Published in final edited form as:
Nat Cell Biol. 2021 April 01; 23(4): 401–412. doi:10.1038/s41556-021-00657-2.

Disruption of the MSL complex inhibits tumor maintenance by exacerbating chromosomal instability

Josep Monserrat[1], **Cristina Morales Torres**[1], **Louise Richardson**[1], **Thomas S. Wilson**[1], **Harshil Patel**[2], **Marie-Charlotte Domart**[3], **Stuart Horswell**[2], **Ok-Ryul Song**[4], **Ming Jiang**[4], **Margaret Crawford**[5], **Minh Bui**[6], **Yamini Dalal**[6], **Paola Scaffidi**[1,*]

[1]Cancer Epigenetics Laboratory, Francis Crick Institute, London, NW1 1AT, UK

[2]Bioinformatics and Biostatistics, Francis Crick Institute, London, NW1 1AT, UK

[3]Electron Microscopy, Francis Crick Institute, London, NW1 1AT, UK

[4]High Throughput Screening, Francis Crick Institute, London, NW1 1AT, UK

[5]Advanced Sequencing, Francis Crick Institute, London, NW1 1AT, UK

[6]Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20892, USA

Abstract

Rewiring of cellular programs in malignant cells generates cancer-specific vulnerabilities. Here, using an unbiased screening strategy aimed at identifying non-essential genes required by tumor cells to sustain unlimited proliferative capacity, we identify the Male-Specific Lethal (MSL) acetyltransferase complex as a vulnerability of genetically unstable cancers. We find that disruption of the MSL complex and consequent loss of the associated H4K16ac mark do not substantially alter transcriptional programs, but compromise chromosome integrity and promote chromosomal instability (CIN) that progressively exhausts the proliferative potential of cancer cells through a p53-independent mechanism. This effect is dependent on pre-existing genomic instability and normal cells are insensitive to MSL disruption. Using cell- and patient-derived xenografts from multiple cancer types, we show that excessive CIN induced by MSL disruption inhibits tumor maintenance. Our findings suggest that targeting of MSL may be a valuable means to increase CIN beyond the level tolerated by cancer cells without inducing severe adverse effects in normal tissues.

*Correspondence to: Paola.Scaffidi@crick.ac.uk.

Author Contributions Statement
JM performed most of the experiments and analyses, with help from CMT and PS. LR, TSW and MJ generated Cas9-expressing cell lines. HP performed NGS analysis of the CRISPR-KO screens and RNA-seq. MCD performed scanning electron microscopy. SH performed karyotyping analysis. OS performed high-content microscopy analysis. MC performed low-pass genome sequencing. MB analyzed chromatin fibers under the supervision of YD. PS conceived and supervised the study and wrote the manuscript with input from other authors.

Competing Insterests Statement
The authors have no competing interests.

Europe PMC Funders Author Manuscripts

During cancer development, cells acquire a range of biological properties that allow them to initiate and maintain a growing tumor. Establishment and maintenance of malignant phenotypes require rewiring of cellular programs, which creates cancer-specific dependencies that can be exploited for therapeutic purposes[1]. Unbiased identification of cancer dependencies has typically relied on drop-out viability screening employing pooled lentiviral libraries[1]. This approach identifies genes essential for cell survival or cell cycle progression and, in combination with suitable counter-screens, can reveal cancer-specific, driver-specific[2, 3], or cancer-type-specific[4] dependencies. However, viability-based read-outs cannot identify non-essential proteins that cancer cells depend on to maintain unlimited proliferative capacity, a key cellular property underpinning disease maintenance, which is mechanistically distinct from cell cycle progression [5–8].

Here, we devised a strategy to identify new cancer-specific dependencies combining *in vitro* CRISPR-Cas9 screening to assess long-term proliferative capacity of malignant cells and *in vivo* tumor maintenance assays. Given the therapeutic potential of epigenetic regulators, their broad expression patterns across cancer types and their role in regulating cellular states [9, 10], we focused our search for dependencies on chromatin and DNA modifiers. We report that malignant cells from various cancer types require a functional MSL complex to maintain unlimited proliferative potential, and that MSL disruption inhibits tumor maintenance. Surprisingly, this inhibitory effect is not mediated by changes in cellular differentiation programs, but by the accumulation of chromosomal abnormalities that progressively exhaust cells' proliferative potential. Since MSL disruption has deleterious effects only in cells with a basal level of genomic instability, normal cells are unaffected, making MSL a specific vulnerability of chromosomally unstable cancers.

Results

A strategy to assess cancer cells' proliferative potential

Within individual cancers, only subsets of cells can proliferate indefinitely and maintain tumor growth[11–15]. Because of the challenges associated with isolating these cells from clinically-derived samples and using them for large-scale screening, we opted to begin our investigation using a well-characterized and tractable experimental system and subsequently validate our findings using patient-derived xenograft models of multiple cancer types. We have previously shown that xenografts generated by human *de novo* transformed dermal fibroblasts (TDF) are maintained by a small subset of cells (1-5% of the tumor cell population) characterized by a primitive phenotype and unlimited proliferative potential (Extended Fig. 1a) [16, 17]. The tumor-maintaining cells are marked by the surface antigen Stage Specific Embryonic Antigen 1 (SSEA1), which allows their isolation and characterization (Extended Fig. 1a) [16]. TDF xenografts recapitulate features of hierarchically-organized tumors from patients and have enabled the identification of epigenetic mechanisms relevant for numerous cancer types [16–19].

To comprehensively identify dependencies of tumor-maintaining cells on epigenetic regulators, we designed the CRISPR-Cas9 screen such that we could identify two distinct sets of proteins: (i) factors whose loss results in death or immediate cycle arrest of tumor-maintaining cells, excluding non-specific hits based on a counter-screen performed using an

isogenic, non-tumorigenic cell line (viability arm); (ii) proteins whose knock-out inhibit cell long-term proliferative capacity (unlimited proliferation arm) (Fig. 1a). Whereas reduced fitness within the population of transduced cells enables easy detection of the first class of hits, the unlimited proliferation arm required the development of a more complex assay based on combined phenotypic and functional readouts. In addition to monitoring SSEA1 expression, we also engineered TDF cells such that they express a fluorescent reporter when they lose long-term proliferative capacity (Fig. 1a). To this end, we introduced the GFP coding region in the 3'untranslated region (UTR) of *DCN* (Extended Fig. 1b,c), a gene previously found to be lowly expressed in SSEA1$^+$ cells, which undergoes a 15-fold upregulation in SSEA1$^-$ tumor bulk cells characterized by limited proliferative capacity [16] (Extended Fig. 1d). Treatment of the engineered TDF cells with quisinostat, a histone deacetylase inhibitor that drives cancer cell differentiation[19], induced GFP expression, showing that the fluorescent reporter responds to differentiation cues (Extended Fig. 1e). We then analyzed TDF-induced tumors and observed that SSEA1$^-$ tumor bulk cells showed the expected increase in fluorescent GFP signal, indicating correct regulation of the knocked-in *DCN-GFP* allele *in vivo* (Extended Fig. 1f,g). Furthermore, transplantation assays for secondary tumor formation showed a reduced tumorigenic potential for GFP$^+$ tumor cells compared with SSEA1$^+$ cells, functionally validating the fluorescent reporter *in vivo* and indicating that the two phenotypic markers distinguish tumor cell subpopulations with distinct proliferative potential (Extended Fig. 1f,h). Although useful for enrichment of cells that have lost unlimited proliferative capacity, we reasoned that selection based purely on marker expression would yield many false positive hits since epigenetic modifiers targeted in the screen are involved in transcriptional regulation. We thus combined a functional readout based on clonogenic assays, a standard means to functionally assess cancer cell long-term proliferative capacity *in vitro* and predict tumorigenicity *in vivo* [20], and searched for sgRNAs that induced GFP expression and deprived cells of clonogenic ability (Fig. 1a).

As a positive control for the screen, knock-out of HRAS^{V12}, the oncogenic driver in TDF cells, resulted in a 4-fold increase in the expression of the *DCN*-GFP reporter, a 16-fold reduction in clonogenic ability and impaired tumorigenic potential (Extended Fig. 1i-k). These effects were not a consequence of reduced cell viability, confirming that the different screen readouts detect distinct consequences of gene knock-out on cell survival and proliferative potential (Extended Fig. 1l). To efficiently knock-out chromatin and DNA modifiers, we employed a previously characterized, focused sgRNA library[21] that targets 3,759 nuclear genes, including 346 epigenetic regulators, with up to 10 guides per gene (median number of sgRNAs per gene: 7) (Supplementary Table 1). The library also contained negative controls (39 non-targeting guides), and two distinct sets of positive controls (sgRNAs targeting 83 ribosomal genes, to assess screen saturation, and 3 *HRAS*-targeting sgRNAs) (Supplementary Table 1). To be able to temporally control KO induction, we introduced a doxycycline-inducible Cas9 in TDF cells.

We performed the screen in duplicate, favoring transduction of individual sgRNAs in each cell. Upon selection of transduced cells (T_0 population), gene knock-out was induced and cells grown for 14 days (T_{14} population). In parallel, an isogenic control cell line not expressing HRAS^{V12}, which is not clonogenic and does not induce tumor formation *in vivo* [16], was used for a viability counter-screen to exclude non-specific hits (Fig. 1a). At day 14,

half of the cell population was harvested for the viability arm, and the remaining cells were used for the unlimited proliferation arm following the strategy indicated above. sgRNAs were amplified from the cell population at various time points during the selection process and sequenced by next-generation sequencing (NGS) (Fig. 1a). Assessment of controls indicated highly consistent biological replicates at T_{14} (Extended Fig. 2a) ($r_s = 0.93$, p-value < 0.0001), efficient depletion of ribosomal genes (89% ± 1.26% depleted genes) and other essential genes such as *POLR2A* (Extended Fig. 2b,c and Fig. 1b), absence of non-targeting sgRNAs in the depleted set (Fig. 1b, Extended Fig. 2b) and depletion of the positive *HRAS* control in the unlimited proliferation arm (Fig. 1c). Overall, the viability arm identified 290 essential genes, 180 of which were also found in the counter-screen, leaving 110 hits specific to the tumor-maintaining cells (Fig. 1d, Extended Fig. 2d,e, Supplementary Table 2). An additional 181 genes were identified through the unlimited proliferation arm (Fig. 1c,d and Extended Fig. 2e, Supplementary Table 2). Among the 291 specific hits, 19 have been implicated in nuclear processes involving chromatin or DNA modifications (Fig. 1d, Supplementary Table 2).

Broad and cancer-specific dependency on the MSL complex

Inspection of the epigenetic-related hits revealed the presence of 3 out of 4 subunits of the acetyltransferase complex Male-Specific Lethal (MSL) [22] (p-value of the enrichment: 2.8×10^{-5}, hypergeometric test): *KAT8* from the viability arm, *MSL1* and *MSL2* from the unlimited proliferation arm (Fig. 1b,c). To validate the results of the screen, we performed tumor maintenance assays in which we injected TDF cells containing inducible Cas9 and sgRNAs targeting the three genes in NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice, induced gene KO *in vivo* in established tumors in half of the animals, and compared the growth rate of induced and uninduced tumors for each gene (Fig. 1e,f and Extended Fig. 3a). Knock-out of *KAT8*, *MSL1* and *MSL2* using sgRNAs distinct from those used in the screen significantly impaired tumor growth, while knock-out of the GFP reporter, as a control, had no effect, confirming reliable identification of genes important for tumor maintenance (Fig. 1f, Extended Fig. 3a).

The MSL complex plays a key role in dosage compensation in *Drosophila melanogaster*, where it mediates chromosome-wide gene upregulation on the X chromosome in male flies through deposition of histone H4K16ac, a major driver of chromatin decondensation[22, 23]. Although its molecular function and residue specificity are conserved in mammals, little is known about the physiological role of the complex in mammalian cells, where dosage compensation relies instead on gene silencing in females[24]. The mammalian MSL complex is ubiquitously expressed and acts as a homotetramer composed of four subunits with distinct molecular functions: the scaffold protein MSL1, the ubiquitin-ligase MSL2, the H4K20me-binding protein MSL3, and the acetyltransferase KAT8 (Fig. 1g). KAT8 is also part of the Non-Specific Lethal (NSL) complex that regulates expression of housekeeping genes, and interference with this subunit affects both complexes, preventing dissection of the specific role of MSL[25, 26]. In line with a role for MSL in sustaining cancer cell long-term proliferative potential, we were unable to derive and maintain knock-out monoclonal cell lines for any of the complex subunits. To overcome this issue, we introduced MSL-targeting sgRNAs in TDF cells expressing inducible Cas9 to be able to acutely knock-out the genes

upon doxycycline treatment. All subsequent experiments thus employed polyclonal populations in which the percentage of edited cells ranged from ~70% to ~90% (Extended Fig. 3b). As expected, knock-out of each subunit resulted in loss-of-function of the MSL complex (MSL^{LOF}) and widespread loss of H4K16ac within the population (Fig. 2a). Low H4K16ac levels were also observed upon generation of mutants of MSL subunits lacking specific domains, confirming that a fully functional MSL complex is required to maintain homeostatic levels of the histone marks [22] (Extended Fig. 3c,d, Supplementary Table 3). Despite efficient gene editing, a minority of cells escaping knock-out and maintaining high H4K16ac levels were present in all edited populations and their relative abundance increased over time (Extended Fig. 3e,f). While supporting a role of the complex in preserving cell long-term proliferative capacity, the negative selection of MSL^{LOF} cells creates an experimental challenge as it limits the time frame suitable for functional assays and leads to an underestimation of the full effect of MSL disruption. The optimal temporal window allowing MSL^{LOF} cells to manifest proliferation defects prior to being substantially depleted was between 14 and 21 days after knock-out.

In agreement with the screen results, *KAT8* knock-out led to cell death, due to the simultaneous disruption of both MSL and NSL complexes (Extended Fig. 4 a-c), whereas inactivation of MSL-specific subunits did not affect cell survival but impaired the proliferative capacity of cells, as assessed by limiting dilution clonogenic assays (Extended Fig. 4d,e). To further characterize how disruption of the MSL complex affects tumor cells *in vivo*, we repeated the tumor maintenance assays and harvested tumors 4 weeks after knock-out induction (Fig. 2b). Sequencing of genomic DNA revealed that 11 out of 24 tumors did not show substantial editing of the targeted MSL genes, indicating that they mainly contained cells that escaped gene knock-out (Extended Fig. 3g). We therefore selected edited tumors in which MSL loss-of-function (MSL^{LOF}) had been achieved and assessed their tumor organization, comparing them with unedited tumors used as controls. As expected, flow cytometry analysis showed that MSL disruption resulted in a significant decrease in the fraction of $SSEA1^{+}$ cells and a concomitant increase in the percentage of cells expressing the GFP reporter (Fig. 2c,d). In agreement, edited tumors contained 5-fold fewer cells endowed with unlimited proliferative potential, as assessed by limiting dilution transplantation assays for secondary tumor formation (Fig. 2e, Extended Fig. 4f). To validate these findings in clinically-relevant samples, we disrupted the MSL complex in patient-derived xenografts (PDX) from two different cancer types characterized by altered H4K16ac patterns: gastric and pancreatic cancer[27, 28], selecting models that shared the same genetic drivers as TDF-induced xenografts (activating *RAS* mutations and inactivating *TP53* mutations, see Supplementary Table 4). As observed in TDF-induced xenografts, disruption of the MSL complex inhibited tumor growth in both PDXs (Fig. 2f). Limiting dilution clonogenic assays showed impaired proliferative capacity of MSL^{LOF} cells from various other cancer types, including melanoma, breast cancer and osteosarcoma (Fig. 2g). HT-1080 fibrosarcoma cells, a p53-proficient, genetically stable line did not show a significant inhibition of clonogenic ability, suggesting that high levels of genetic instability may sensitize cells to MSL loss-of-function (Fig. 2g). As a more sensitive readout of proliferation defects, negative selection of $H4K16ac^{low}$ cells over time indicated that despite a reduced

sensitivity compared with genetically unstable lines, HT-1080 cells were also affected by MSL disruption (Extended Fig. 4g,h).

The observation that MSL components were not identified as hits in the counter-screen performed using a non-transformed isogenic line suggests that cancer cells may be specifically dependent on MSL to preserve long-term proliferative potential. To further examine this, we disrupted the complex in an independent non-cancerous cell line: telomerase-immortalized normal mammary epithelial cells (HME1). Consistent with the counter-screen results, normal HME1 cells did not show impaired proliferative potential upon MSL disruption and clones of MSL^{LOF} cells could be readily isolated and maintained (Fig. 2g,h). This observation is in line with the findings that *Msl1*-deficient ESCs have unaffected self-renewal ability[29] and that *Msl3*-null mice are viable and do not display overt defects[30], indicating that the MSL complex is largely dispensable for normal organismal function (Extended Fig. 4i, Supplementary Table 5). Altogether, these results indicate that a functional MSL complex is required to sustain tumor maintenance in various cancer types and that disruption of the complex impairs the ability of cells to proliferate indefinitely. Since many analyzed tumor models lack a functional p53, the effect of MSL disruption does not rely on p53-dependent tumor suppressive pathways [31]. The dependency on MSL is specific to cancer cells and normal cells preserve long-term proliferative capacity in the absence of a functional complex.

