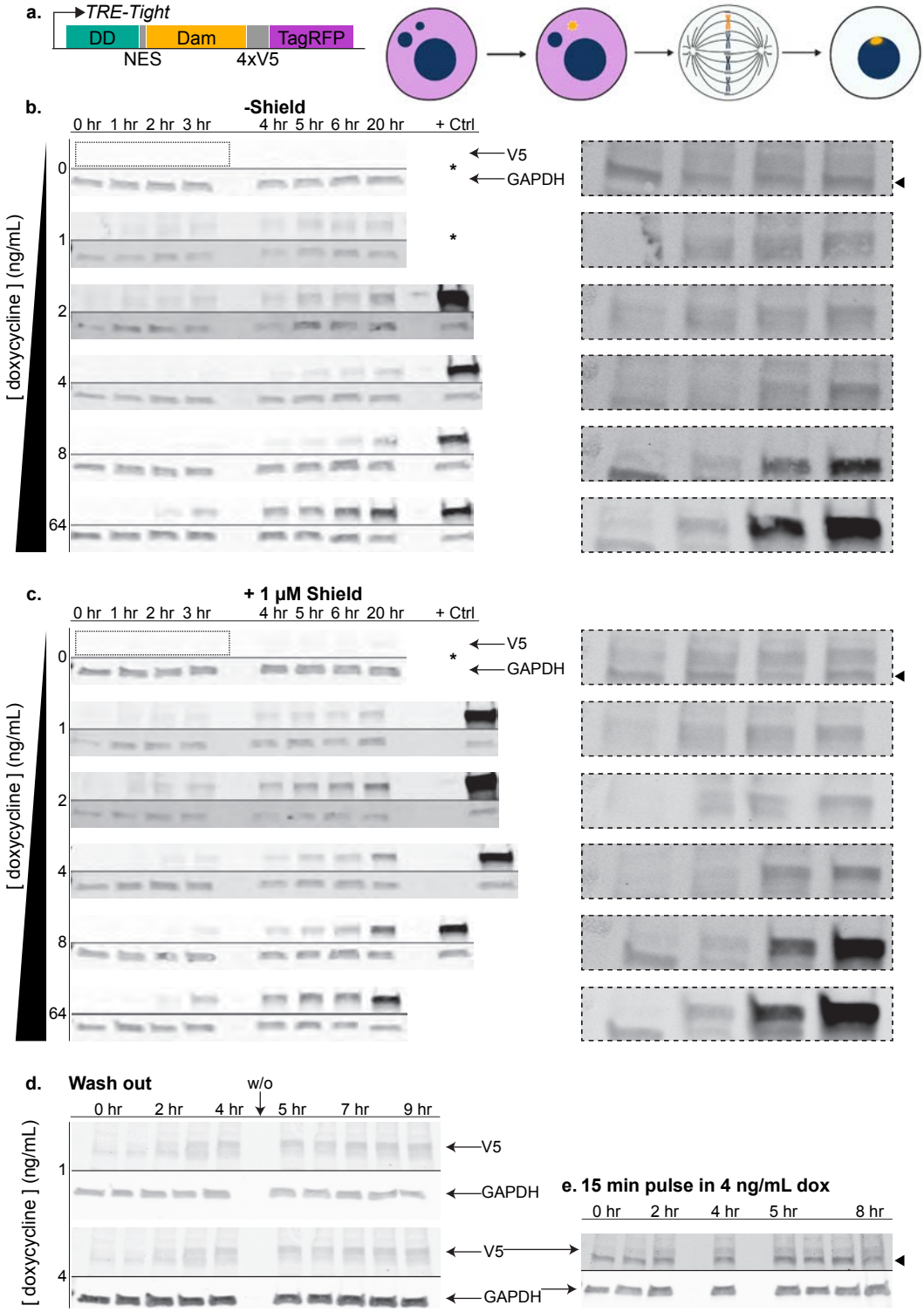

Figure 1: Dam can be rapidly induced by doxycycline, but is degraded very slowly
a) Diagrams of construct transduced into hTERT RPE-1 cells and how it can be used to mark and track chromatin over time. Graphic created with BioRender.com. **b)** Western blot of a timecourse tracking V5 expression over time after addition of doxycycline, absent Shield1. Positive control is lysate taken from cells treated with 64 ng/mL doxycycline and 1 µM Shield1 for 20 hrs. Asteriskss mark where no positive control was used. Dashed square shows zoomed in region of blot, shown on the right with increased contrast to visualize low levels of V5. Arrowhead marks non-specific band. **c)** Same as b), but with addition of 1 µM Shield1. **d)** Time course tracking