MSL loss induces genomic, not transcriptional, abnormalities

To understand how disruption of the MSL complex impairs cancer cell long-term proliferative potential, we performed transcriptional profiling of TDF cells in which the three MSL-specific subunits were knocked-out and compared them to two controls: wild-type TDF cells (WT) and cntr-KO cells, in which the GFP reporter was targeted, controlling for possible CRISPR-induced non-specific effects (Fig. 3a). We found 314 differentially expressed genes (DEGs) in MSL^{LOF} cells (False discovery rate (FDR) q-value ≤ 0.01, maximum transcripts per million (TPM) > 1), showing only moderate fold changes (FC) (median $\log_2$ FC: 0.31 and -0.42 for upregulated and downregulated genes, respectively) (Fig. 3a, Supplementary Table 6). Surprisingly, considering the genome-wide accumulation of H4K16ac at active promoters and enhancers[32], gene set enrichment analysis (GSEA)[33] did not show major alterations in specific biological processes or cancer-related pathways, suggesting that MSL disruption, and consequent loss of the mark did not trigger substantial changes in gene expression programs (Fig. 3b and Supplementary Table 7). In particular, we did not detect transcriptional changes consistent with cell differentiation, a biological process often associated with loss of unlimited proliferative potential within tumors [10] (Supplementary Table 7). However, GSEA revealed enrichment of 21 positional gene sets (FDR q-value < 0.01), indicating that neighboring genes on chromosomes showed concordant changes in mRNA levels upon MSL disruption (Fig. 3b,c and Supplementary Table 7). We thus examined whether the detected changes in mRNA levels were due to underlying differences in DNA content in MSL^{LOF} cells, and performed low-pass whole genome sequencing to identify possible copy number alterations. MSL^{LOF} cells were highly aneuploid (Extended Fig. 5a) and displayed several genomic alterations compared to both WT and cntr-KO cells (Fig. 3d and Supplementary Table 8). Importantly, DNA changes in

MSL^{LOF} cells correlated with the detected mRNA changes, with 13 out of 21 positional gene sets identified by GSEA showing consistent DNA copy number changes, and chromosomes displaying similar large-scale alterations in DNA and mRNA levels (Extended Fig. 5b and Fig. 3e). Thus, the apparent gene upregulation and downregulation detected by RNA-seq analysis in fact reflected increased aneuploidy in MSL-disrupted cells. Karyotype abnormalities of MSL^{LOF} cells were confirmed by analysis of chromosome spreads. As expected, due to p53 inactivation in TDF cells, WT and cntr-KO cells displayed a basal level of aneuploidy, with cells containing between 15 and 83 chromosomes (median: 64), but MSL^{LOF} cells showed additional numerical abnormalities and an overall increase in chromosome number (median: 69) (Fig. 3f).

Aneuploidy is the product of chromosomal instability (CIN)[34]. We therefore examined whether MSL-disrupted cells displayed an increased rate of mitotic defects and chromosome mis-segregation. Both in TDF cells and in cells derived from the pancreatic, gastric and melanoma PDXs, MSL disruption resulted in significantly increased frequency of cells containing micronuclei and mitotic divisions displaying lagging chromosomes or anaphase bridges, established hallmarks of CIN[34] (Fig. 3g). As expected, the p53-deficient PDX samples showed a basal level of CIN similarly to TDF cells. As a control, a low frequency of micronuclei-containing cells was observed in non-cancerous HME1 cells, which did not increase upon MSL disruption (Fig. 3g). To confirm these findings with a complimentary approach, we induced MSL loss-of-function in a CIN reporter cell line that allows quantification of CIN through detection of a GFP-encoding human artificial chromosome (HAC)[35] (Fig. 3h). MSL^{LOF} cells showed increased rate of HAC loss, regardless of which MSL-specific subunit was targeted, confirming enhanced CIN (Fig. 3i). We conclude that MSL disruption impairs the fidelity of chromosome segregation during cell division, and increases the rate of mitotic defects in cancer cells.

Accumulation of ssDNA promoting chromosome fragility

To dissect how disruption of the MSL complex leads to CIN we first analyzed the content of micronuclei (MN), assessing for centromere presence, indicative of whole chromosome mis-segregation, or γH2AX, indicative of damaged chromosomes. We found that 14% ± 4% of MN in MSL^{LOF} cells contained centromeric regions, indicating relatively infrequent presence of whole chromosomes. In contrast, 70% ± 5% were marked by γH2AX, suggesting compromised genome integrity as a major cause of MN formation (Fig. 4a,b, Extended Fig. 6a). In agreement, chromosome fragments and broken chromatids were enriched in metaphase spreads from MSL^{LOF} cells (Fig. 4c,d). γH2AX-marked MN did not contain active components of the double-strand break response pathway, such as 53BP1 and phosphorylated ATM[36] (Extended Fig. 6b,c), but were often marked by phosphorylated Replication Protein A (pRPA), a major single-stranded DNA (ssDNA) sensor (68% of γH2AX-marked MN, N: 213), and its downstream effector phosphorylated Checkpoint Kinase 1 (pCHK1)[36], suggesting the presence of ssDNA in MN (Fig. 4a, Extended Fig. 6d). In agreement, DAPI intensity was significantly lower in MN than in nuclei (Extended Fig. 6e). pRPA and pCHK1 foci were also found at high frequency within nuclei, a pattern consistent with widespread replication defects[36], with foci as big as one tenth of the nuclear area (Fig. 4e,f, Extended Fig. 6f,g). Increased frequency of pRPA-foci was also observed in

all other cancer models that had shown a dependency on MSL (Fig. 4f). To confirm that MSL disruption increases the endogenous level of replication stress we treated TDF cells with the DNA polymerase inhibitor aphidicolin, and indeed observed hypersensitivity of MSL^{LOF} cells to the drug (Fig. 4g). As a control, non-cancerous HME1 cells characterized by low replication stress were more resistant to aphidicolin treatment (Extended Fig. 6h). Importantly, while HME were insensitive to MSL disruption under normal growth conditions (Fig. 2g, Fig. 4h), induction of replication stress impaired the growth of MSL^{LOF} cells over five days, suggesting that basal levels of genomic stress critically sensitize cells to the loss of MSL function (Fig. 4h, Extended Fig. 6h). In line with the notion that under-replicated DNA can escape the G2/M checkpoint and interfere with chromosome segregation[37], γH2AX- and pRPA-marked regions were also detected in mitotic cells (Fig. 4i-l), with large DNA regions characterized by low DAPI intensity being strongly stained (Fig. 4i) and γH2AX foci marking anaphase bridges and chromatids with compromised integrity (Fig. 4j,l). Given the key role of H4K16ac in chromatin decondensation, a pre-requisite for efficient DNA replication[23, 38], these findings suggest that MSL loss-of-function and consequent loss of the histone mark promote CIN by leading to an accumulation of under-replicated DNA that cannot be properly segregated during mitosis.

Mis-segregation of whole chromosomes in MSL-disrupted cells

Although clastogenic events appear to be the main source of CIN in MSL^{LOF} cells, the numerical abnormalities detected in MSL^{LOF} cells suggests that cells may also suffer whole-chromosome mis-segregation. In agreement, MSL^{LOF} TDF cells displayed prolonged mitosis compared to both wild-type and cntr-KO cells, with cell divisions lasting up to 2.5 hours (Fig. 5a,b and Supplementary videos 1 and 2). In particular, complete alignment of chromosomes required up to 120 min and unattached chromosomes could often be detected (Fig. 5a, Supplementary videos 3 and 4). Similar defects were also observed in PDX-derived cells (Fig. 5b). This delay in mitotic progression was not due to alterations in the spindle assembly checkpoint (SAC), as shown by unaffected expression of the SAC components[39] and by the ability of cells to efficiently arrest in metaphase upon nocodazole treatment (Fig. 5c,d). Analysis of chromosome ultrastructure by scanning electron microscopy revealed that H4K16ac-depleted chromosomes, as expected, were more compacted than control ones, but these structural abnormalities did not result in altered centromeric chromatin (Fig. 5e,f). We then asked whether accumulation of under-replicated DNA may interfere with chromosome alignment. Indeed, mitotic MSL^{LOF} cells displayed unaligned chromosomes marked by strong γH2AX staining and containing a single centromere (Fig. 5g), and interphase cells showed an increased frequency of micronuclei positive for both centromere and γH2AX staining, likely containing under-replicated whole chromosomes (Fig. 5h,i). All together, these observations support the notion that replication stress and whole-chromosome mis-segregation may be mechanistically linked[37, 40, 41] and suggest that accumulation of ssDNA induced by MSL disruption promotes both structural and numerical CIN.

Enhanced CIN exhausts cells' proliferative capacity

The observations that MSL disruption increases the rate of chromosome mis-segregation suggests that a fitness cost associated with excessive CIN may underlie the loss of unlimited proliferative potential in MSL-deficient cells. Although CIN is overall beneficial for tumors,

as it generates cellular diversity that promotes cancer evolution, abnormal chromosome content is often deleterious at the individual-cell level[42–44]. Cancer cells, especially in the absence of functional p53, can tolerate aneuploidy, but high CIN levels generate broad cellular stress, including genotoxic, proteotoxic, metabolic and osmotic stress[42–44]. We therefore examined whether the proliferative potential of individual MSL-disrupted cells correlated with their CIN level. To do so, we grew clonal populations over 14 days and used DNA fluorescence *in situ* hybridization (FISH) to visualize five chromosomes (Chr.3, Chr. 7, Chr. 11, Chr. 12 and Chr. 17). For each clone, we quantified the number of cells and a CIN score based on the heterogeneity in chromosome content across cells (see Methods). As expected, MSL^{LOF} clones were smaller than WT and cntr-KO clones (p-value < 0.0001) (Fig. 6a, Extended Fig. 6i). Furthermore, confirming increased CIN, MSL^{LOF} clones were highly heterogeneous with respect to chromosome copy numbers, leading to significantly higher CIN scores (p-value < 0.0001) (Fig. 6b,c). Importantly, CIN score and clone size showed an inverse relationship, with clones characterized by high CIN containing significantly less cells than more stable clones, linking MSL-disruption, increased CIN and impaired long-term proliferative potential of cancer cells (Fig. 6d). To establish a causal link between enhanced CIN and impaired proliferative capacity, we treated TDF cells with reversine, an inhibitor of the mitotic kinase MSP1 that induces chromosome mis-segregation[45], and performed clonogenic assays. Reversine-treated cells grew over 12 days but progressively exited the cell cycle, as indicated by loss of Ki67 staining, and formed significantly smaller clones compared to DMSO-treated cells, phenocoping MSL genetic disruption and confirming that excessive CIN is overall detrimental to cancer cells (Fig. 6e-j). Reversine treatment also phenocopied MSL disruption *in vivo*, significantly delaying tumor development compared with control tumors (Fig. 6k, p = 0.009). To finally examine the relationship between enhanced CIN and cell differentiation state, we performed a similar analysis using HCC1569 breast cancer cells, whose differentiation state can be assessed by quantifying the phenotypic markers CD24 and CD44[46]. Although reversine impaired the cells' proliferative capacity, no change was observed in the markers, while the expected increase in $CD24^{+}$ cells was observed when cells were treated with quisinostat, which induces cell differentiation[19] (Fig. 6j,l). All together, these results show that excessive CIN induced by disruption of the MSL complex exhausts the proliferative capacity of cancer cells by compromising their fitness over multiple cell divisions, without affecting their differentiation state (Fig. 6m).

Discussion

We show here that disruption of the MSL complex and consequent loss of H4K16ac, a major driver of chromatin decondensation[23], impair the proliferative capacity of malignant cells from various cancer types. Surprisingly, this effect is not mediated by transcriptional changes and activation of a differentiation program, but rather by the accumulation of chromosomal abnormalities and aneuploidy that are progressively detrimental to cellular fitness. Increasing evidence from the analysis of both clinical samples and experimental systems suggests that the extent of CIN in cancer cells determines its tumor-promoting or tumor-suppressive functional output[44]. While moderate CIN levels fuel cancer evolution, high CIN levels are deleterious for tumor cells and correlate with good outcome in various

cancer types[44]. Based on these observations, strategies aimed at exacerbating CIN for therapeutic purposes have been explored[44]. Although inhibition of mitotic checkpoints effectively induces CIN[39, 47], the clinical usefulness of this approach is limited by the severe adverse effects on tissue homeostasis, as highly proliferative normal cells are also sensitive to the treatment[48–51]. We find that the sensitivity to MSL loss-of-function depends on pre-existing genetic instability and that non-cancerous cells preserve long-term proliferative potential upon disruption of the complex. In agreement with our findings, *Msl1*-deficient mouse embryonic stem cells display normal self-renewal ability[29] and *Msl3*-null mice are viable and do not display overt defects[30]. Owing to the selective effect of MSL disruption on malignant cells, inhibition of MSL function may be a well-tolerated means to induce extreme CIN in unstable cancer cells and deprive them of long-term proliferative capacity.

In agreement with observations made in Drosophila[22], we find that a fully functional MSL complex is required to maintain homeostatic levels of H4K16ac levels, offering multiple opportunities for pharmacological targeting. Extensive characterization of the complex at the biochemical and structural levels[22, 52, 53] aids the development of targeting strategies, which may include interference with protein-protein interactions among complex subunits, binding to chromatin and DNA mediated by MSL3, or the ubiquitin-ligase activity of MSL2. Targeting of KAT8's acetyltransferase activity is complicated by the simultaneous effect that such a strategy would have on the NSL complex, which contributes to the regulation of housekeeping genes[25, 26]. Nevertheless, encouraging results obtained recently with inhibitors targeting the related MYST acetyltransferases KAT6A/B and KAT7[54, 55] suggest that despite the general role of these proteins in sustaining cellular homeostasis, there may me a therapeutic window in cancer.

In addition to revealing MSL as a cancer-specific vulnerability relevant for various cancer types, our findings also suggest that accumulation of genomic abnormalities during cancer growth may be a differentiation-independent mechanism that generates non-self-renewing cells within tumors. We show that absence of long-term proliferative potential detected by standard self-renewal assays – clonogenic and limiting dilution transplantation assays - does not necessarily imply a differentiated phenotype and could reflect compromised cell fitness. These observations indicate that results from self-renewal assays may need to be cautiously interpreted in the absence of molecular evidence of differentiation, especially when analyzing solid tumors, which are often characterized by genetic instability.

Methods

Generation of the *DCN*-GFP reporter cell line

Transformed dermal fibroblasts (TDF) and the isogenic non-tumorigenic line lacking $HRAS^{V12}$ (DF)[16] were grown as indicated in Supplementary Table 9. To generate the fluorescent reporter cell line expressing *DCN*-GFP, 4 different plasmids were obtained from GenScript: pcDNA3.3-Cas9 for Cas9 expression, pCR Blunt II TOPO-sgRNA1 and 2 against *DCN*'s 3'UTR (Supplementary Table 10) and pUC57-*DCN*-GFP donor plasmid to be used for homologous recombination. The donor plasmid contained an IRES (sequence as in Addgene 64784), followed by the GFP coding sequence, and two 800 bp flanking regions homologous to the *DCN* 3'UTR (chr12:91539008-91539807 and

chr12:91539808-91540607). The donor plasmid harbored point-mutations in the sequence of the sgRNAs used for CRISPR-mediated editing to prevent re-editing of the integrated cassette (Supplementary Table 10). Six million non-tumorigenic DF cells, which express 15 times higher *DCN* levels compared to TDF cells, were electroporated with the four plasmids (5 μg each, with the donor plasmid linearized with NdeI [NEB R0111S] to increase editing efficiency). Six days after transfection, GFP^+ cells were sorted into individual wells of 96-well plates to isolate clonal populations. Individual clones were screened for correct editing by PCR and gel electrophoresis using primers specific to the knocked-in allele (Supplementary Table 11). The selected clone (Cl. C3) was then transduced with pBabe-$HRAS^{V12}$ as previously described[16] to generate TDF cells containing the knocked-in fluorescent reporter sequence. As expected, transformation induced downregulation of the reporter, leading to undetectable GFP levels in TDF cells. Dox-inducible 3xFLAG-Cas9 was finally introduced in the reporter TDF cell line by transducing cells with a modified pCW-HygroCas9 vector (Addgene 50661) in which the sequence encoding resistance to puromycin had been replaced with the sequence encoding resistance to hygromycin B (Supplementary Table 11). Individual clones were isolated and a clone with minimal leakiness and high expression levels upon induction with 1 μg/mL Doxycycline (Sigma D9891) was selected for use in the screen (Cl.14). In a similar way, inducible Cas9 was introduced in the DF reporter line and the resulting clone (Cl. L) was used for the counter-screen.

sgRNA library quality control and generation of positive and negative control sgRNAs

The sgRNA library (33,829 guides) targeting 3,759 nuclear genes was obtained from Addgene[21] and propagated by transforming highly-competent Stbl3 cells (Thermo Fisher C737303) ensuring at least 100 times coverage of the library complexity. Sequencing of the library, as described below, revealed the presence of other sgRNAs in addition to those belonging to the nuclear pool. These included sgRNAs targeting ribosomal genes and 5,692 sgRNAs from other pools of the genome-wide library[21] (Supplementary Table 1). Furthermore, non-targeting sgRNA controls (NTC) were underrepresented (only 9 out of 100 were detected). We therefore selected 30 NTC sequences from the non-targeting sgRNA subpool from the original published library, cloned them into the pLX-sgRNA backbone (Addgene 50662) as previously described[21] (https://media.addgene.org/data/08/61/acb3ad96-8db6-11e3-8f62-000c298a5150.pdf) and added them to the library obtained from Addgene at an equimolar ratio. Similarly, three positive control sgRNAs were designed using the CRISPR MIT Designer tool (crispr.mit.edu) against *HRAS*, cloned in the lentiviral vector and added to the library.

Virus generation and titer estimation for the screen

Virus production was performed by transfecting 80% confluent HEK-293T cells, grown as described in Supplementary Table 9, with pLenti- or pLX-sgRNAs[21, 57] (Supplementary Table 10), pCW-HygroCas9 (Addgene 50661) or H2B-mCherry (Addgene 20972) alongside packaging plasmids (psPax2 and pMD2.G, Addgene 12260 and 12259, respectively). Viral particles were used to infect cells with 5 μg/mL Polybrene (Santa Cruz sc-134220). Transduced cells were selected with the appropriate antibiotic for 5 d to eliminate non-infected cells (DF and TDF cells: 6 μg/mL Blasticidin S, Calbiochem 203350; 100 μg/mL

Hygromycin B, Thermo Fisher 10687010. HME1 and PDX-derived cells: 5 μg/mL Blasticidin S [GXA 3067 and MEXF 2090] or 7.5 μg/mL [PAXF 1997], Calbiochem 203350; 100 μg/mL Hygromycin B, Thermo Fisher 10687010). To ensure that cells were transduced with individual sgRNAs in the screens, the library viral titer was estimated by infection of 100,000 cells with decreasing viral concentrations and selection with Blasticidin S for 5 d. The percentage of surviving cells following selection relative to a non-selected control was used as a measure of infection efficiency. A viral dilution yielding an infection efficiency of 30% equivalent to a Multiplicity Of Infection (MOI) of 0.3 was used for the screens.

CRISPR-Cas9 screen, counter-screen and assessment of sgRNA abundance

The screen was performed in duplicate, favoring transduction of individual sgRNAs in each cell (multiplicity of infection: ~ 0.3) and ensuring that at least 1,000-fold coverage of the library complexity was maintained at each step. One-hundred and twenty-six million TDF cells (main screen) or DF cells (counter-screen) were transduced with the sgRNA library and selected with 6 μg/mL Blasticidin S (Calbiochem 203350) for 5 d. A sample of 50 million cells was isolated for use as T_0 population. Subsequently, Cas9 expression and gene KO was induced through continued administration of 1 μg/mL Dox for 14 d for TDF cells, or 24 d for the slower cycling DF cells to equalize population doublings, at which stage 50 million cells were isolated as end-point for the viability arm of the screen and counter-screen (T_{14} and T_{24}, respectively). A further 50 million TDF cells were stained with an eFluor 660-conjugated anti-SSEA1 antibody (eBioscience 50-8813-42) and analyzed by flow cytometry (see protein immunodetection section) to isolate $SSEA1^-/GFP^+$ cells by cell sorting in the unlimited proliferation arm. Sorted cells were grown for 48 h, after which half the population was harvested for genomic DNA isolation and the other half was plated in semi solid medium (Generon CBA-155, see clonogenic assay section). Cells were grown for 10 d, the resulting colonies collected following the manufacturer's instructions and their genomic DNA extracted. For all samples, genomic DNA (gDNA) extraction was performed using the DNeasy Blood & Tissue Kit (Qiagen 69506).

To prepare the sgRNA libraries for NGS, a 2-step nested PCR-based protocol originally described by Wang *et al.* was followed[21] using primers listed in Supplementary Table 11. For initial quality control of the sgRNA plasmid library, 5 ng of plasmid DNA were used as a template and amplified in 3 independent reactions that were pooled prior to sequencing. For the screen samples, either 85 μg (viability arm, corresponding to ~13 million cells) or all available gDNA (unlimited proliferation arm) was amplified. To ensure efficient amplification of the sgRNAs, multiple PCR reactions were run for each sample, using a maximum of 2 μg gDNA in 100 μl reactions with 20 cycles of amplification. Following the first PCR, all reactions were pooled and 5 μL were used as template for the second PCR, run in triplicates for 22-30 cycles. Final products were pooled, run on a 2% agarose gel, excised and purified using QIAquick gel extraction kit (Qiagen 28706) prior to sequencing. The libraries were analyzed on a DNA 1000 BioAnalyser 2,100 chip (Agilent) to ensure good quality, and sequenced on an Illumina HiSeq 4000 platform using custom primers (Supplementary Table 11) generating ~30-50 million reads per sample. PhiX DNA was added to the sequencing lanes at 35% to increase read base diversity and enable efficient

sequencing. Raw sequencing reads were trimmed to 20 bp using cutadapt with the "--cut -<trim_size>" parameter to identify the sgRNA sequences. These were then mapped to the sgRNA library using BWA (version 0.5.9-r16)[58] with the parameters "-l 20 -k 2 -n 2". sgRNA counts were obtained after de-multiplexing samples based on i7 indexes and excluding mapped reads with mismatches. To quantify the relative abundance of sgRNAs in each sample, raw reads for each sgRNA were normalized to the overall read counts.

Hit identification

Viability arm—sgRNAs depleted upon gene knock-out were identified by comparing the normalized sgRNA counts in the population 14 d after Cas9 induction (T_{14}) and in the initial population (T_0). Only sgRNAs with at least 1 raw count in both biological replicates at T_0 were used for analysis. Raw reads for each sgRNA were normalized to total read counts for each sample and the fold change (FC) between T_{14} and T_0 calculated for each replicate. Depleted sgRNAs showing a $\log_2FC \leq -1$ in both biological replicates at T_{14} compared to T_0 were selected. With a similar approach, depleted sgRNAs showing a $\log_2FC \leq -0.6$ between T_{24} and T_0 in the counter-screen were selected. The different threshold used in the counter-screen is to account for the overall lower sgRNA depletion observed when using slower-cycling DF cells. The $\log_2$ FC ≤ -0.6 threshold was chosen as it resulted in a fraction of depleted ribosomal genes comparable to that observed in the main screen. To robustly identify genes important for survival and cell cycle progression of tumor-maintaining cells, multiple filters were applied to exclude: (i) genes with less than 3 depleted sgRNAs; (ii) lowly expressed genes (TPM ≤ 5), as assessed by RNA-seq; (iii) genes with at least 3 depleted sgRNAs in the counter-screen after FC filtering. Positive and negative controls were used for quality control purposes and are excluded from hit count and related tables.

Unlimited proliferation arm—Depleted sgRNAs were identified by comparing the normalized sgRNA counts in cells retrieved from colonies grown in semi-solid medium and sorted $SSEA1^-/GFP^+$ cells. Only sgRNAs with at least 1 raw count in both biological replicates in the sorted population were used for analysis. Raw reads for each sgRNA were normalized to total read counts for each sample and the FC between colonies and sorted cells was calculated for each replicate. Depleted sgRNAs showing a $\log_2FC \leq -4$ in both biological replicates in the colonies compared with the sorted population were selected. To robustly identify genes important for sustaining unlimited proliferative capacity of tumor-maintaining cells, multiple filters were applied to exclude: (i) genes with less than 2 depleted sgRNAs; (ii) lowly expressed genes (TPM ≤ 5), as assessed by RNA-seq. Positive and negative controls were used for quality control purposes and are excluded from hit count and related tables.

Generation of Cas9-expressing cell lines

Cells were grown as indicated in Supplementary Table 9. For hTERT-immortalized human mammary epithelial cells (hTERT-HME1, ME16C) (ATCC CRL-4010), the CIN reporter HAC cells[35], and HCC-1569 breast cancer cells (ATCC CRL-2330, obtained from the Crick Institute common repository), doxycycline-inducible Cas9 was introduced as described for TDF cells. The CIN reporter line was transduced with a plasmid encoding Puromycin resistance and was selected with 1 μg/ml puromycin for 5 d. For U2OS osteosarcoma cells

(ATCC HTB-96, obtained from the Crick Institute common repository), and HT-1080 fibrosarcoma cells (ATCC CCL-121, obtained from the Crick Institute common repository) doxycycline-inducible Cas9 was knocked-in into the AAVS1 locus using the Genome-CRISPR Human AAVS1 Safe Harbor Gene Knock-in system (GeneCopoeia plasmids SH100 (AAVS1 CRISPR-Cas9 clone) and SH304 (AAVS1 Cas9 knock-in donor clone-TRE3G-Puro). TransIT-X2 (Mirus, MIR 6003) and Fugene HD (Promega E2311) transfection reagents were used to introduce plasmids into U2OS and HT1080 cells, respectively, and three days after transfection puromycin was added at concentration of 2 μg/ml for U2OS or 1 μg/ml for HT-1080 to select transfected cells. Individual clones were isolated and clones with minimal leakiness and high editing activity upon induction with 1 μg/ml doxycycline were selected.

Generation of knock-out and MSL-mutant cell lines

Knock-out lines were generated either by transducing Cas9-expressing cells with lentiviral constructs expressing sgRNA pools specific to the target genes[57] or by transfecting synthetic guide RNAs (Edit-R crRNA, Horizon) (Supplementary Table 10). pLenti-sgRNA constructs were generated as previously described[57] and transduction and selection performed as described above. To induce gene KO in the selected population, Cas9 expression was induced by treatment of cells with 1 μg/mL Doxycycline (Sigma D9891). For transfection of synthetic guide RNAs, each CRISPR RNA (crRNA) and the trans-activating CRISPR RNA (tracrRNA) were resuspended in 1x siRNA buffer (Horizon B-002000-UB-100) at a concentration of 20 μM. Each crRNA was then mixed with an equal amount of tracrRNA and the mix diluted 1:100 in Optimem. 10 μl of the cr/tracrRNA mix were added to a well of a 96-well plate (to achieve a final concentration of 20 nM in a 100 μl of final volume) and mixed with 10 μl of transfection reagent diluted in Optimem (Supplementary Table 9). After 15 min, 4,000 cells resuspended in 80 μl of complete medium containing 1 μg/ml doxycycline were added to each well. To maximize editing efficiency, cells were pre-treated with doxycycline for 24 h. The efficiency of knock-out in the populations ranged from 70% to 95% of cells, as assessed by sequencing of the targeted locus and/or quantification of H4K16ac by immunofluorescence microscopy. Because of progressive negative selection of MSL^{LOF} cells in the population, most experiments were performed between 7 and 21 days after knock-out induction, the optimal temporal window that allowed MSL^{LOF} cells to manifest proliferation defects prior to being substantially depleted. MSL-mutant lines were generated using the synthetic guide RNAs indicated in Supplementary Table 10.

Patient-derived xenograft (PDX) models

Information about the GXA 3067, PAXF 1997 and MEXF 2090 PDX models are listed in Supplementary Table 4. Models were obtained from the Charles Rivers tumor model compendium https://compendium.criver.com/compendium2/cancertype?species.name=Human and propagated in NSG mice. Cells derived from each PDX were grown as indicated in Supplementary Table 9. Doxycycline-inducible Cas9 was introduced in PDX-derived cells as described above for TDF cells.

Animal studies

Animal studies were subject to ethical review by the Francis Crick Animal Welfare and Ethical Review Body and regulation by the UK Home Office project license PPL 70/8167 and PC2165EA4. NSG mice were maintained under pathogen-free conditions, and food and water were provided *ad libitum*. For generation of primary tumors, 5×10^5 TDF or PDX-derived cells and 1×10^5 carrier hTERT-immortalized fibroblasts were injected intradermally in both flanks of 8-10 week-old male NSG mice in 50 μL of PBS. For validation of the reporter cell line, 1,000 sorted cells and 1×10^5 carrier hTERT-immortalized fibroblasts were injected. For tumor maintenance experiments, upon appearance of ~2 × 2 mm tumors, mice were randomly segregated into a +/- Dox treatment (2 mg/mL Doxycycline in drinking water supplemented with 1% sucrose, or 1% sucrose alone, changed every 2-3 days) and tumor volume was measured biweekly using electronic calipers until animals were humanely sacrificed. For experiments directly testing the effect of CIN on tumor development, TDF cells were treated with DMSO (control) or 250 nM reversine (Generon, HY-14711-1mL) for 4 days and then injected intradermally into both flanks of 8-10 week-old female NSG mice (2.5×10^6 cells in 50 μl of sterile PBS). For generation of secondary tumors in limiting dilution transplantation experiments, 10, 100 or 1,000 cells from dissociated TDF-induced primary tumors were re-injected into the flanks of new recipient NSG mice together with 1×10^5 non-tumorigenic carrier hTERT-immortalized fibroblasts in 50 μL of PBS. In experiments comparing unedited and edited tumors induced by TDF cells containing MSL subunit-targeting sgRNAs, cells from all unedited (< 50% edited sequence) or edited (> 50% edited sequence) *MSL3*-KO dissociated primary tumors were pooled and injected into mice. Tumor appearance was scored over 8 weeks, when mice were humanely sacrificed. Frequency of tumor-propagating cells in primary tumors was estimated by limiting dilution analysis using ELDA software (http://bioinf.wehi.edu.au/software/elda/).

Tumor characterization

For tumor dissociation, subcutaneous tumors were collected, cut into small pieces of ~2-3 mm in diameter with a scalpel, transferred into GentleMacs C tubes (Miltenyi Biotec 130-096-334) containing 3 volumes of RPMI (Gibco 21875034) supplemented with 0.5x Liberase (Sigma 5401020001), dissociated with the GentleMacs dissociator (Human tumor 1.1 program) and incubated for 30 min in fast agitation at 37 °C. Cells were further dissociated with the GentleMacs dissociator (Human tumor 2.1 program) and incubated for 30 min at 37 °C with 100 U/mL of DNase I (NEB M0303). Following one last dissociation step (Human tumor 3.1 program), cells were filtered through a 70 μm cell strainer (Fisher 10788201), washed 3 times with RPMI and viable cells counted. Cells were then used for flow cytometry analysis or limiting dilution transplantation assays.

For validation of the *DCN*-GFP reporter cell line, cells from dissociated primary tumors were stained with an eFluor-660-conjugated anti-SSEA1 antibody (eBioscience 50-8813-42) and $SSEA1^+$ and GFP^+ cells sorted by flow cytometry. One thousand $SSEA1^+$ or GFP^+ sorted cells were then injected into both flanks of new recipient NSG mice together with 1×10^5 non-tumorigenic carrier hTERT-immortalized cells in 50 μL of PBS, and tumor growth was scored over 4 weeks, when mice were humanely sacrificed.

For analysis of genome editing in tumor samples, tumors were cut into small pieces of ~2-3 mm in diameter, transferred into gentleMACS M tubes (GentleMACS 130-096-335) containing 2 volumes of ATL buffer supplemented with 1:20 proteinase K (Qiagen 69506) and blended at high speed with a gentleMACS dissociator (RNA_01.01 Program). Genomic DNA was extracted using DNeasy Blood & Tissue Kit as per manufacturer's protocol (Qiagen 69506) and used for assessment of genome editing by TIDE analysis[59] https://tide.deskgen.com/).

Clonogenic assays

Colony formation assays in semi-solid medium were performed using the CytoSelect Clonogenic Tumor Cell Isolation kit (Generon CBA-155) in MEM following the manufacturer's instructions. To generate the bottom layer, either 16 mL or 1.5 mL of semi solid medium were added to 15 cm plates (unlimited proliferation screen arm) or 6-well plates (*HRAS*-KO), respectively, and let solidify at 4 °C for 30 min. The top layer containing 80,000 (unlimited proliferation screen arm) or 5,000 (*HRAS*-KO) cells was then poured, let solidify at 4 °C for 5 min and plates were then placed in an incubator at 37°C. Cells were grown for 10 d, after which colonies were either recovered from the semi solid medium following the manufacturer's instructions (Generon CBA-155, unlimited proliferation screen arm), or stained with 100 μg/mL 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Sigma M2128) overnight at 37 °C (*HRAS*-KO). To quantify the number of colonies, stained plates were scanned with an Epson Perfection V700 scanner (Epson) and images processed with ImageJ.

For limiting dilution two-dimensional clonogenic assays, suitable also for non-transformed cells, a matrix of decreasing numbers of cells was seeded in 96-well plates. Starting from 4000 cells transferred in the A1 well in 200 μl and through 2-fold serial dilutions of cells from top to bottom rows and from left to right columns, the final estimated number of cells seeded in 100 μl of medium in each well ranged from 2000 (A1) to 0.015 (J12), with diagonal wells containing identical estimated numbers of cells. Cells were plated in three 96-well plates per condition. 11-16 days after plating, depending on the cell line, whole wells were imaged using an IncuCyte Zoom live cell imager (Essen bioscience) with a 4x objective and populated wells, containing at least one clone of more than 20 cells, were scored manually. Due to edge effects affecting cell growth, the 36 external wells were excluded from the analysis and they are not shown in Extended Fig. 4d. Cell clonogenic potential was estimated by the percentage of populated internal wells. To quantify clone size, the number of cells per colony after 10 d of growth was manually quantified in a double-blind manner. For visualization purposes, clones were stained with Crystal Violet solution (0.5% crystal violet powder [Sigma C0775] and 20% methanol in ddH2O).

In vitro proliferation assays

Cell growth kinetics were assessed using the MTS-based CellTiter 96 Aqueous One Solution (Promega G3582) following the manufacturer's protocol. Five thousand cells per well, or 20,000 for day 0 used for normalization, were plated in 96-well plates in triplicates and grown for 3 d. Daily, 20 μL of the CellTiter solution were added per well and incubated for 2.5 h at 37 °C. The optical density of the wells at 490 nm was then measured with an iMark

Microplate Absorbance Reader (Biorad 168-1130) and used as a measure of the number of viable cells in each well.

HAC loss assays

A monoclonal CIN reporter line kindly shared by V. Larionov[35] and modified to express inducible Cas9 (see above) was transfected with cr/tracrRNAs targeting MSL-subunits or a control cr/tracrRNAs targeting the *LMNA* gene as indicated above. Three days after transfection cells were plated in three 24-well plates and cells grown for an additional 13 days in the absence of Blasticidin. While growing, cells were split at a 1:3 ratio when they approached confluence. After 16 d, cells were trypsinized, resuspended in FACS buffer (see below) and analyzed using an LSR Fortessa analyzer (BD Biosciences) to quantify the fraction of GFP-negative cells. FACSDiva and FlowJo software were used to acquire and analyze data, respectively.

Immunofluorescence microscopy

Immunostaining of cultured cells grown on coverslips and fixed in 4% paraformaldehyde (PFA, Alfa Aesar 43368.9) in PBS was performed following standard protocols using the following primary antibodies: anti-H4K16ac (Cell Signaling Technology 13534, 1:500), anti-γH2A.X (Millipore, 05636, 1:1000), human anti-crest/centromere (Immunovision, HCT-0100, 1:500), anti-pATM S1981 (Abcam ab81292, 1:200), anti-53BP1 (Novus Biologicals 100-304, 1:200), anti-pRPA S33 (Bethyl Laboratories A300-246A, 1:200), anti-pCHK1 S345 (Cell Signaling technology 2348 1:50) and anti-Ki67 (BD Bioscience 610968, 1:500). Secondary antibodies, used at a 1:400 dilution, were: Alexa fluor 568 donkey anti rabbit (Thermo Fisher Scientific, A10042); Alexa fluor 488 donkey anti mouse (Thermo Fisher Scientific, A21202), Dylight 650 goat anti-human (Thermo Fisher Scientific, SA5-10137). After staining, coverslips were mounted using vectashield mounting medium containing DAPI (vector laboratories, H-1200-10). To assess H4K16ac levels following *KAT8* KO, 4 d after Cas9 induction, detached cells collected from the medium or trypsinized cells from the adherent population were plated onto poly-L-lysine-coated coverslips (BD Biocoat 354085) for 30 min prior to fixation with 4% PFA and staining as previously described[17]. Cells were imaged using an Axio Observer microscope (Zeiss) using 10X, 40X or 63X objectives or a Zeiss LSM 710 confocal microscope using a 63X objective. Quantification of H4K16ac signal was performed using MetaMorph v7.7.7.0 software.