V5 expression over time after addition of doxycycline and 1 µM Shield1. After 4 hrs, cells were washed 3x with PBS and replaced with doxycycline-free media. **e)** Time course tracking V5 expression after incubation in 4 ng/mL doxycycline with no Shield1 for 15 min followed by 3x PBS washout. Arrowhead marks non-specific band.

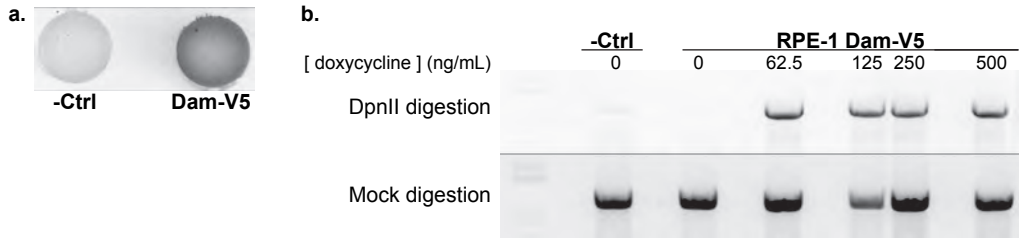

Figure 2: Dam expression induces m6A
a) DNA dot blots for m6A of genomic DNA sampled from RPE-1 TO cells (-Ctrl) or Dam-V5 cells treated with 62.5–125 ng/mL doxycycline for 24 hrs. **b)** Restriction digest PCR of genomic DNA sampled from RPE-1 TO cells (-Ctrl) or Dam-V5 cells treated with varying levels of doxycycline for 24 hrs. DNA was subject to mock or DpnII digestion before running a PCR for the *LMNA* locus.