For high-throughput imaging and quantification of pRPA foci, cells grown and stained in 96-well plates (Miltenyi Biotec 130-098-264) were imaged using an Opera Phenix High Content Screening System (PerkinElmer) with 63x/NA 1.15 water-immersion lens. Z-stacks from 2 to 4.1μm with a step size of 0.3μm were acquired using excitation lasers at 405, 488, 561, and 640nm, and emission filters at 450, 540, 600, and 690nm, respectively. Cell segmentation and quantification analysis were performed using Harmony software 4.9 detecting the parameters indicated in Supplementary Table 12. The presence of γH2A.X, pRPA and centromeres in micronuclei was scored manually analyzing images acquired through the Phenix system.

Flow cytometry

Flow cytometry analysis was performed using anti-SSEA1-eFluor 660 (eBioscience 50-8813-42, 1:20), anti-CD44-FITC (IM7) (eBioscience, 11-0441-81, 0.5 mg per test) and anti-CD24-VioBlue (clone: 32D12) (Miltenyi biotech, 130-099-150, 1:11). For tumor-derived samples, anti-mouse MHC Class I H2-K^d (eBioscience 17-5957-82, 1:200) followed by PerCP-Cy5.5-conjugated streptavidin (eBioscience 45-4317-80, 1:200) was used to exclude host cells. Cells were incubated with antibodies at a concentration of 1×10^6 cells/100 μL in sort buffer (15 mM HEPES buffer [Sigma H3375], 1% Bovine Serum Albumin [Sigma A9647], 2 mM EDTA [VWR 20302.260], 100 U/mL DNase I (NEB M0303), 100 U/mL penicillin and 100 μg/mL streptomycin [Thermo Fisher 15140122] in PBS) for 45 min on ice, washed 3 times with sort buffer and analyzed using LSR Fortessa (BD Biosciences) or sorted using FACSAria II (BD Biosciences). FACSDiva and FlowJo software were used to acquire and analyzed data, respectively. Gating strategy used for tumor analysis is shown in Supplementary Fig. 1.

RNA sequencing

RNA extraction was performed using RNeasy Plus Mini Kit (Qiagen 74136) as per manufacturer's instructions. RNA-seq analysis was performed on TDF cells transduced with sgRNAs targeting MSL subunits or GFP 14 d after Cas9 induction by doxycycline treatment or non-transduced wild-type cells. At that stage, ~ 70% of cells were knocked-out, as assessed by H4K16ac immunostaining. Libraries were prepared using KAPA mRNA Hyper Prep Kit (Roche KK8581), assessed on a DNA 1000 BioAnalyser 2100 chip (Agilent) to ensure good quality, and sequenced on an Illumina HiSeq 4000 platform generating ~37 million 76 bp strand-specific single-end reads per sample. Adapter trimming was performed with cutadapt (version 1.9.1)[60] with parameters "--minimum-length = 25 --quality-cutoff = 20 -a AGATCGGAAGAGC". The RSEM package (version 1.3.0)[61] in conjunction with the STAR alignment algorithm (version 2.5.2a)[62] was used for the mapping and subsequent gene-level counting of the sequenced reads with respect to hg19 RefSeq genes downloaded from the UCSC Table Browser[63] on 7th June 2017. The parameters used were "--star-output-genome-bam --forward-prob 0". Differential expression analysis was performed with the DESeq2 package (version 1.12.3)[64] within the R programming environment (version 3.3.1). An adjusted FDR p-value of ≤ 0.01 was used as the significance threshold for the identification of differentially expressed genes. Normalized transcripts per million (TPM) values were generated for all replicates, which were averaged for the same gene across the three biological replicates for each corresponding condition. Genes with a maximum TPM value lower than 1 across all the samples were excluded from the analysis. To identify differentially expressed genes (DEGs) in MSL-disrupted cells, pairwise comparisons between each MSL-KO sample (*MSL1*-KO, *MSL2*-KO or *MSL3*-KO cells) and each control sample (*GFP*-KO or WT cells) were first performed to identify genes upregulated or downregulated in each MSL-KO sample with an FDR ≤ 0.01 relative to each control. DEGs were defined as genes showing differential expression in all MSL-KO samples compared to both controls (overlap of all individual gene lists, analyzing upregulated and downregulated separately). Gene set enrichment analysis was performed using the GSEA software (version 3.0, Broad Institute) using default settings. Fragments Per Kilobase of transcript per Million (FPKM) values for MSL-KO and control samples were provided to the algorithm and tested

for enrichment of Gene Ontology (GO), hallmarks, and oncogenic signatures from the MSigDB database or custom-made positional gene sets generated based on hg19 RefSeq genes and cytogenetic sub-band coordinates downloaded from the UCSC genome browser. Gene signatures were considered differentially expressed when a FDR q-value ≤ 0.01 was obtained after 1,000 permutations.

Karyotype analysis

To perform karyotypic analysis, genomic DNA from 10 biological replicates containing 100,000 cells each was extracted 14 d after Cas9 induction using the DNeasy Blood & Tissue Kit (Qiagen 69506). Whole-genome sequencing libraries were generated using the Nextera DNA Flex Library Prep Kit (Illumina 20018705) and were confirmed to be of good quality using a TapeStation 4200 with D1000 ScreenTape (Agilent G2991AA). Libraries were equalized in DNA quantity, pooled and sequenced on an Illumina HiSeq 2500 platform generating ~ 250 million paired-end 100 bp reads per sample. Raw sequenced reads were adapter/quality trimmed using TrimGalore 0.6.0 and then aligned to the GRCh37/hg19 genome using bwa mem (bwa 0.7.15)[58] with subsequent sorting and merging of replicates performed using samtools 0.1.19[65] (total coverage per sample: ~8x). Single sample copy number variation estimates were performed using the QDNASeq R package (QDNASeq-Bioconductor 3.8 for R 3.5)[66] with a bin size of 1,000 kb. This analysis takes the approach that in the absence of copy number variation, genomic loci of similar sequence content should have roughly equal read counts. The method applies a correction for sequence mappability and GC content. The mappability bed file used to construct the QDNASeq bins file for Hg19 was obtained from the encode project: https://www.encodeproject.org/files/ENCFF011ZFS/@@download/ENCFF011ZFS.bed.gz). Since the X and Y chromosomes affect the performance of QDNASeq, they were removed prior to the normalization/smoothing stage. QDNASeq operates on a per-sample basis and performs a 2-dimensional smoothing by regressing observed values for each GC-content/Mappability bin onto all comparable bins within that sample. Paired (relative) copy number analysis directly comparing wild-type or cntr-KO cells and MSL-disrupted cells was performed using CNVKit (0.9.5)[56] with –target-avg-size set to 3,000 nucleotides. For each chromosome, adjacent segments with common estimated absolute copy number were merged if their boundaries were no more than 5 Mb apart. Merged segments of total length < 2Mb were discarded. Each segment logR is reported as the mean logR of the intervals it contains.

Aphidicolin sensitivity assays

Three thousand cells were plated in triplicates in 96-well plates, excluding external wells, or 24-well plates, treated with increasing concentrations of Aphidicolin (Sigma A4487-1ML) for 16 hours, and then cultured for 5 days. When using 96 well plates, cells were split 1:3 after 48 h. Five days after plating, cells were fixed with 4% PFA in PBS, permeabilized with 0.5% Triton in PBS and stained with SYTOX Green Nucleic Acid Stain (Thermo Fisher S7020). After a wash in PBS, whole wells were imaged using an IncuCyte Zoom live cell imager (Essen bioscience) with a 4x objective. Nuclei were quantified using IncuCyte software (version 2020B) and EC50 estimated using the built-in concentration response function.

Fluorescence *in-situ* hybridization (FISH)

Cells were plated sparsely on glass slides and grown for 14 d to generate clonal populations. Slides were treated with a hypotonic solution of 0.075 M KCl for 7 min at 37 °C, fixed immediately in ice cold 3:1 methanol:glacial acetic acid for 20 min at RT and dried for 1 week at RT. Samples were then treated with SPoT-Light tissue pre-treatment kit (Thermo Fisher 8401) at 98 °C for 10 min, washed twice in PBS, dehydrated in a graded ethanol series of 70%, 85%, 100% EtOH for 5 min each and air dried for 10 min at RT. Genomic DNA was co-denatured with centromeric probes for chromosomes 3, 7 and 12 (Cytotest CT-CCP007-10-O, CT-CCP012-10-G, CT-CCP03-10-R) or chromosomes 11 and 17 (CT-CCP011-10-O, CT-CCP017-10-R) at 72 °C for 2 min and hybridized overnight at 37 °C in a humidified chamber. Slides were then washed 3 times in 0.5x Saline-Sodium Citrate (SCC) firstly at RT, then at 75 °C and finally at RT again for 5 min each. After a final wash in nuclease-free dH_2O for 5 min at RT, slides were dehydrated in a graded ethanol series of 70%, 85% and 100% EtOH for 5 min each and air dried for 10 min at RT. Finally, slides were mounted with Vectashield containing DAPI (Vector H-1200). Images were acquired with an Olympus VS120 Slide Scanner (Olympus) and the number of foci per cell detected in each channel manually counted in a double-blind manner using the QuPath v0.2.2 software. Individual images were manually contrasted to remove background fluorescence and identify *bona fide* peri-centromeric regions for quantification. For each clone, a chromosomal instability (CIN) score was calculated (variance in the number of foci per cell across the population).

Analysis of chromosome segregation defects and SAC activity

To quantify chromosome segregation defects, cells were seeded at a 60% confluence onto glass coverslips (MarineField Superior 0111520), incubated overnight, fixed with fresh 4% PFA (Alfa Aesar 43368.9) for 15 min, washed with PBS and permeabilized with 0.5% Triton-X100 (Sigma T8787) in PBS for 5 min. Coverslips were then mounted onto glass slides with Vectashield containing DAPI (Vector H-1200). Cells were imaged with a Axio Observer microscope (Zeiss) using a 40x objective. Mitotic cells with lagging chromosomes and interphase cells with micronuclei were manually scored in a double-blind manner. In experiments involving only TDF cells, editing of the integrated GFP reporter was used as a control. In experiments involving also PDX-derived lines, the *LMNA* gene was chosen as an endogenous control for gene editing given its ubiquitous expression across tissues and the harmless effects of its loss on cells. To probe SAC activity, cells were treated with 100nM nocodazole for 24h and the percentage of mitotic cells scored upon imaging with an IncuCyte Zoom live cell imager (Essen bioscience).

Metaphase spreads analysis

To quantify the number of chromosomes per cell, cells growing in T-25 flasks were treated with 0.5 μg/mL KaryoMAX Colcemid (Thermo Fisher 15212012) for 30 min and plates tapped to detach mitotic cells. Cells recovered from the medium were then treated with a hypotonic solution of 0.075 M KCl for 20 min at 37 °C and immediately fixed and washed 5 times with 3:1 methanol:glacial acetic acid. A drop of resuspended cells was seeded on a humidified glass slide at a 35° angle, dried at RT for 30 min and mounted with Vectashield

with DAPI (Vector H-1200). Images of mitotic chromosomes were acquired with an Axio Observer microscope (Zeiss) or with a LSM 710 confocal microscope (Zeiss) using a 63x objective. The number of chromosomes per cell was manually quantified in a double-blind manner. For experiments involving immunostaining of chromosome spreads, cells were not treated with colcemid to avoid changes in chromosome structure or mitotic signaling. 2 ml of cold 0.075 M KCl were added to cells growing in T-25 flasks, and flasks tapped to detach mitotic cells. Detached cells were transferred into a 15 ml tube and incubated on ice for 30 minutes. $1x10^5$ cells were then resuspended in 1ml of 0.075 M KCl + 0.1% tween-20 and spun onto SuperFrost Plus glass slides (Menzel-Glaser) at 800 rpm for 7 min in a Cytospin 3 (Thermo Shandon). Slides were immediately transferred into a coplin jar containing potassium chromosome medium (KMC: 120 mM KCl, 20 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 0.1% (v/v) Triton X-100), incubated for 15 minutes and then blocked for 1h at room temperature with 10% BSA in KCM medium. Slides were stained with the following primary antibodies diluted in 10% BSA-KCM for 1h at room temperature: rabbit anti-H4K16Ac (cell signaling, 13534S) 1:500, mouse anti-γH2A.X (Millipore, 05636) 1:1000, human anti-crest/centromere (Immunovision, HCT-0100) 1:500. Slides were then washed twice for 5 minutes with KCM, and incubated for 30 min at room temperature with secondary antibodies diluted 1:500 in KCM medium: Alexa fluor 568 donkey anti rabbit (Thermo Fisher Scientific, A10042); Alexa fluor 488 donkey anti mouse (Thermo Fisher Scientific, A21202), Dylight 650 goat anti-human (Thermo Fisher Scientific, SA5-10137). After two washes with KCM, slides were fixed for 15 minutes with KCM containing 4% (v/v) paraformaldehyde at room temperature, washed 5 minutes with distilled water and finally mounted with vectashield mounting medium containing DAPI (vector laboratories, H-1200-10). Slides were images on an Axio Observer microscope (Zeiss) using a 63x objective.

Live-cell imaging

For fluorescent imaging of mitotic divisions, 1×10^4 TDF cells transduced with H2B-mCherry (Addgene 20972), were seeded in 24-well glass bottom plates (Cellvis P24-1.5H-N) in phenol red-free media (Thermo Fisher 51200038) and grown overnight. Prior to imaging, cells were treated with 50 nM SiR-Tubulin-647 (Cytoskeleton, CY-SC002) to visualize microtubules. Cells were maintained in an Okolab environment chamber at 37 °C and 5% CO_2 and images acquired every 5 min for 24 h using a Nikon Ti2 inverted microscope and a 40x objective with Perfect Focus System and ASI motorized XY stage with Piezo Z, capturing Z-stacks (11 × 1 μm steps). Images were processed using Micromanager 2.0 and videos generated with ImageJ. To quantify the length of mitosis, cells were seeded in duplicates in 6-well plates at a ~30% confluence and incubated overnight. Cells were then transferred into an IncuCyte Zoom live cell imager (Essen bioscience) and imaged every 5 min for 4 h. The length of mitosis was manually quantified in a double-blind manner from the time a cell rounded up to enter mitosis until two distinct daughter cells formed.

Scanning electron microscopy

Chromosome spreads generated as described above (methanol:glacial acetic acid fixation), but not air dried, were incubated in SPoT-Light tissue pre-treatment kit (Thermo Fisher

8401) for 10 min at 50 °C to remove cell debris. Samples were then fixed in 4% formaldehyde/2.5% glutaraldehyde in 0.1 M phosphate buffer (PB) pH 7.4 for 30 min at RT, washed 3 times in PB for 5 min, stained in 1% OsO4/1.5% potassium ferricyanide for 1 h at RT, washed in water, dehydrated in a graded ethanol series of 70%, 90% and 100% EtOH twice at each step for 5 min, critical point dried with CO_2 using a Leica EM CPD300 and mounted on aluminum stubs using adhesive carbon tabs. Samples were coated with 2 nm of platinum using a Quorum Q150 R S sputter coater and the quality of the preparation assessed by imaging on a Leica M205C stereomicroscope. Samples from *MSL3*-KO and *GFP*-KO cells were processed side by side to minimize inter-preparation variability, and two independent preparations for each genotype were analyzed. Scanning electron microscopy images were recorded at 5 kV and a working distance of approximately 7 mm with a FEI Quanta 250 FEG scanning electron microscope and ETD detector.

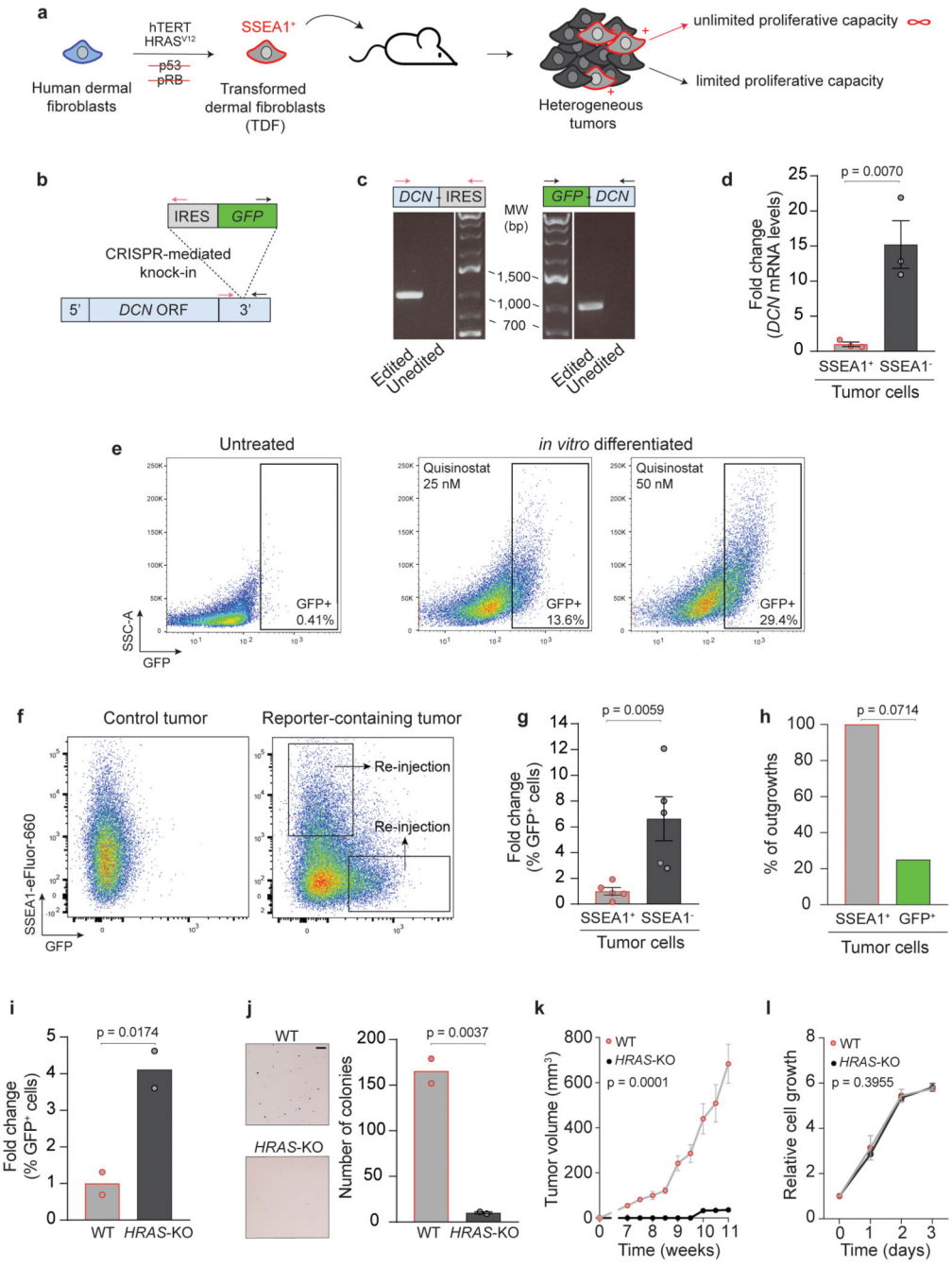
Analysis of chromatin fibers

Extended fiber with double lysis immuno-fluorescence microscopy were performed previously described[67]. Briefly, 50,000 cells were resuspended in 500 µL hypotonic buffer (0.075 M KCl) and incubated for 10 min at room temperature. Cells were then cytospun for 4 min at 800 rpm onto a poly-lysine coated glass slide. Slides were transferred to a Coplin jar with salt-detergent lysis buffer (25 mM Tris, pH: 9.5; 500 mM NaCl; 500 mM Urea; 1% Triton-X; 1 mM PMSF) for 20 min, transferred to a second jar with PBS-T for 20 min, and repeated again with salt-detergent lysis buffer. In separate jars, fibers were fixed with 3.7% paraformaldehyde for 10 min at room temperature, washed twice and then stored in PBS until immunostaining was performed using an anti-CENP-A antibody (Abcam ab13939 1:200) and an anti-H4K16Ac (cell signaling, 13534S 1:500) as previously described[68].

Statistics and reproducibility

The type of statistical tests performed in this study, whether they were one- or two-tailed, the exact value of N, and what N represents are indicated in the main text, in figure legends, or in the relevant methods section. Unless otherwise stated, all values are the average of individual values ± standard error of the mean (SEM). Statistical analysis was performed using the indicated NGS-related packages or GraphPad software. All experiments were independently repeated at least two times with similar results obtained.

Extended Data



Extended Data Fig. 1. Generation of a reporter cell line to assess loss of unlimited proliferative capacity

a. Schematics of the experimental system used for the screen.

b. Diagram illustrating the design of the *DCN*-driven fluorescent reporter. Arrows: approximate location of primers used in **c.** ORF: open reading frame.

c. Amplicons detecting the correctly integrated IRES-GFP cassette within the *DCN* 3'UTR. Edited: TDF reporter line; unedited: unedited parental cells. Schematics at the top show the sequence junctions probed. MW: molecular weight.

d. Quantification of *DCN* upregulation in SSEA1$^-$ non-tumorigenic tumor cells as assessed by microarray analysis [16]. Values are average ± SEM from n = 3 biologically independent tumors. P-value from one-tailed Student's t-test.

e. Flow cytometry analysis of untreated or quisinostat-treated, differentiated TDF reporter cells.

f. Flow cytometry analysis of a tumor containing the *DCN*-GFP reporter and of a control tumor induced by unedited TDF cells. Gates used for sorting cells used in **h**.

g. Quantification of *DCN* upregulation in SSEA1$^-$ non-tumorigenic tumor cells, as assessed by flow cytometry (FACS) detecting expression of the *DCN*-GFP reporter. Values are average ± SEM from n = 5 biologically independent tumors. P-value from one-tailed Student's t-test.

h. Transplantation assays for secondary tumor formation. Quantification of tumors induced by injection of 1,000 SSEA1$^+$ or GFP$^+$ sorted cells. N = 4 independent injections per condition. P-value from one-sided Fisher's exact test.

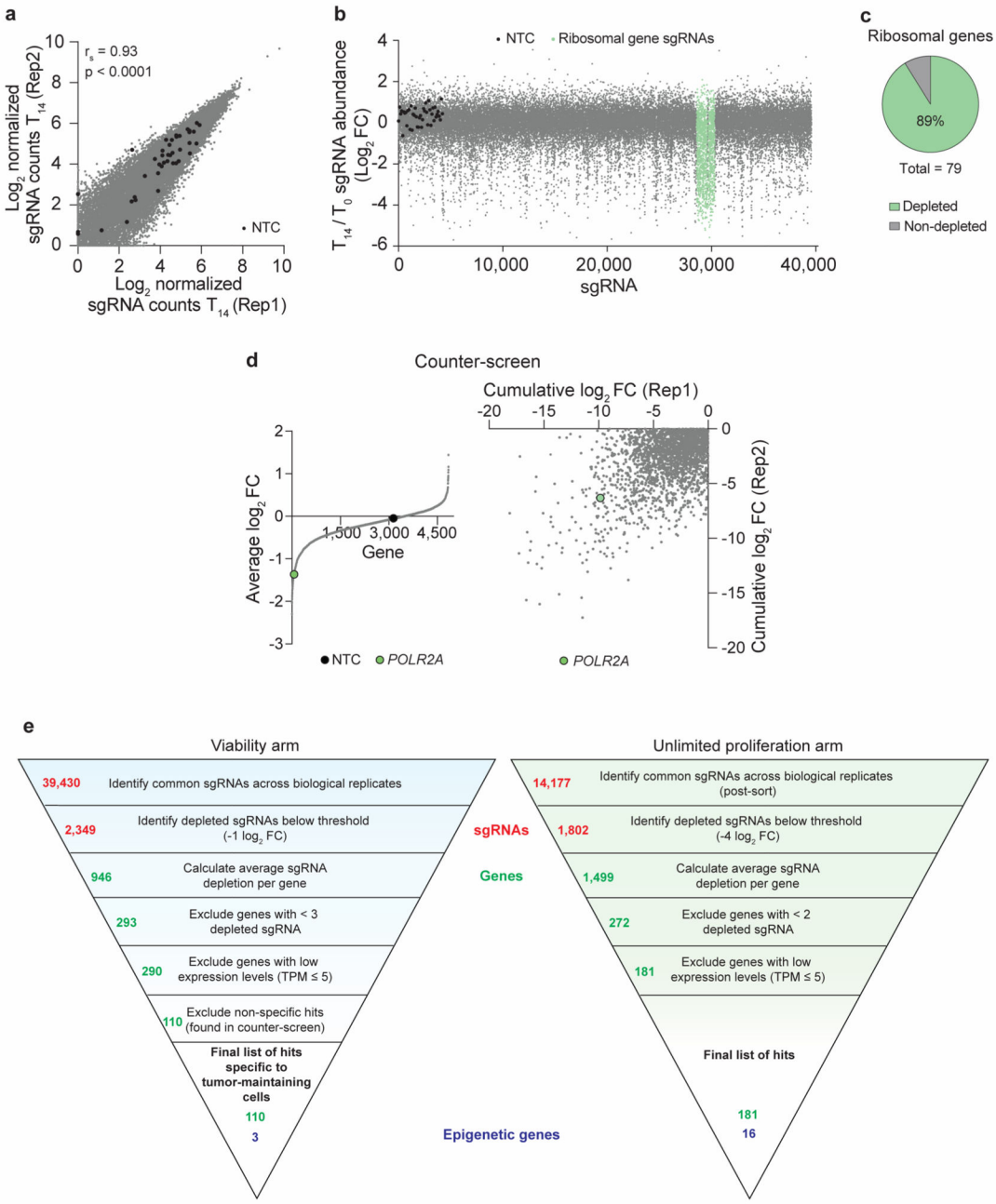
i. Quantification of *DCN* upregulation in *HRAS*-KO tumors as assessed by flow cytometry detecting expression of the *DCN*-GFP reporter. Individual values from two biologically independent samples. P-value from one-tailed Student's t-test.

j. Clonogenic assays showing loss of unlimited proliferative capacity by *HRAS*-KO TDF cells. Individual values from 2 biologically independent samples. P-value from one-tailed Student's t-test. Scale bar: 100 μm.

k. Growth kinetics of tumors induced by wild-type (WT) or *HRAS*-KO TDF cells. Values are average ± SEM from n = 4 biologically independent tumors. P-value from one-tailed Student's t-test at the last time point.

l. Growth kinetics of WT or *HRAS*-KO TDF cells as assessed by MTS assays. Values normalized to initial plating show the average ± SEM from n = 3 biologically independent samples. P-value from one-tailed Student's t-test at the last time point.

Source data are provided.



Extended Data Fig. 2. CRISPR-Cas9 screen controls

a. Relative abundance of sgRNAs in the two biological replicates of TDF cells transduced with the sgRNA library 14 d after Cas9 induction (T_{14}). NTC: non-targeting sgRNA controls. The statistical significance of the correlation between the two replicates (two-sided Spearman correlation test) and the corresponding correlation coefficient (r_s) are indicated.

b. Fold change (FC) in sgRNA abundance comparing the T_{14} and T_0 cell populations of replicate 2 (Rep 2). Individual, alphabetically ordered sgRNAs are shown. Non-targeting

Europe PMC Funders Author Manuscripts

sgRNA controls (NTC) are spread among targeting sgRNAs rather than clustered to ease their visualization.

c. Proportion of ribosomal genes depleted (average $\log_2$ FC ≤ -0.5) at T_{14}, indicating near saturation of the screen. Four out of 83 ribosomal genes targeted in the library are not expressed in TDF cells and are excluded from the analysis. The average value from both replicates is shown.

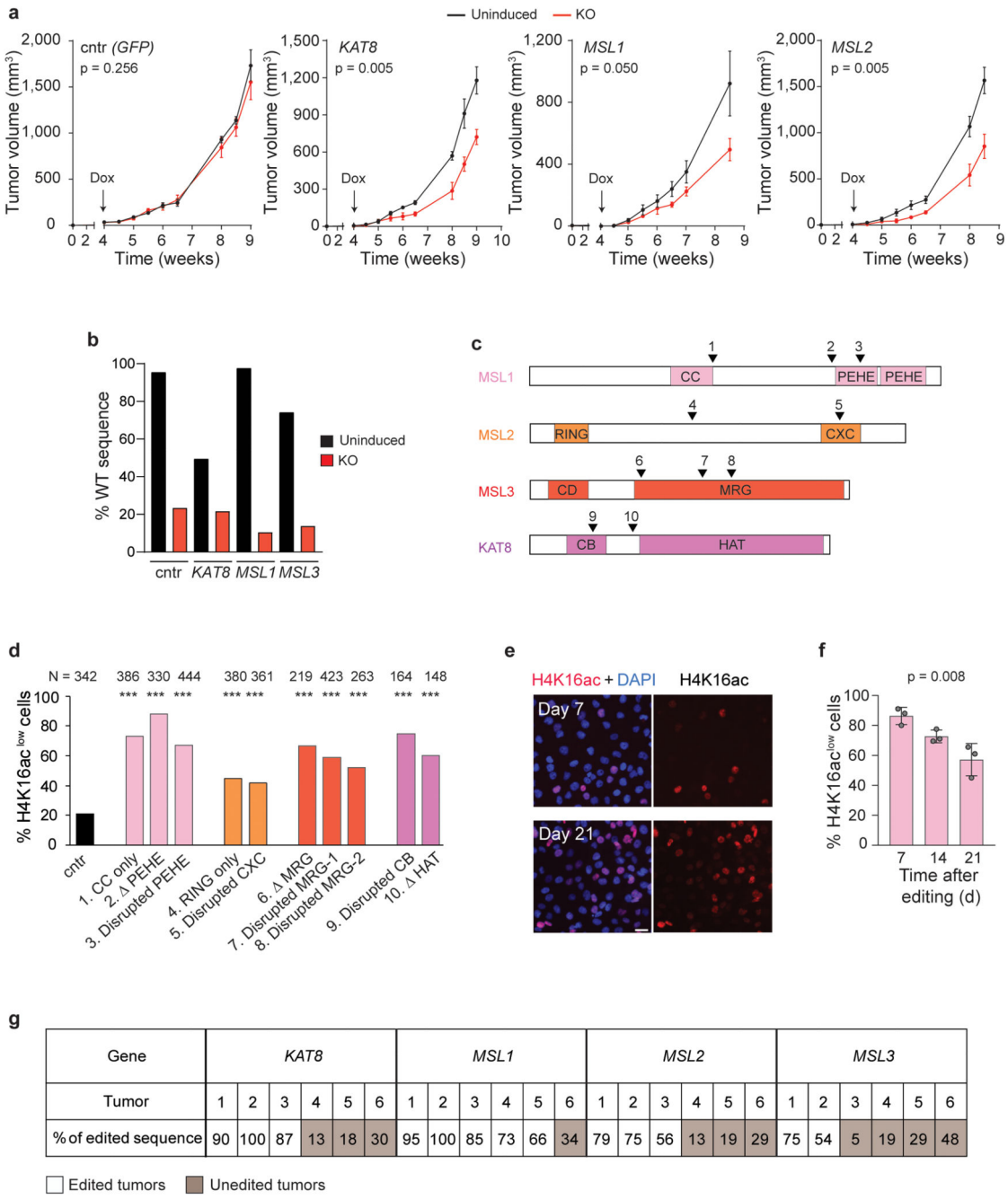
d. Fold change in sgRNA abundance comparing the T_{24} and T_0 cell populations in the counter-screen (see Methods). The average (left) or cumulative (right) value of all sgRNAs targeting each gene is plotted. The average of both replicates is shown in the graph on the left. The graph on the right displays only depleted genes and omits ribosomal genes due to the large number of sgRNAs targeting each gene (1-42 sgRNAs, median: 14 sgRNAs/gene), which makes them not comparable with other targeted genes. The essential gene *POLR2A* is shown as a positive control of sgRNA depletion. NTC: non-targeting sgRNA controls.

e. Filtering strategy used to select hits from both screen arms.

Source data are provided.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



g

Gene	*KAT8*						*MSL1*						*MSL2*						*MSL3*					
Tumor	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
% of edited sequence	90	100	87	13	18	30	95	100	85	73	66	34	79	75	56	13	19	29	75	54	5	19	29	48

Edited tumors Unedited tumors

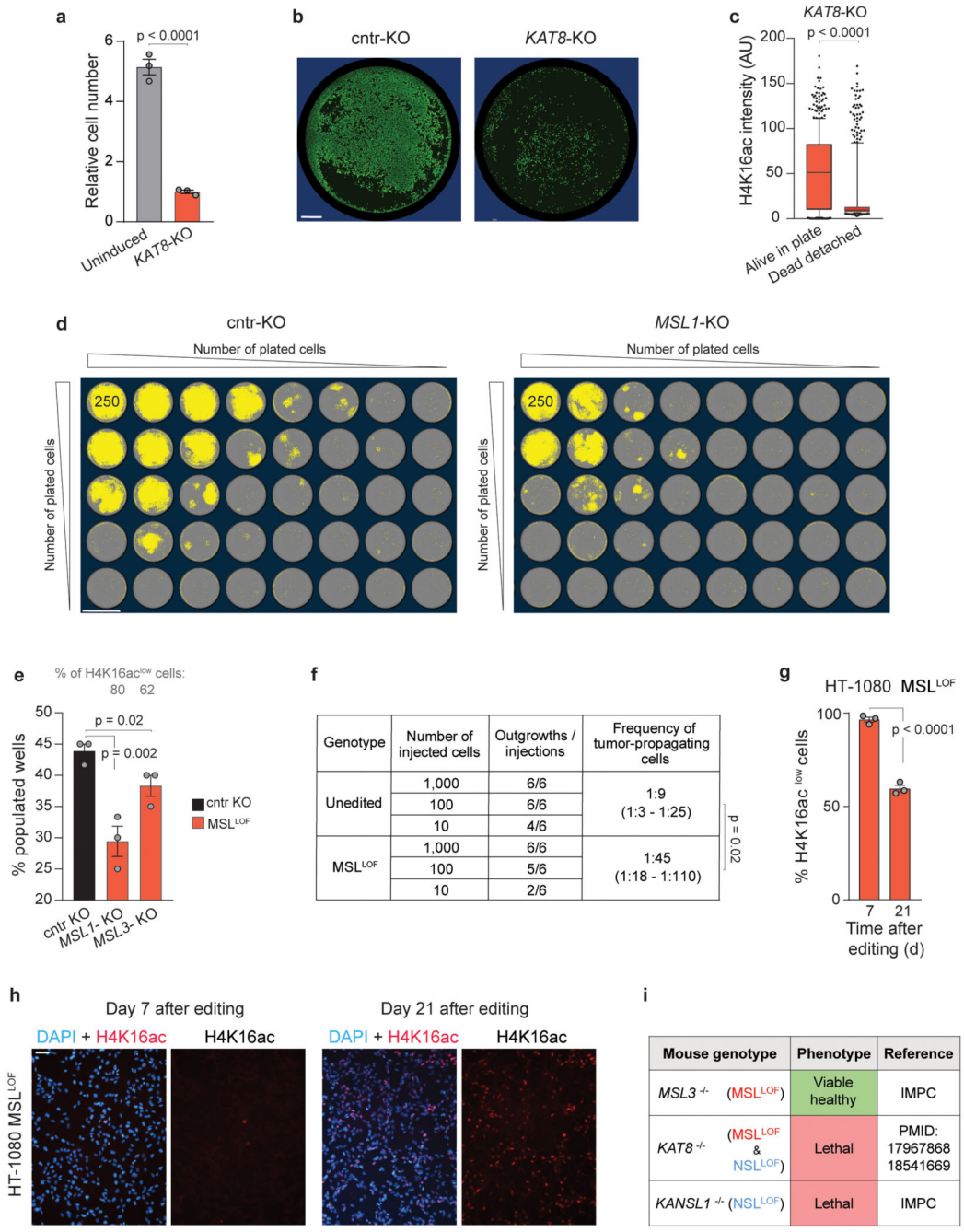
Extended Data Fig. 3. *In vivo* validation and characterization of primary hits

a. Growth kinetics of the indicated conditions in tumor maintenance assays performed as indicated in Fig. 1e. N = 4 biologically independent tumors. P-value from one-tailed Student's t-test at the last time point.

b. Quantification of gene editing efficiency of TDF cells of the indicated samples as assessed by Sanger sequencing and subsequent TIDE analysis https://tide.deskgen.com/ [52]. Cntr: control cells expressing GFP-targeting sgRNAs.

c. Schematic representation of the four MSL subunits and the mutants generated by genome editing. Arrowheads indicate the approximate location of sgRNAs used to generate truncated proteins lacking functional domains used in **d**. The characterized function of each domain is described in Supplementary Table 3. CC: coiled-coil. CD: chromo. CB: chromobarrel.
d. Quantification of MSL loss-of-function in the indicated MSL-mutant cell populations generated using the sgRNAs indicated in **c**. H4K16ac levels were measured by quantitative immunofluorescence and used as a readout of MSL function. The effect of *KAT8*-KO is underestimated as many cells H4K16ac^{low} died and detached from the plate prior to staining. Note that slight differences in the fraction of H4K16ac^{low} cells across mutants may be due to differences in sgRNA activity. All mutants resulted in reduction of the histone mark. The number of analyzed cells in each condition is shown. The statistical significance of the differences compared to a control sgRNA are indicated by three asterisks ($p < 0.0001$, two-tailed Fisher's test)
e-f. Representative images (**e**) and quantification (**f**) of H4K16ac levels in TDF cells at the indicated times after induction of *MSL1* knock-out. Values are average ± SEM from n = 3 biologically independent samples. P-value from one-way ANOVA. Scale bar: 20 μm.
g. Quantification of gene editing in the indicated tumors by Sanger sequencing and subsequent TIDE analysis. Tumors showing less than 50% edited sequence were considered unedited. Note the varying editing efficiency across all tumors, indicating the presence of wild-type cells also in tumors classified as "edited". Wild-type cells sustain tumor growth, leading to an underestimation of the effect induced by MSL disruption.
Source data are provided.