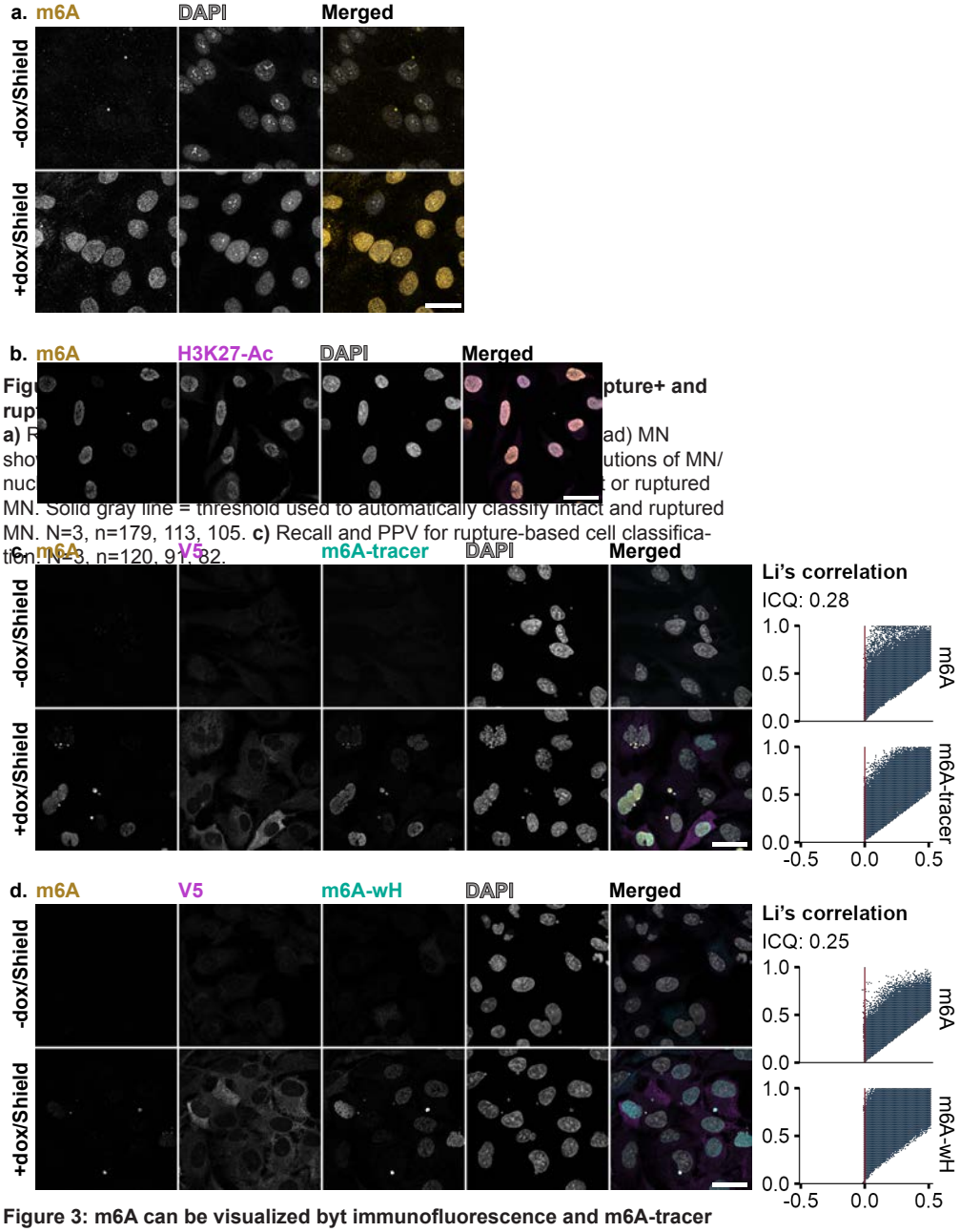

Figure 3: m6A can be visualized byt immunofluorescence and m6A-tracer
a) Immunofluorescence of m6A in Dam-4xV5 cells treated without doxycycline or with 500 ng/mL doxycycline + 1 µM Shield1. DNA denaturation done with heat and formamide. **b)** Gentle immunofluorescence of m6A and acetylated histone 3 lysine 27 (H3K27-Ac). DNA denaturation done with DpnI digestion. **c)** Comparison of m6A signal and m6A-tracer in Dam-4xV5 cells expressing m6A-tracer with and without doxycycline/Shield1 treatment. **d)** Same as c) but in cells expressing m6A-wH.

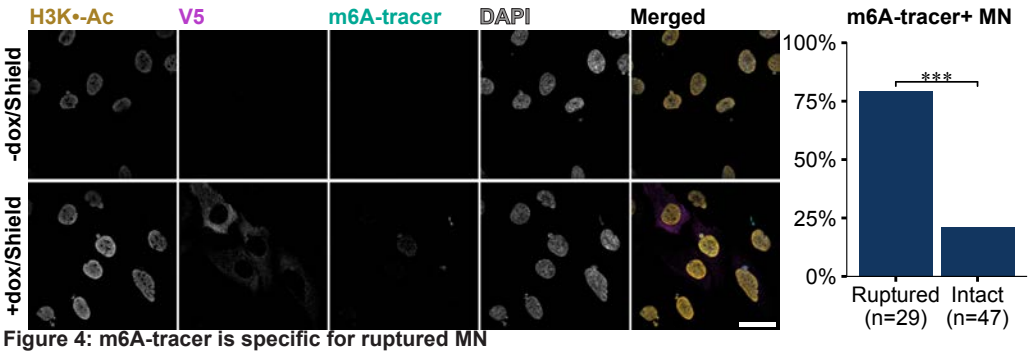

Figure 4: m6A-tracer is specific for ruptured MN
Representative image of immunofluorescence of Dam-4xV5 cells expressing m6A-tracer with and without doxycycline/Shield treatment. MN rupture status determined by manual inspection of acetylated histone 3 lysines (H3K•-Ac). Proportions of m6A-tracer+ MN in V5+ cells with m6A-tracer- nuclei on the right.