Europe PMC Funders Author Manuscripts



Extended Data Fig. 4. Disruption of the MSL complex impairs the long-term proliferative capacity of cancer cells

a. Quantification of the number of adherent, viable cells 8 d after induction of *KAT8* knock-out with doxycycline (*KAT8*-KO) and in the corresponding uninduced cells. Values are average ± SEM from n = 3 biologically independent samples. P-value from one-tailed Student's t-test.

b. Images of TDF cells 4 d after transfection of synthetic crRNAs-trRNAs targeting *KAT8* (*KAT8*-KO) or the GFP reporter (cntr-KO). Scale bar: 1 mm.

c. Quantification of H4K16ac levels in adherent, alive cells and detached cells recovered from the medium 4 d after induction of *KAT8* knock-out by doxycycline treatment. For boxplots, top, middle and bottom delimiters: 75^{th}, 50^{th}, 25^{th} percentiles; top and bottom whiskers: 90th and 10th percentiles. N = 471 and 488 biologically independent adherent and detached cells, respectively. P-value from one-tailed Mann Whitney's U-test. AU: arbitrary units.

d. Limiting dilution clonogenic assays of the indicated TDF cells 11 d after plating. Representative images of 96-well plates in which 2-fold serial dilutions of cells were plated. External wells are not shown (see Methods). Yellow shows the IncuCyte cell mask detecting growing clones. Scale bar: 6 mm.

e. Limiting dilution clonogenic assays using TDF cells. The percentage of wells containing at least one clone larger than 20 cells is plotted. Values are average ± SEM from n = 3 biologically independent samples. The knock-out efficiency, as assessed by quantification of $H4K16ac^{low}$ cells in each population is indicated. P-value from one-tailed Student's t-test.

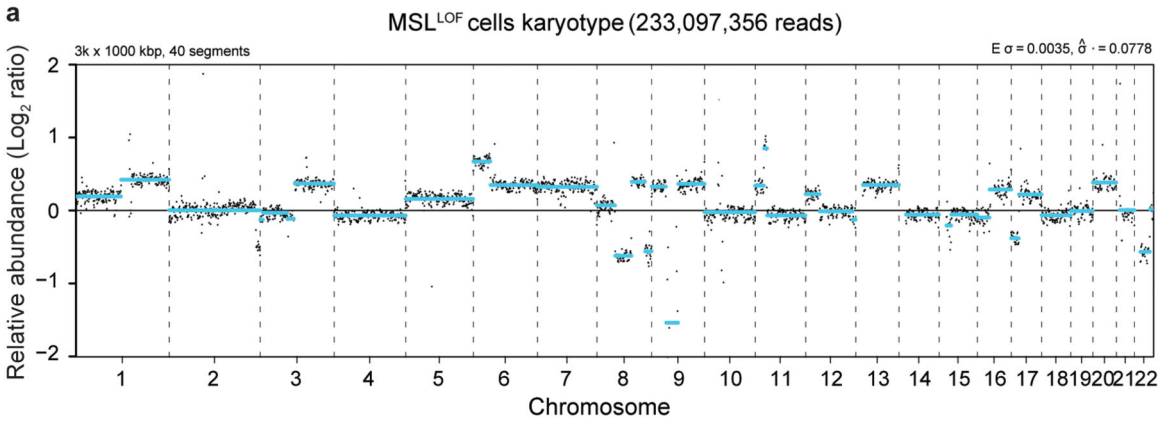
f. Limiting dilution transplantation assay into NSG mice using cells from unedited or MSL^{LOF} (*MSL3*-KO) tumors. The frequency of tumor-propagating cells (estimate with upper and lower limits) is indicated. P-value from χ^2 test.

g-h. Quantification (**g**) and representative images (**h**) of H4K16ac levels in HT-1080 cells at the indicated times after induction of *MSL1* knock-out. Values are average ± SEM from n = 3 areas in the well for a total of over 800 cells per condition. P-value from two-tailed Fisher's exact test. Scale bar: 100 μm.

i. Viability phenotype of the indicated knock-out mice. IMPC: International Mouse Phenotyping Consortium (see: https://www.mousephenotype.org/data/genes/MGI:1341851). PMID: Pubmed ID.

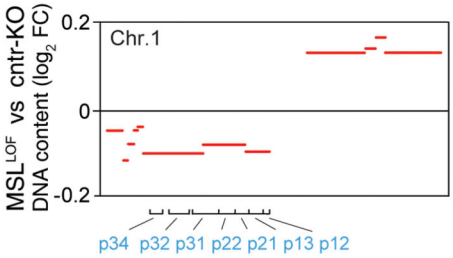
Source data are provided.

a



b

Chromosomal region	Differential expression (GSEA) MSL^{LOF} vs both controls	FDR q-value (GSEA)	Copy number changes (CNVkit) MSL^{LOF} vs WT	Copy number changes (CNVkit) MSL^{LOF} vs cntr-KO
CHR11P11	Upregulated	0	Gained (CN +1)	Gained (CN +1)
CHR11P13	Upregulated	0	Gained (CN +1)	Gained (CN +1)
CHR11P15	Upregulated	0	Gained (CN +1)	Gained (CN +1)
CHR11Q11	Downregulated	0	Lost (CN -1)	Lost (CN -1)
CHR15Q11	Upregulated	0	Gained (CN +1)	Gained (CN +1)
CHR17P11	Downregulated	0	Lost (CN -1)	Lost (CN -1)
CHR17P12	Downregulated	0	Lost (CN -1)	Lost (CN -1)
CHR17P13	Downregulated	0	Lost (CN -1)	Lost (CN -1)
CHR17Q25	Upregulated	0		Gained (CN +1)
CHR5P15	Upregulated	0	Gained (CN +1)	Gained (CN +1)
CHRXP21	Upregulated	0		
CHR17Q21	Upregulated	1.02E-04	Gained (CN +1)	Gained (CN +1)
CHR17Q22	Upregulated	1.96E-04		
CHR1P34	Downregulated	2.11E-04		
CHR1P13	Downregulated	1.90E-03		
CHR1P32	Downregulated	2.22E-03		
CHR1P21	Downregulated	2.44E-03	Lost (CN -1)	Lost (CN -1)
CHR1P22	Downregulated	3.17E-03		
CHR12Q21	Upregulated	3.19E-03		
CHR1P12	Downregulated	5.86E-03		
CHR17Q12	Upregulated	8.26E-03	Gained (CN +1)	Gained (CN +1)

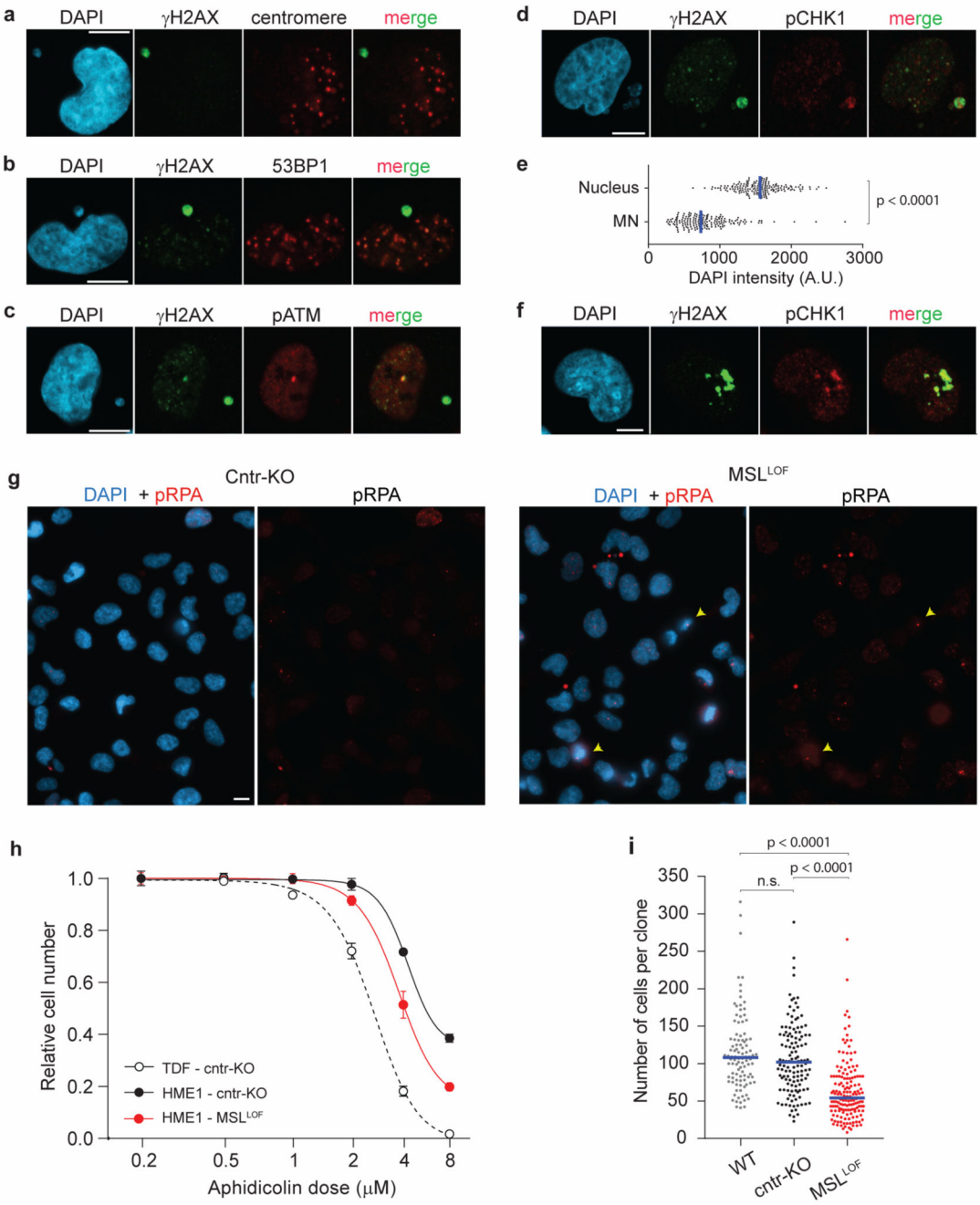


Extended Data Fig. 5. Gene expression and genomic abnormalities in MSL-disrupted cells

a. Karyotyping of MSL^{LOF} (*MSL3*-KO) TDF cells. The lines represent the consensus value for genomic segments that are either gained (positive values) or lost (negative values) relative to the rest of the genome (see Methods). The black dots are experimental values for 1 Mb intervals. X and Y chromosomes, which affects the performance of QDNASeq [66], are excluded from the analysis.

b. Correlation between large-scale changes in mRNA levels and DNA copy number alterations in MSL^{LOF} (*MSL3*-KO) cells. For each positional gene set identified as upregulated or downregulated in MSL^{LOF} cells by GSEA analysis, the relative copy number changes detected by CNVkit are shown. Chromosome 1 is shown to illustrate that even though CNVKit did not call copy number changes, likely because only a subset of cells in the population were affected by consistent gains or losses, the p arm showed a relative reduction compared to control cells (cntr-KO), correlating with apparent downregulation of

the corresponding genes. The slight difference compared to the Chr.1 pattern observed in the comparison with WT cells is likely due to subsampling of the population upon sgRNA transfection in the cntr-KO sample. Note that the genetic drivers in TDF cells are located in chromosomal regions showing highly significant changes, which indicates high prevalence in the population of MSL^{LOF} cells: HRAS on chr11p15 - gained, p53 on chr17p13 – lost, and hTERT on chr5 p15 – gained. This pattern suggests selection of less deleterious karyotypes in the population, compared to random gains or losses of other regions.



Europe PMC Funders Author Manuscripts

Extended Data Fig. 6. Accumulation of single-stranded DNA upon MSL disruption

a-g. Immunofluorescence microscopy of MSL^{LOF} TDF cells using the indicated antibodies (**a-d**, **f-g**) and quantification of DAPI intensity in nuclei and micronuclei (**e**). Nuclei were counterstained with DAPI. Yellow arrowheads in **g** indicate mitotic cells displaying p-RPA foci on chromosomes. N = 172 biologically independent cells collected from two experiments. P-value from one tailed paired Student's t-test. Scale bar: 10 μm.

h. Aphidicolin sensitivity assay comparing TDF and normal HME1 cells. Values are average ± SEM from n = 3 biologically independent samples after 5 days of growth. The non-linear fit of the experimental points is shown. Minor differences in TDF IC50 compared to Fig. 4g are due to a slightly different set up of the experiment (see Methods).

i. Clonogenic assays comparing the proliferative capacity of individual TDF cells of the indicated genotypes (See also Fig. 6a). N = 103, 131, 161 biologically independent clones for WT, cntr-KO and MSL^{LOF}(*MSL3*-KO), respectively, collected from two experiments. Horizontal lines: median values. P-value from one-way ANOVA followed by Dunn's test. N.s.: non-significant (p-value = 0.9495).

Source data are provided.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the Crick Biological Research Facility for help with animal work, the Crick Advanced Sequencing for preparing and sequencing NGS libraries, the Crick Light Microscopy for help with fluorescent microscopy, and the Flow Cytometry for help with cell sorting. We thank Thomas Metz and Oncotest-Charles River for sharing PDX lines and Panos Zalmas and Antonio Tedeschi for useful discussions and sharing reagents. This work was supported by The Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001152), the UK Medical Research Council (FC001152), and the Wellcome Trust (FC001152), and by the CRUK Drug Discovery Award (C50796/A19448) provided by Cancer Research UK and Bayer Healthcare. YD and MB were funded by the Intramural Research Program of the National Institutes of Health, USA. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Data Availability

The accompanying RNA-seq and DNA-seq data set is available through GEO: GSE144019. Gene sets used for GSEA are available through the MSigDB database v7.2 at https://www.gsea-msigdb.org/gsea/msigdb/index.jsp. Source data are provided with this paper. All other data is supporting the findings of this study are available from the corresponding author on reasonable request.

References

1. Mair B, Moffat J, Boone C, Andrews BJ. Genetic interaction networks in cancer cells. Curr Opin Genet Dev. 2019; 54:64–72. [PubMed: 30974317]
2. Witkiewicz AK, et al. Targeting the Vulnerability of RB Tumor Suppressor Loss in Triple-Negative Breast Cancer. Cell Rep. 2018; 22:1185–1199. [PubMed: 29386107]
3. Kumar MS, et al. The GATA2 transcriptional network is requisite for RAS oncogene-driven non-small cell lung cancer. Cell. 2012; 149:642–655. [PubMed: 22541434]

4. Li L, et al. Identification of DHODH as a therapeutic target in small cell lung cancer. Sci Transl Med. 2019; 11
5. He S, Nakada D, Morrison SJ. Mechanisms of stem cell self-renewal. Annu Rev Cell Dev Biol. 2009; 25:377–406. [PubMed: 19575646]
6. Dick JE. Stem cell concepts renew cancer research. Blood. 2008; 112:4793–4807. [PubMed: 19064739]
7. Hayflick L. Mortality and immortality at the cellular level. A review. Biochemistry (Mosc). 1997; 62:1180–1190. [PubMed: 9467840]
8. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell. 1997; 88:593–602. [PubMed: 9054499]
9. Dawson MA. The cancer epigenome: Concepts, challenges, and therapeutic opportunities. Science. 2017; 355:1147–1152. [PubMed: 28302822]
10. Wainwright EN, Scaffidi P. Epigenetics and Cancer Stem Cells: Unleashing, Hijacking, and Restricting Cellular Plasticity. Trends in Cancer. 2017; 3:15.
11. Komlodi-Pasztor E, Sackett D, Wilkerson J, Fojo T. Mitosis is not a key target of microtubule agents in patient tumors. Nat Rev Clin Oncol. 2011; 8:244–250. [PubMed: 21283127]
12. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nature reviews Cancer. 2012; 12:323–334. [PubMed: 22513401]
13. Tirosh I, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature. 2016; 539:309–313. [PubMed: 27806376]
14. Barabe F, Kennedy JA, Hope KJ, Dick JE. Modeling the initiation and progression of human acute leukemia in mice. Science. 2007; 316:600–604. [PubMed: 17463288]
15. Lan X, et al. Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. Nature. 2017; 549:227–232. [PubMed: 28854171]
16. Scaffidi P, Misteli T. In vitro generation of human cells with cancer stem cell properties. Nature cell biology. 2011; 13:1051–1061. [PubMed: 21857669]
17. Torres CM, et al. The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity. Science. 2016; 353
18. Mortimer T, et al. Redistribution of EZH2 promotes malignant phenotypes by rewiring developmental programmes. EMBO Rep. 2019; 20:e48155. [PubMed: 31468686]
19. Morales Torres C, et al. Selective inhibition of cancer cell self-renewal through a Quisinostat-histone H1.0 axis. Nature communications. 2020; 11:1792.
20. Menyhart O, et al. Guidelines for the selection of functional assays to evaluate the hallmarks of cancer. Biochim Biophys Acta. 2016; 1866:300–319. [PubMed: 27742530]
21. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014; 343:80–84. [PubMed: 24336569]
22. Keller CI, Akhtar A. The MSL complex: juggling RNA-protein interactions for dosage compensation and beyond. Curr Opin Genet Dev. 2015; 31:1–11. [PubMed: 25900149]
23. Shogren-Knaak M, et al. Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science. 2006; 311:844–847. [PubMed: 16469925]
24. Galupa R, Heard E. X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation. Annu Rev Genet. 2018; 52:535–566. [PubMed: 30256677]
25. Gupta A, et al. Involvement of human MOF in ATM function. Molecular and cellular biology. 2005; 25:5292–5305. [PubMed: 15923642]
26. Gupta A, et al. The mammalian ortholog of Drosophila MOF that acetylates histone H4 lysine 16 is essential for embryogenesis and oncogenesis. Molecular and cellular biology. 2008; 28:397–409. [PubMed: 17967868]
27. McDonald OG, et al. Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. Nat Genet. 2017; 49:367–376. [PubMed: 28092686]
28. Zhu L, et al. Expression of hMOF, but not HDAC4, is responsible for the global histone H4K16 acetylation in gastric carcinoma. Int J Oncol. 2015; 46:2535–2545. [PubMed: 25873202]

29. Ravens S, et al. Mof-associated complexes have overlapping and unique roles in regulating pluripotency in embryonic stem cells and during differentiation. Elife. 2014; 3
30. Dickinson ME, et al. High-throughput discovery of novel developmental phenotypes. Nature. 2016; 537:508–514. [PubMed: 27626380]
31. Tonnessen-Murray CA, Lozano G, Jackson JG. The Regulation of Cellular Functions by the p53 Protein: Cellular Senescence. Cold Spring Harb Perspect Med. 2017; 7
32. Sammons MA, Zhu J, Drake AM, Berger SL. TP53 engagement with the genome occurs in distinct local chromatin environments via pioneer factor activity. Genome research. 2015; 25:179–188. [PubMed: 25391375]
33. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:15545–15550. [PubMed: 16199517]
34. Potapova TA, Zhu J, Li R. Aneuploidy and chromosomal instability: a vicious cycle driving cellular evolution and cancer genome chaos. Cancer Metastasis Rev. 2013; 32:377–389. [PubMed: 23709119]
35. Lee HS, et al. A new assay for measuring chromosome instability (CIN) and identification of drugs that elevate CIN in cancer cells. BMC Cancer. 2013; 13:252. [PubMed: 23694679]
36. Blackford AN, Jackson SP. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. Molecular cell. 2017; 66:801–817. [PubMed: 28622525]
37. Wilhelm T, Said M, Naim V. DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. Genes (Basel). 2020; 11
38. Groth A, Rocha W, Verreault A, Almouzni G. Chromatin challenges during DNA replication and repair. Cell. 2007; 128:721–733. [PubMed: 17320509]
39. Thompson SL, Bakhoum SF, Compton DA. Mechanisms of chromosomal instability. Curr Biol. 2010; 20:R285–295. [PubMed: 20334839]
40. Burrell RA, et al. Replication stress links structural and numerical cancer chromosomal instability. Nature. 2013; 494:492–496. [PubMed: 23446422]
41. Umbreit NT, et al. Mechanisms generating cancer genome complexity from a single cell division error. Science. 2020; 368
42. Tang YC, Amon A. Gene copy-number alterations: a cost-benefit analysis. Cell. 2013; 152:394–405. [PubMed: 23374337]
43. Zhu J, Tsai H-J, Gordon MR, Li R. Cellular Stress Associated with Aneuploidy. Developmental cell. 2018; 44:420–431. [PubMed: 29486194]
44. Sansregret L, Vanhaesebroeck B, Swanton C. Determinants and clinical implications of chromosomal instability in cancer. Nat Rev Clin Oncol. 2018; 15:139–150. [PubMed: 29297505]
45. Santaguida S, Tighe A, D'Alise AM, Taylor SS, Musacchio A. Dissecting the role of MPS1 in chromosome biorientation and the spindle checkpoint through the small molecule inhibitor reversine. The Journal of cell biology. 2010; 190:73–87. [PubMed: 20624901]
46. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:3983–3988. [PubMed: 12629218]
47. Janssen A, Kops GJ, Medema RH. Elevating the frequency of chromosome mis-segregation as a strategy to kill tumor cells. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:19108–19113. [PubMed: 19855003]
48. Tardif KD, et al. Characterization of the cellular and antitumor effects of MPI-0479605, a small-molecule inhibitor of the mitotic kinase Mps1. Mol Cancer Ther. 2011; 10:2267–2275. [PubMed: 21980130]
49. Martinez R, et al. Mitotic Checkpoint Kinase Mps1 Has a Role in Normal Physiology which Impacts Clinical Utility. PloS one. 2015; 10:e0138616. [PubMed: 26398286]
50. Wang Q, et al. BUBR1 deficiency results in abnormal megakaryopoiesis. Blood. 2004; 103:1278–1285. [PubMed: 14576056]

51. Kusakabe K, et al. Discovery of imidazo[1,2-b]pyridazine derivatives: selective and orally available Mps1 (TTK) kinase inhibitors exhibiting remarkable antiproliferative activity. J Med Chem. 2015; 58:1760–1775. [PubMed: 25625617]
52. Kim D, et al. Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. Nature structural & molecular biology. 2010; 17:1027–1029.
53. Kadlec J, et al. Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. Nature structural & molecular biology. 2011; 18:142–149.
54. Baell JB, et al. Inhibitors of histone acetyltransferases KAT6A/B induce senescence and arrest tumour growth. Nature. 2018; 560:253–257. [PubMed: 30069049]
55. MacPherson L, et al. HBO1 is required for the maintenance of leukaemia stem cells. Nature. 2020; 577:266–270. [PubMed: 31827282]
56. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS computational biology. 2016; 12:e1004873. [PubMed: 27100738]
57. Henser-Brownhill T, Monserrat J, Scaffidi P. Generation of an arrayed CRISPR-Cas9 library targeting epigenetic regulators: from high-content screens to in vivo assays. Epigenetics. 2017; 12:1065–1075. [PubMed: 29327641]
58. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]
59. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic acids research. 2014; 42:e168. [PubMed: 25300484]
60. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. 2011; 17:10–12.
61. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics. 2011; 12:323. [PubMed: 21816040]
62. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21. [PubMed: 23104886]
63. Karolchik D, et al. The UCSC Table Browser data retrieval tool. Nucleic acids research. 2004; 32:D493–D496. [PubMed: 14681465]
64. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014; 15:550. [PubMed: 25516281]
65. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]
66. Scheinin I, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. Genome research. 2014; 24:2022–2032. [PubMed: 25236618]
67. Kyriacou E, Heun P. High-resolution mapping of centromeric protein association using APEX-chromatin fibers. Epigenetics & chromatin. 2018; 11:68. [PubMed: 30445992]
68. Bui M, et al. Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. Cell. 2012; 150:317–326. [PubMed: 22817894]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

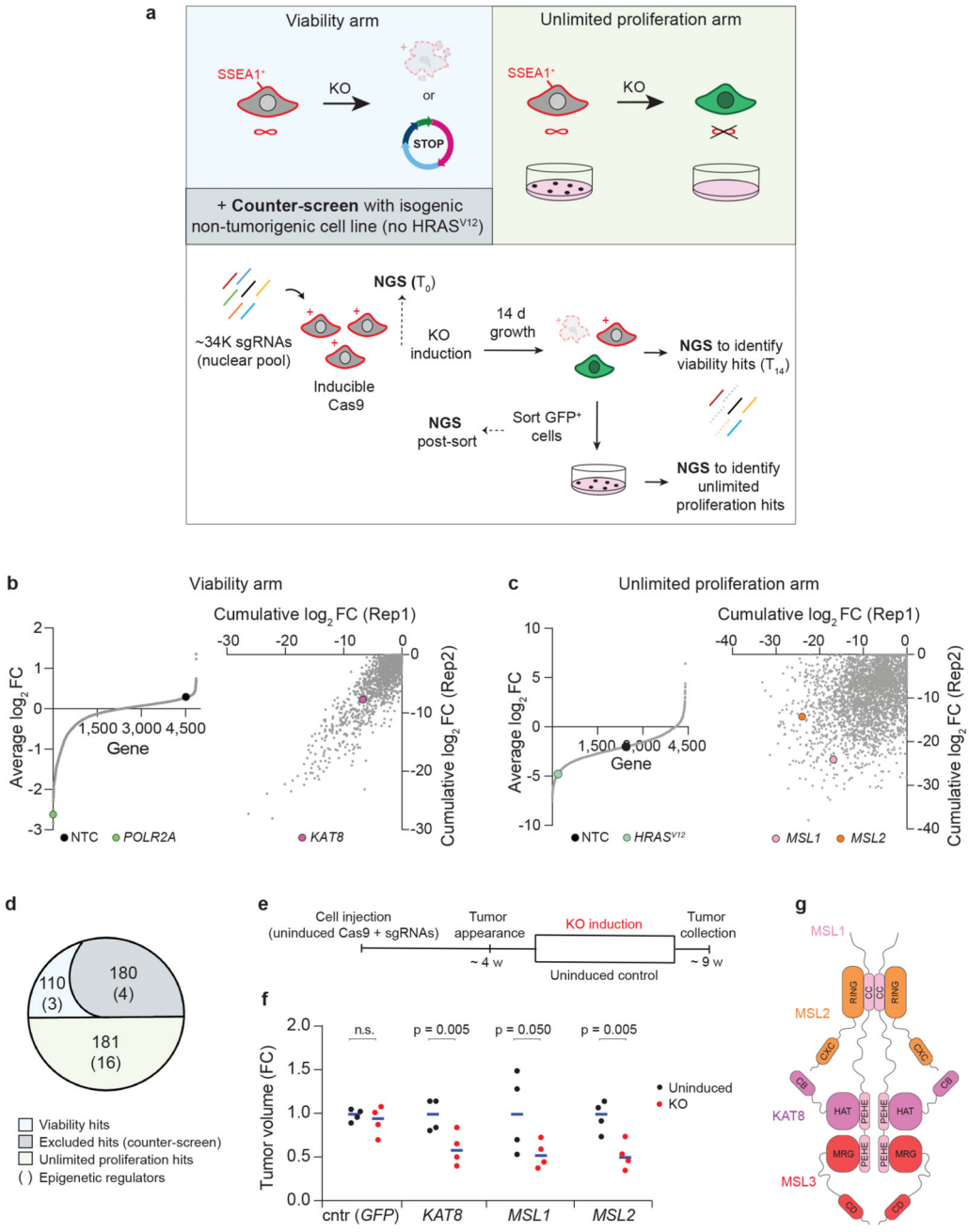


Figure 1. A CRISPR-Cas9 screen identifies proteins important for tumor maintenance.

a. Schematic of the CRISPR-Cas9 screen. Top panels: experimental design to identify sgRNAs inducing SSEA1$^+$ cell death (transparent cell) or cell cycle arrest (colored circle) specifically (left), and sgRNAs depriving cells of unlimited proliferative capacity (red infinity sign), as assessed by expression of the *DCN*-GFP reporter and loss of clonogenic ability (right). Bottom panel: key steps of the selection process. NGS: next generation sequencing.

b. Fold change (FC) in sgRNA abundance comparing the T_{14} and T_0 cell populations. The average (left) or cumulative (right) value of all sgRNAs targeting each gene is plotted. The average of two biologically independent replicates is shown in the graph on the left. The graph on the right displays only depleted genes and omits ribosomal genes. The essential gene *POLR2A* is shown as a positive control. NTC: non-targeting sgRNA controls.
c. Fold change in sgRNA abundance comparing cells that generated colonies in semi-solid medium and sorted $SSEA1^-/GFP^+$ cells prior to plating. The average (left) or cumulative (right) value of all sgRNAs targeting each gene is plotted. Data for the average of two biologically independent replicates is shown in the graph on the left. The graph on the right only displays depleted sgRNAs. $HRAS^{V12}$ is shown as a positive control. NTC: non-targeting sgRNA controls.
d. Number of identified hits in both screen arms.
e. Experimental design to assess the effect of *in vivo* gene knock-out on tumor maintenance.
f. Endpoint tumor volume for the indicated conditions in tumor maintenance assays performed as indicated in **e**. Cntr: control cells in which GFP was targeted. Values are shown as relative to the median (blue lines) of each uninduced condition. N = 4 biologically independent tumors. See also Extended Fig. 3a. The statistical significance of the difference between the uninduced and KO condition for each gene is indicated (one-tailed Student's t-test). N.s.: non-significant (p-value = 0.256).
g. Schematic representation of the MSL complex, a homotetramer of the four indicated subunits. Protein domains with characterized functions are labelled. Diagram adapted from Keller *et al.*[22]. CC: coiled-coil. CD: chromo. CB: chromobarrel.
Source data are provided.

Europe PMC Funders Author Manuscripts

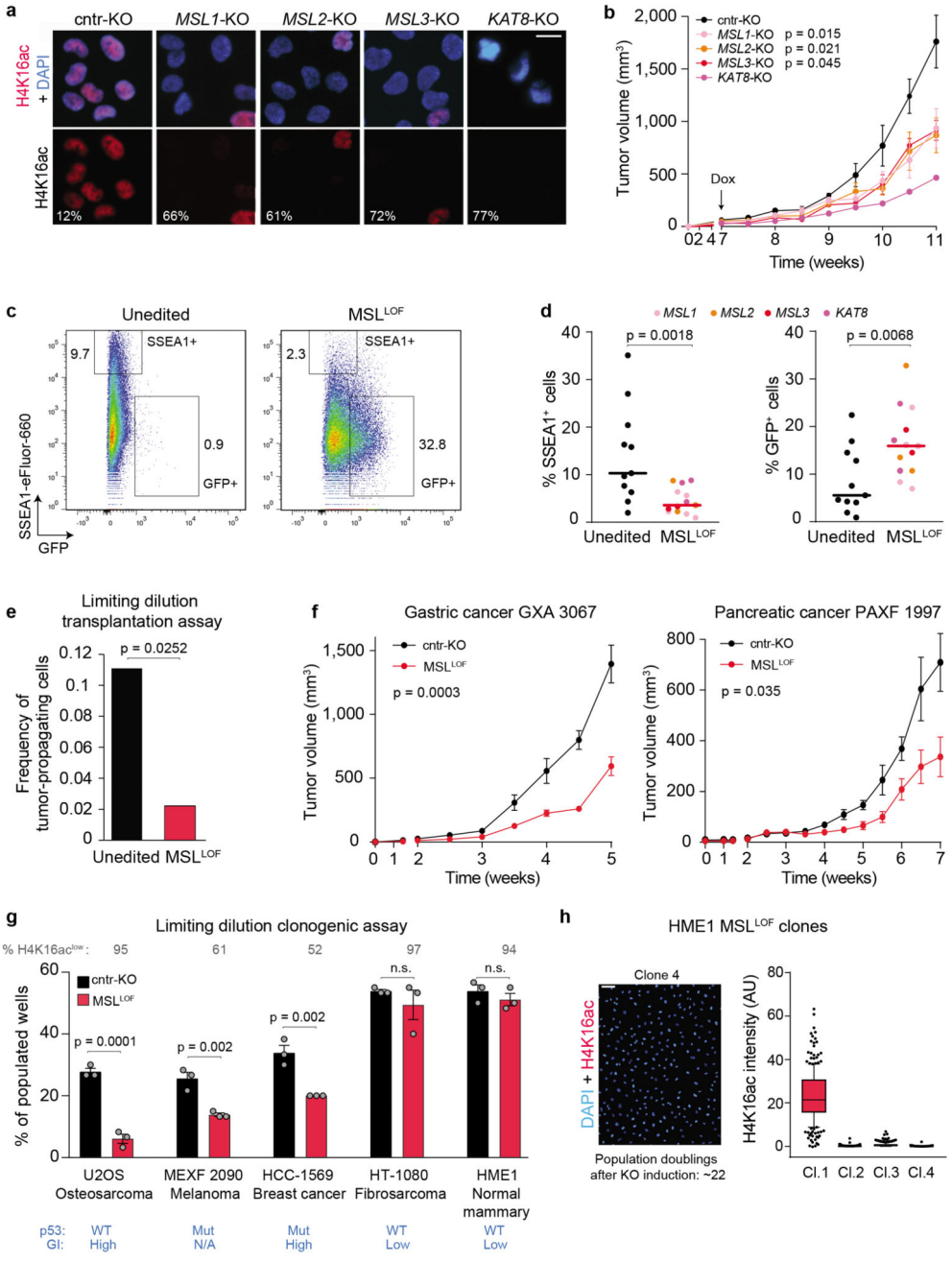


Figure 2. A functional MSL complex is required for tumor maintenance.

a. Immunodetection of H4K16ac in the indicated TDF cells 14 d after knock-out, with percentages of H4K16ac^{low} cells indicated. Scale bar: 10 μm.

b. Tumor maintenance assay. Growth kinetics of tumors induced by the indicated TDF cells. Dox: doxycycline. Values are average ± SEM from n = 4, 5, 3, 3 and 2 biologically independent tumors for cntr-KO, *MSL1*-KO, *MSL2*-KO, *MSL3*-KO, *KAT8*-KO, respectively. Only edited tumors, as assessed in Extended Fig. 3g, are averaged. P-value compared to cntr-KO at last time point from one-tailed Student's t-test.

c-d. Flow cytometry analysis (**c**) and quantification of the indicated subsets of cells in tumor samples (**d**). N = 11 and 13 biologically independent tumors for unedited and MSL^{LOF}, respectively. P-value from one-tailed Student's t-test.

e. Limiting dilution transplantation assay into NSG mice using cells from unedited or MSL^{LOF} (*MSL3*-KO) tumors (n = 18 independent injections). P-value from χ^2 test.

f. Growth kinetics of PDX models of gastric and pancreatic cancer, comparing MSL^{LOF} (*MSL3*-KO) and control tumors. Values are average ± SEM from n = 6 biologically independent tumors. P-value from one-tailed Student's t-test at last time point.

g. Limiting dilution clonogenic assays quantifying the long-term proliferative capacity of the indicated cell lines upon *MSL1* knock-out (MSL^{LOF}). The percentage of wells containing at least one clone larger than 20 cells is plotted. Values are average ± SEM from n = 3 biologically independent samples. P-value from one-tailed Student's t-test. The percentage of $H4K16ac^{low}$ cells in each MSL^{LOF} sample, the p53 status and the level of genomic instability (GI) of each line are indicated. Mut: mutated. N/A: not available. N.s.: non-significant ($p > 0.1$).

h. Representative image and quantification of H4K16ac levels in monoclonal populations isolated from HME1 cells after induction of *MSL1* knock-out. One clone that escaped knock-out and three with complete H4K16ac loss are shown. For boxplots, top, middle and bottom delimiters: 75^{th}, 50^{th}, 25^{th} percentiles; top and bottom whiskers: 90^{th} and 10^{th} percentiles. N = 250 biologically independent cells. Scale bar: 100 μm.

a, **h**, Data are representative of ten and two independent experiments, respectively. Source data are provided.