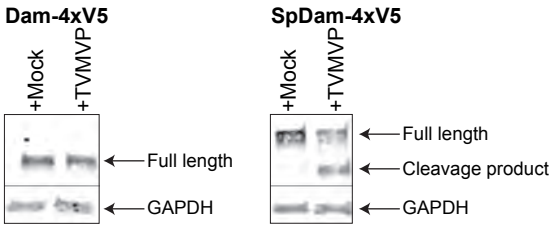

Figure 5: Splittable Dam construct can be cleaved by TVMVP
Western blot of lysate from U2OS TO cells transfected with either DD-NES-Dam-4xV5-TagRFP (Dam-4xV5) or DD-NES-SpDam-4xV5-emiRFP703 (SpDam-4xV5) and either an empty plasmid (mock) or TVMV protease (TVMVP) and incubated in 500 ng/mL doxycycline for 24 hours.

Figure S3: Dendra ratios persist for many hours
a) The Dendra-red/Dendra-green ratios from images taken post-photoconversion, pre-FACS during the experiment in Fig 4 showing 3 discernible populations even after 8 hours. Representative images from each time point are color-coded by Dendra ratio.

Conclusion

For decades, micronuclei were seen as the passive indicator of exposure to genotoxins or of CIN. In recent years, research has intensified as it was discovered that micronuclei were not merely symptoms but causes of CIN as well as other processes, including inflammation, oncogenesis, and chromothripsis. Because they arise through mechanisms that can also produce other, overlapping phenotypes, micronuclei can only be identified through visual inspection. This has historically forced researchers to adopt hypothesis-driven, labor-intensive approaches to study, limiting the ability to perform unbiased screens or conduct hypothesis discovery. While this work does not fundamentally change the fact that micronuclei can only be identified through visual inspection, we were able to tap into another area of intense research—deep neural networks and artificial intelligence—to reduce the labor required. The power of neural networks to not only parse visual information but to even create works of art has exploded in the last few years. By using a computer to automate the identification of micronucleated cells, we can for the first time isolate micronucleated cells and cells that have experienced micronucleus rupture regardless of which chromosomes micronucleate and, to a large extent, without the need for a specific cell line—only for that cell line to express two simple constructs, easily integrated through transfection or transduction.

While the Dam Tracker tool alone would not have helped with the identification of cells *en masse*, it could have allowed us to identify *chromatin* that had experienced micronucleus rupture *en masse*, using ChIP. It is also possible, using specialized FACS that could sort punctate GFP signal—*ie*, that arising from m6A-tracer marking a ruptured micronucleus—from diffuse GFP signal that we would be able to sort cells with ruptured micronuclei. The problem of mitosis is still a challenging one, though solvable as the Pellman Lab was able to do. I have hopes that a protease-based method for cell cycle control could work and would obviate the need to synchronize cells or use cell cycle inhibitors.

My hope is that this work will enable future researchers to understand the full consequences of micronucleation and micronucleus rupture. And, I hope it will inspire future researchers to explore the possibilities of integrating computational and wet lab biology in a single project.

Acknowledgements

This work was supported by a National Institutes of Health grant (R35GM124766-02, awarded to E.M. Hatch), a National Human Genome Research Institute grant (RM1HG010461, awarded to D.M. Fowler), a training grant (T32CA009657, awarded to L. DiPeso), the Rita Allen Foundation Scholars program, and the Fred Hutch Cancer Center Bioinformatics and Genomics cores (funded by NIH grant P39CA015704).

Bibliography

1. Guo, X. *et al.* The molecular origins and pathophysiological consequences of micronuclei: New insights into an age-old problem. *Mutat. Res. - Rev. Mutat. Res.* **779**, 1–35 (2019).
2. Howell, W. The life history of the formed elements of the blood, especially the red blood corpuscles. *J. Morphol.* **4**, 57–116 (1890).
3. Kirsch-Volders, M., Bolognesi, C., Ceppi, M., Bruzzone, M. & Fenech, M. Micronuclei, inflammation and auto-immune disease. *Mutat. Res. Rev. Mutat. Res.* **786**, 108335 (2020).
4. MacDonald, K. M., Benguerfi, S. & Harding, S. M. Alerting the immune system to DNA damage: micronuclei as mediators. *Essays Biochem.* **64**, 753–764 (2020).
5. Mackenzie, K. J. *et al.* cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**, 461–465 (2017).
6. Vázquez-Diez, C., Yamagata, K., Trivedi, S., Haverfield, J. & FitzHarris, G. Micronucleus formation causes perpetual unilateral chromosome inheritance in mouse embryos. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 626–631 (2016).
7. Zhang, C.-Z. *et al.* Chromothripsis from DNA damage in micronuclei. *Nature* **522**, 179–184 (2015).
8. Boettcher, B. & Barral, Y. The cell biology of open and closed mitosis. *Nucleus* **4**, 160–165 (2013).
9. Liu, S. & Pellman, D. The coordination of nuclear envelope assembly and chromosome segregation in metazoans. *Nucleus* **11**, 35–52 (2020).
10. Maciejowski, J., Li, Y., Bosco, N., Campbell, P. J. & de Lange, T. Chromothripsis and Kataegis Induced by Telomere Crisis. *Cell* **163**, 1641–1654 (2015).
11. Hoffelder, D. R. *et al.* Resolution of anaphase bridges in cancer cells. *Chromosoma* **112**, 389–397 (2004).
12. Rao, X. *et al.* Multiple origins of spontaneously arising micronuclei in HeLa cells: Direct evidence from long-term live cell imaging. *Mutat. Res. Mol. Mech. Mutagen.* **646**, 41–49 (2008).

13. Sepaniac, L. A. *et al.* Micronuclei in Kif18a mutant mice form stable micronuclear envelopes and do not promote tumorigenesis. *J. Cell Biol.* **220**, e202101165 (2021).
14. Fenech, M. *et al.* Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis* **26**, 125–132 (2011).
15. Hatch, E. M., Fischer, A. H., Deerinck, T. J. & Hetzer, M. W. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell* **154**, 47–60 (2013).
16. Mammel, A. E., Huang, H. Z., Gunn, A. L., Choo, E. & Hatch, E. M. Chromosome length and gene density contribute to micronuclear membrane stability. *Life Sci. Alliance* **5**, (2022).
17. Liu, S. *et al.* Nuclear envelope assembly defects link mitotic errors to chromothripsis. *Nature* **561**, 551–555 (2018).
18. Guo, X., Dai, X., Wu, X., Cao, N. & Wang, X. Small but strong: Mutational and functional landscapes of micronuclei in cancer genomes. *Int. J. Cancer* **148**, 812–824 (2021).
19. Hatch, E. M. & Hetzer, M. W. Nuclear envelope rupture is induced by actin-based nucleus confinement. *J. Cell Biol.* **215**, 27–36 (2016).
20. Denais, C. M. *et al.* Nuclear envelope rupture and repair during cancer cell migration. *Science* **352**, 353–8 (2016).
21. Olmos, Y., Hodgson, L., Mantell, J., Verkade, P. & Carlton, J. G. ESCRT-III controls nuclear envelope reformation. *Nature* **522**, 236–239 (2015).
22. Raab, M. *et al.* ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. *Science* **352**, 359–62 (2016).
23. Vietri, M. *et al.* Unrestrained ESCRT-III drives micronuclear catastrophe and chromosome fragmentation. *Nat. Cell Biol.* **22**, 856–867 (2020).
24. Willan, J. *et al.* ESCRT-III is necessary for the integrity of the nuclear envelope in micronuclei but is aberrant at ruptured micronuclear envelopes generating damage. *Oncogenesis* **8**, 1–14 (2019).
25. Flynn, P. J., Koch, P. D. & Mitchison, T. J. Chromatin bridges, not micronuclei, activate cGAS after drug-induced mitotic errors in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2103585118 (2021).

26. MacDonald, K. M. *et al.* Antecedent chromatin organization determines cGAS recruitment to ruptured micronuclei. *Nat. Commun.* **14**, 556 (2023).
27. Mohr, L. *et al.* ER-directed TREX1 limits cGAS activation at micronuclei. *Mol. Cell* **81**, 724-738.e9 (2021).
28. Nader, G. P. de F. *et al.* Compromised nuclear envelope integrity drives TREX1-dependent DNA damage and tumor cell invasion. *Cell* **184**, 5230-5246.e22 (2021).
29. MacDonald, K. M., Nicholson-Puthenveedu, S., Tageldein, M. M., Arrowsmith, C. & Harding, S. M. cGAS recruitment to micronuclei is dictated by pre-existing nuclear chromatin status. 2022.01.13.476191 Preprint at https://doi.org/10.1101/2022.01.13.476191 (2022).
30. Volkman, H. E., Cambier, S., Gray, E. E. & Stetson, D. B. Tight nuclear tethering of cGAS is essential for preventing autoreactivity. *eLife* **8**, e47491 (2019).
31. Terzoudi, G. I. *et al.* Stress induced by premature chromatin condensation triggers chromosome shattering and chromothripsis at DNA sites still replicating in micronuclei or multinucleate cells when primary nuclei enter mitosis. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **793**, 185–198 (2015).
32. Ly, P. *et al.* Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. *Nat. Cell Biol.* **19**, 68–75 (2017).
33. Papathanasiou, S. *et al. Transgenerational transcriptional heterogeneity from cytoplasmic chromatin*. http://biorxiv.org/lookup/doi/10.1101/2022.01.12.475869 (2022) doi:10.1101/2022.01.12.475869.
34. Luijten, M. N. H., Lee, J. X. T. & Crasta, K. C. Mutational game changer: Chromothripsis and its emerging relevance to cancer. *Mutat. Res. Mutat. Res.* **777**, 29–51 (2018).
35. Asanami, S., Shimono, K. & Kaneda, S. Transient hypothermia induces micronuclei in mice. *Mutat. Res. Toxicol. Environ. Mutagen.* **413**, 7–14 (1998).
36. Hintzsche, H., Riese, T. & Stopper, H. Hyperthermia-induced micronucleus formation in a human keratinocyte cell line. *Mutat. Res. Mol. Mech. Mutagen.* **738**–**739**, 71–74 (2012).

37. Maass, K. K. *et al.* Altered nuclear envelope structure and proteasome function of micronuclei. *Exp. Cell Res.* **371**, 353–363 (2018).
38. Santaguida, S., Tighe, A., D'Alise, A. M., Taylor, S. S. & Musacchio, A. Dissecting the role of MPS1 in chromosome biorientation and the spindle checkpoint through the small molecule inhibitor reversine. *J. Cell Biol.* **190**, 73–87 (2010).
39. Crasta, K. *et al.* DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **482**, 53–58 (2012).
40. Dertinger, S. D., Torous, D. K., Hayashi, M. & MacGregor, J. T. Flow cytometric scoring of micronucleated erythrocytes: an efficient platform for assessing in vivo cytogenetic damage. *Mutagenesis* **26**, 139–145 (2011).
41. Hutter, K.-J. & Stöhr, M. Rapid detection of mutagen induced micronucleated erythrocytes by flow cytometry. *Histochemistry* **75**, 353–362 (1982).
42. Hasle, N. *et al.* High-throughput, microscope-based sorting to dissect cellular heterogeneity. *Mol. Syst. Biol.* **16**, e9442 (2020).
43. hatch-lab/mnfinder: MN segmentation neural net. https://github.com/hatch-lab/mnfinder.
44. He, Q. *et al.* Chromosomal instability-induced senescence potentiates cell non-autonomous tumourigenic effects. *Oncogenesis* **7**, 1–18 (2018).
45. Santaguida, S. *et al.* Chromosome Mis-segregation Generates Cell-Cycle-Arrested Cells with Complex Karyotypes that Are Eliminated by the Immune System. *Dev. Cell* **41**, 638-651.e5 (2017).
46. Dürrbaum, M. *et al.* Unique features of the transcriptional response to model aneuploidy in human cells. *BMC Genomics* **15**, 139 (2014).
47. Matlashov, M. E. *et al.* A set of monomeric near-infrared fluorescent proteins for multicolor imaging across scales. *Nat. Commun.* **11**, 239 (2020).
48. Lopez-Urrutia, A. Deep Retina 3th place solution to Kaggle's 2018 Data Science Bowl. (2022).
49. Vassilev, L. T. Cell Cycle Synchronization at the G2/M Phase Border by Reversible Inhibition of CDK1. *Cell Cycle* **5**, 2555–2556 (2006).

50. Vereecke, L., Beyaert, R. & van Loo, G. The ubiquitin-editing enzyme A20 (TNFAIP3) is a central regulator of immunopathology. *Trends Immunol.* **30**, 383–391 (2009).
51. Verstrepen, L. *et al.* Expression, biological activities and mechanisms of action of A20 (TNFAIP3). *Biochem. Pharmacol.* **80**, 2009–2020 (2010).
52. Baron, V., Adamson, E. D., Calogero, A., Ragona, G. & Mercola, D. The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGFβ1, PTEN, p53, and fibronectin. *Cancer Gene Ther.* **13**, 115–124 (2006).
53. Wang, B. *et al.* The Role of the Transcription Factor EGR1 in Cancer. *Front. Oncol.* **11**, (2021).
54. Rocca, A., Farolfi, A., Bravaccini, S., Schirone, A. & Amadori, D. Palbociclib (PD 0332991): targeting the cell cycle machinery in breast cancer. *Expert Opin. Pharmacother.* **15**, 407–420 (2014).
55. Hai, T., Wolford, C. C. & Chang, Y.-S. ATF3, a Hub of the Cellular Adaptive-Response Network, in the Pathogenesis of Diseases: Is Modulation of Inflammation a Unifying Component? *Gene Expr.* **15**, 1–11 (2018).
56. Ku, H.-C. & Cheng, C.-F. Master Regulator Activating Transcription Factor 3 (ATF3) in Metabolic Homeostasis and Cancer. *Front. Endocrinol.* **11**, (2020).
57. Hasim, M. S., Nessim, C., Villeneuve, P. J., Vanderhyden, B. C. & Dimitroulakos, J. Activating Transcription Factor 3 as a Novel Regulator of Chemotherapy Response in Breast Cancer. *Transl. Oncol.* **11**, 988–998 (2018).
58. Dickey, J. S. *et al.* H2AX: functional roles and potential applications. *Chromosoma* **118**, 683–692 (2009).
59. Nikolova, T. *et al.* The γH2AX Assay for Genotoxic and Nongenotoxic Agents: Comparison of H2AX Phosphorylation with Cell Death Response. *Toxicol. Sci.* **140**, 103–117 (2014).
60. Turinetto, V. & Giachino, C. Multiple facets of histone variant H2AX: a DNA double-strand-break marker with several biological functions. *Nucleic Acids Res.* **43**, 2489–2498 (2015).
61. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, (2021).

62. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS J. Integr. Biol.* **16**, 284–287 (2012).
63. Lin, T.-Y., Goyal, P., Girshick, R., He, K. & Dollar, P. Focal Loss for Dense Object Detection. in 2980–2988 (2017).
64. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. Cell Detection with Star-Convex Polygons. in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (eds. Frangi, A. F., Schnabel, J. A., Davatzikos, C., Alberola-López, C. & Fichtinger, G.) 265–273 (Springer International Publishing, 2018). doi:10.1007/978-3-030-00934-2_30.
65. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (eds. Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F.) 234–241 (Springer International Publishing, 2015). doi:10.1007/978-3-319-24574-4_28.
66. Frank, Hall, M. A., Witten, I. H. & Pal, C. J. The WEKA Workbench. in *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, 2016).
67. Cohen, W. W. Fast Effective Rule Induction. in 115–123 (1995).
68. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
70. Love, M. I. *et al.* Tximeta: Reference sequence checksums for provenance identification in RNA-seq. *PLOS Comput. Biol.* **16**, e1007664 (2020).
71. Stephens, M. False discovery rates: a new deal. *Biostatistics* **18**, 275–294 (2017).
72. Korotkevich, G. *et al.* Fast gene set enrichment analysis. 060012 Preprint at https://doi.org/10.1101/060012 (2021).
73. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

74. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
75. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
76. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
77. The Gene Ontology Consortium *et al.* The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).
78. Greene, D., Richardson, S. & Turro, E. ontologyX: a suite of R packages for working with ontological data. *Bioinformatics* **33**, 1104–1106 (2017).
79. Crameri, F. Geodynamic diagnostics, scientific visualisation and StagLab 3.0. *Geosci. Model Dev.* **11**, 2541–2562 (2018).
80. Højsgaard, S., Halekoh, U. & Yan, J. The R Package geepack for Generalized Estimating Equations. *J. Stat. Softw.* **15**, 1–11 (2006).
81. Aughey, G. N. & Southall, T. D. Dam it's good! DamID profiling of protein-DNA interactions. *WIREs Dev. Biol.* **5**, 25–37 (2016).
82. Palmer, B. R. & Marinus, M. G. The dam and dcm strains of Escherichia coli--a review. *Gene* **143**, 1–12 (1994).
83. Douvlataniotis, K., Bensberg, M., Lentini, A., Gylemo, B. & Nestor, C. E. No evidence for DNA N6-methyladenine in mammals. *Sci. Adv.* **6**, eaay3335 (2020).
84. Liu, X. *et al.* N6-methyladenine is incorporated into mammalian genome by DNA polymerase. *Cell Res.* **31**, 94–97 (2021).
85. Musheev, M. U., Baumgärtner, A., Krebs, L. & Niehrs, C. The origin of genomic N6-methyl-deoxyadenosine in mammalian cells. *Nat. Chem. Biol.* **16**, 630–634 (2020).
86. Xiao, C.-L., Zhu, S., Gu, X.-F., Wang, K. & Yan, G.-R. N 6-Methyladenine DNA Modification in the Human Genome. *Mol. Cell* **71**, 306–318 (2018).
87. Xie, Q. *et al.* N6-methyladenine DNA Modification in Glioblastoma. *Cell* **175**, 1228-1243.e20 (2018).
88. Greil, F., Moorman, C. & van Steensel, B. DamID: Mapping of In Vivo Protein–Genome Interactions Using Tethered DNA Adenine Methyltransferase. *Methods Enzymol.* **410**, 342–359 (2006).

89. van Steensel, B., Delrow, J. & Henikoff, S. Chromatin profiling using targeted DNA adenine methyltransferase. *Nat. Genet.* **27**, 304–308 (2001).
90. Gottschling, D. E. Telomere-proximal DNA in Saccharomyces cerevisiae is refractory to methyltransferase activity in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4062–4065 (1992).
91. Wu, F., Olson, B. G. & Yao, J. DamID-seq: Genome-wide Mapping of Protein-DNA Interactions by High Throughput Sequencing of Adenine-methylated DNA Fragments. *J. Vis. Exp. JoVE* e53620–e53620 (2016) doi:10.3791/53620.
92. Kind, J. *et al.* Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. *Cell* **153**, 178–192 (2013).
93. Mierzejewska, K. *et al.* Structural basis of the methylation specificity of R.DpnI. *Nucleic Acids Res.* **42**, 8745 (2014).
94. Egeler, E. L., Urner, L. M., Rakhit, R., Liu, C. W. & Wandless, T. J. Ligand-switchable Substrates for a Ubiquitin-Proteasome System *. *J. Biol. Chem.* **286**, 31328–31336 (2011).
95. Li, Q. *et al.* A syntaxin 1, Galpha(o), and N-type calcium channel complex at a presynaptic nerve terminal: analysis by quantitative immunocolocalization. *J. Neurosci. Off. J. Soc. Neurosci.* **24**, 4070–4081 (2004).
96. Fernandez-Rodriguez, J. & Voigt, C. A. Post-translational control of genetic circuits using Potyvirus proteases. *Nucleic Acids Res.* **44**, 6493–6502 (2016).
97. Fink, T. *et al.* Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat. Chem. Biol.* **15**, 115–122 (2019).
98. Gao, X. J., Chong, L. S., Kim, M. S. & Elowitz, M. B. Programmable protein circuits in living cells. *Science* **361**, 1252–1258 (2018).
99. Szczesnik, T., Ho, J. W. K. & Sherwood, R. Dam mutants provide improved sensitivity and spatial resolution for profiling transcription factor binding. *Epigenetics Chromatin* **12**, 36–36 (2019).
100. Matthew Hass, A. R., Liow, H., Chen, X., Weirauch, M. T. & Brent, M. R. SpDamID: Marking DNA Bound by Protein Complexes Identifies Notch-Dimer Responsive Enhancers SpDamID enriches for factor-regulated targets and DNA-binding

motifs d SpDamID pairs for major signaling pathways and knockin SpDamID ESCs available. *Mol. Cell* **59**, 685–697 (2015).

101. Cabantous, S. *et al.* A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association. *Sci. Rep.* **3**, 2854 (2013).
102. Horton, R. M. PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. *Mol. Biotechnol.* **3**, 93–99 (1995).
103. Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).