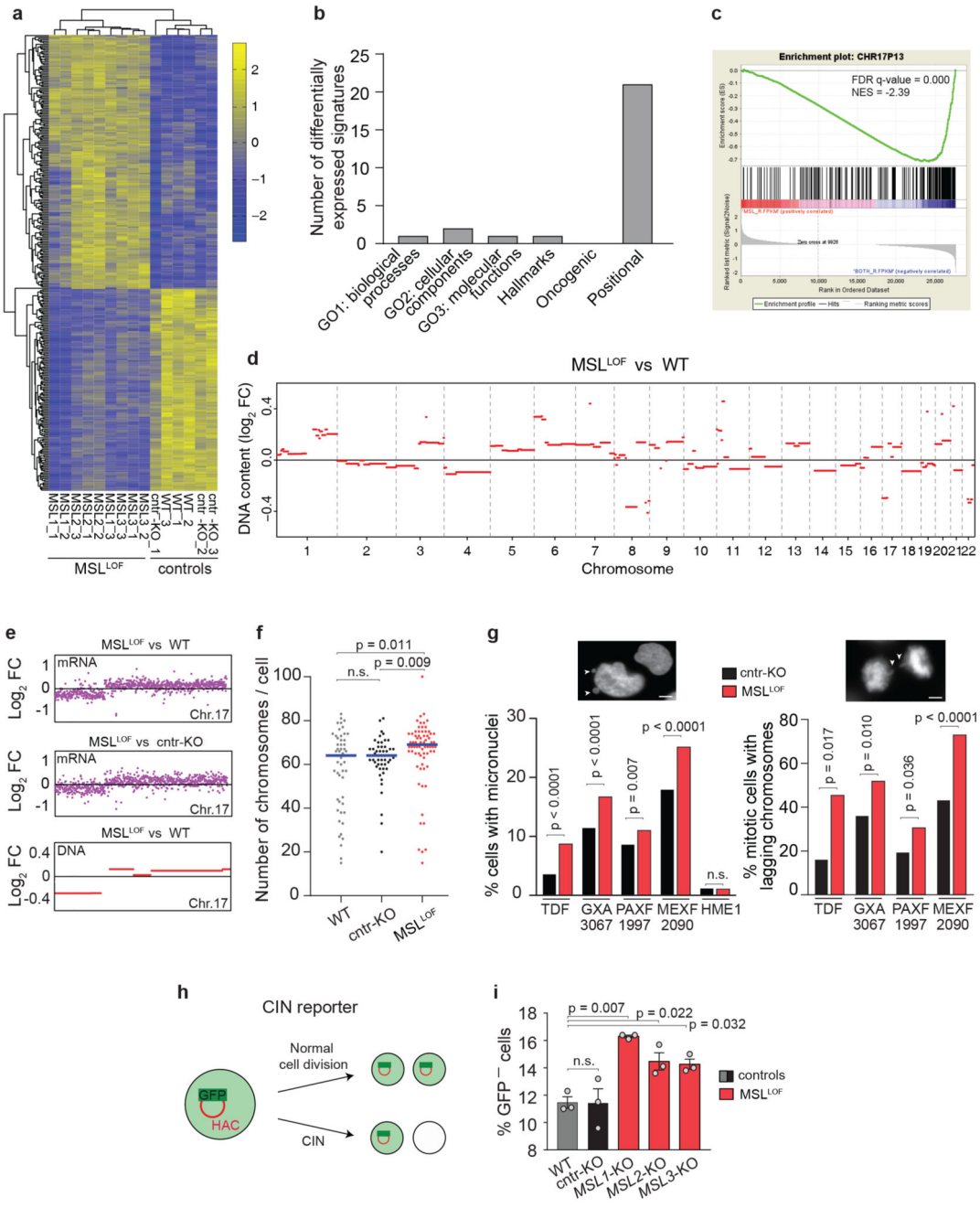


Figure 3. MSL disruption promotes chromosomal instability.

a. Heatmap of relative expression of 314 DEGs in individual biological replicates of the indicated samples. Colors represent row-scaled $\log_2$-transformed TPM values.

b. Quantification of gene signatures from the indicated MSigDB gene set collections that are differentially expressed (FDR ≤ 0.01) in MSL^{LOF} cells (*MSL1*-KO, *MSL2*-KO and *MSL3*-KO) compared to control cells (WT and cntr-KO), as assessed by GSEA. See also Supplementary Table 7.

c. Example of a positional gene set negatively correlating with MSL^{LOF} samples.

Europe PMC Funders Author Manuscripts

d. Copy number alterations in MSL^{LOF} (*MSL3*-KO) TDF cells. Red lines: estimated genomic segments either gained (positive values) or lost (negative values) relative to WT cells. Sex chromosomes, which affect the performance of CNVKit[56], are excluded.
e. Comparison between the fold change (FC) in DNA and mRNA levels in MSL^{LOF} (*MSL3*-KO) TDF cells relative to WT or cntr-KO cells for chromosome 17. See also Extended Fig. 5b. FC in mRNA levels for each expressed gene (TPM > 1), ordered by coordinates are shown. Y axis scales are not comparable.
f. Quantification of chromosome number in MSL^{LOF} (*MSL3*-KO) and control TDF cells. N = 52, 50 and 74 biologically independent cells for WT, cntr-KO and MSL^{LOF}, respectively. Data are representative of two independent experiments. Horizontal lines: median values. P-value from one-way ANOVA followed by Dunn's test. N.s: non-significant (p-value > 0.99).
g. Quantification of mitotic defects in the indicated samples. For each bar, from left to right, N = 2576, 2857, 2119, 2902, 2120, 1651, 1434, 962, 1944, 1686, 25, 33, 125, 112, 114, 157, 123, 85 biologically independent cells collected from two experiments. P-value from one-sided Fisher's exact test. Arrows: micronuclei (left) and lagging chromosomes (right). Scale bar: 5 μm.
h. Schematics of the CIN reporter allowing estimation of CIN rate by loss of the GFP-expressing HAC (red circle).
i. Quantification of HAC loss by flow cytometry analysis of the indicated samples 16 d after gene knock-out. Values are average ± SEM from three biologically independent samples. P-value relative to unedited wild-type cells from one-way ANOVA followed by Dunn's test. N.s: non-significant (p > 0.99).
Source data are provided.

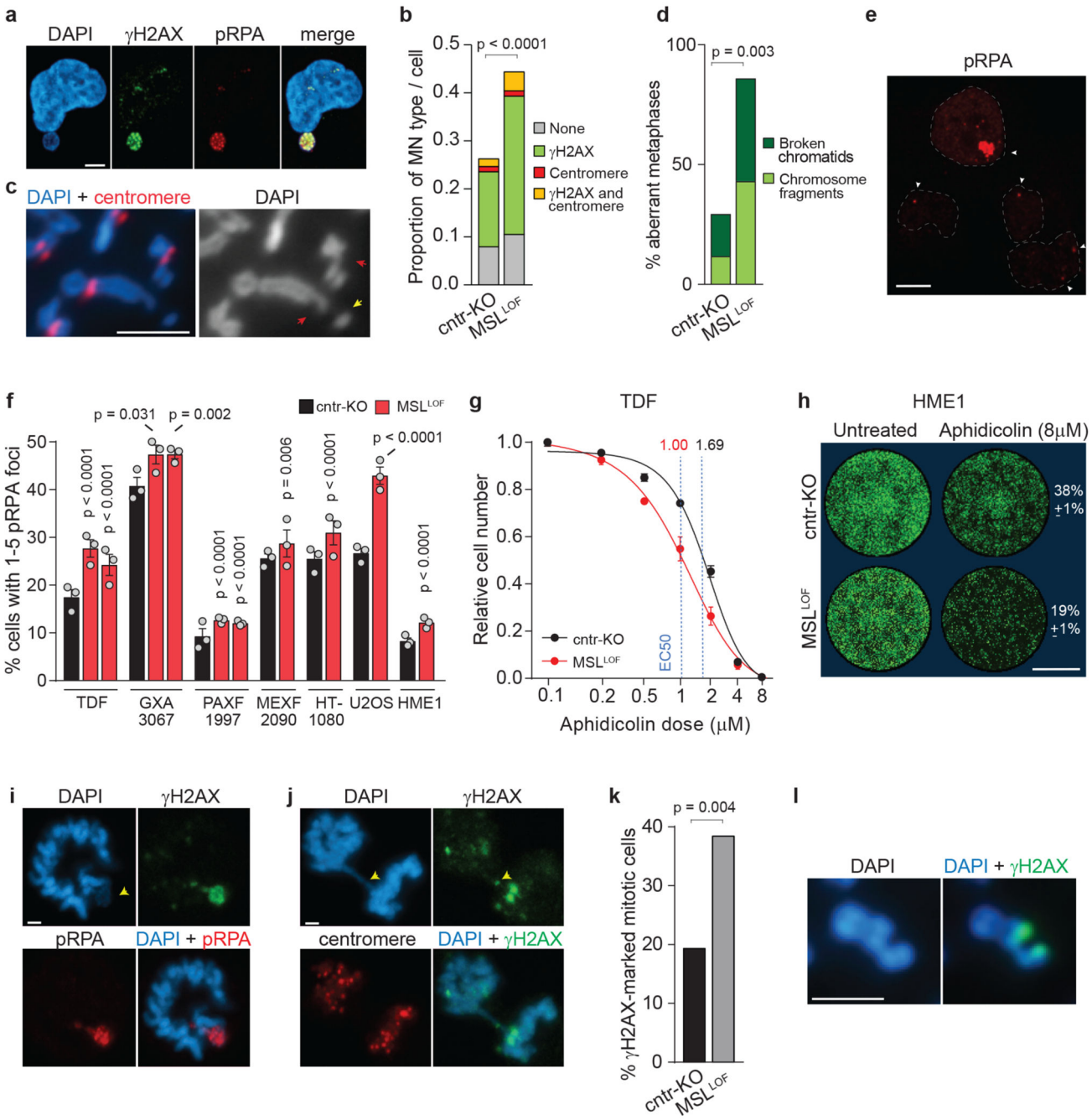


Figure 4. Accumulation of single-stranded DNA upon MSL disruption enhances chromosomal instability.

a-d. Immunofluorescence microscopy of MSL^{LOF} (*MSL1*-KO) TDF cells with the indicated antibodies (**a**, **c**) and quantification (**b**, **d**). N = 1017, 648 independent cells (**b**) or n = 17, 14 independent metaphases (**d**) for cntr-KO and MSL^{LOF}, respectively. P-value for the fraction of cells containing γH2AX-marked micronuclei (**b**) or aberrant chromosomes (**d**) from two-tailed Fisher's exact test. Red and yellow arrowheads: broken chromatids and chromosome fragments. Scale bar: 5 μm.

Europe PMC Funders Author Manuscripts

e. Immunofluorescence microscopy of MSL^{LOF} (*MSL1*-KO) TDF cells with the indicated antibody. Arrowheads: pRPA foci. Dashed lines: nuclear edge. Scale bar: 10 μm.
f. Quantification of interphase cells containing pRPA foci. For TDF, GXA 3067 and PAXF 1997: *MSL1*-KO (left) and *MSL3*-KO (right) shown. For all others: *MSL1*-KO. From left to right, n = 2267, 2019, 1933, 803, 433, 923, 5606, 2730, 4804, 3525, 2270, 2028, 1951, 1285, 718, 3162, 3066 cells across 108 fields. Values are average ± SEM from n = 3 well areas. P-value relative to each cntr-KO sample from two-tailed Fisher's exact test).
g. Aphidicolin sensitivity assay comparing MSL^{LOF} (*MSL1*-KO) and control TDF cells. Values are average ± SEM from n = 3 biologically independent samples after 5 days of growth. The non-linear fit of the experimental points and the estimated EC50 are shown. Similar results were obtained in two additional experiments.
h. Images of MSL^{LOF} (*MSL1*-KO) and control HME1 cells after five days of growth in the indicated conditions. Cells are visualized by SYTOX green. The percentage of cells in the aphidicolin condition relative to the corresponding untreated one (average value from n = 3 biologically independent samples ± SEM) is indicated. See also Extended Fig. 6h. Scale bar: 8 mm.
i-k. Immunofluorescence microscopy of mitotic MSL^{LOF} TDF cells (i-**j**) or chromosome spreads (**i**) using the indicated antibodies, and quantification (**k**). Arrowheads: large DNA region (**i**) and focal area on a chromosome bridge (**j**) with low DAPI intensity. N = 107, 93 independent mitotic cells for cntr-KO and MSL^{LOF}, respectively. P-value from two-tailed Fisher's exact test. Scale bar: 2 μm.
a-**e**, **i**-**l**, Data are representative of two independent experiments. Source data are provided.

Europe PMC Funders Author Manuscripts

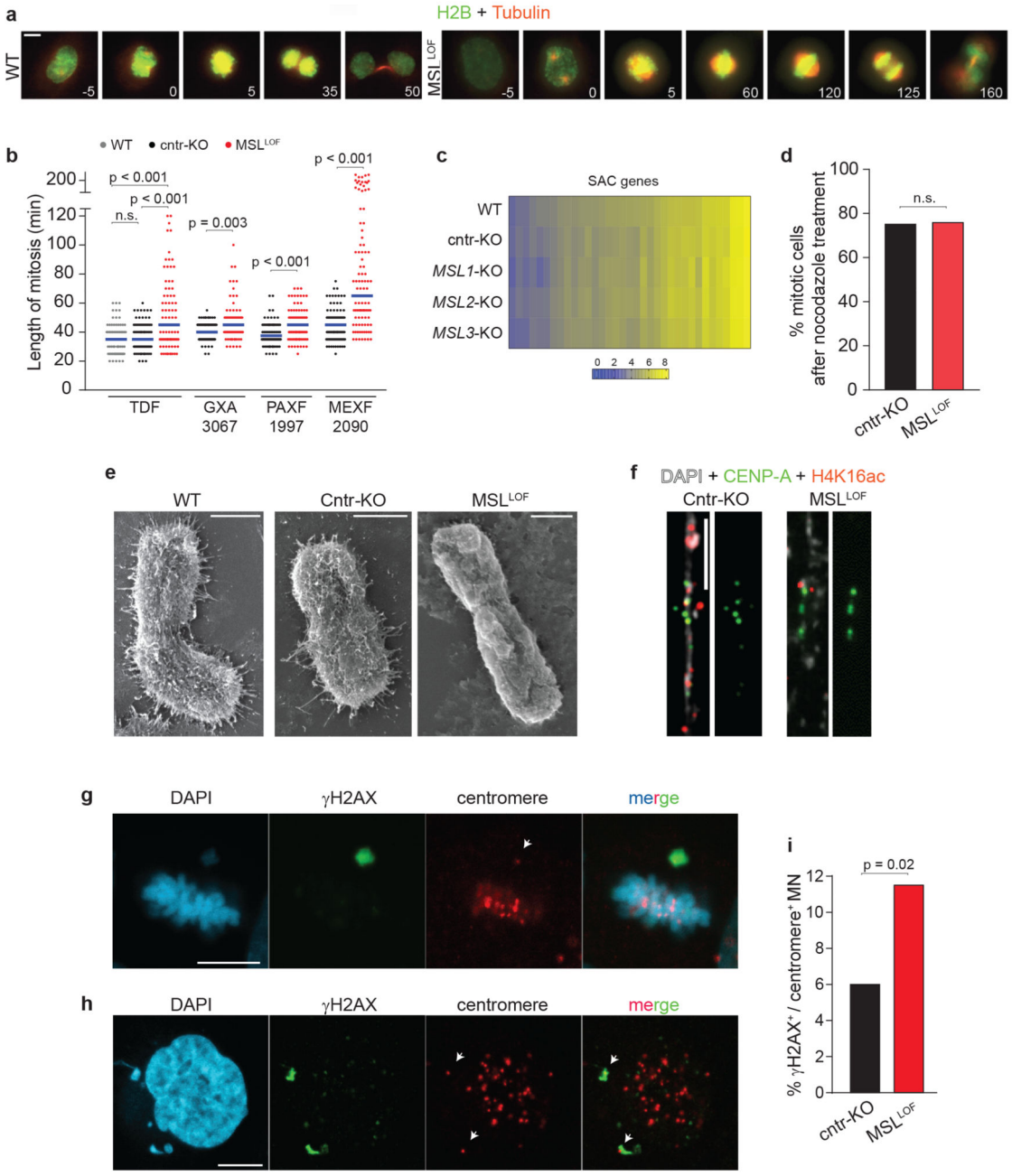


Figure 5. Chromosome mis-segregation in MSL-disrupted cells.

a. Time-lapse live microscopy of mitotic WT and MSL^{LOF} (*MSL3*-KO) TDF cells expressing H2B-mCherry (artificially colored green). Videos synchronized based on the start of chromatin condensation (T: 0 min). Microtubules are visualized with Sir-Tubulin-647 (artificially colored red). Scale bar: 10 μm. See also Supplementary videos 1-4.

b. Quantification of mitosis length. For each sample, from left to right: n = 98, 98, 98, 99, 85, 86,90, 99, 99 biologically independent mitotic events collected from two experiments.

Horizontal lines: median values. P-value from one-way ANOVA followed by Dunn's test (TDF cells) or Mann-Whitney's U-test (PDXs). N.s.: non-significant (p-value = 0.4).

c. Heatmap of expression levels of 35 SAC-related genes in MSL^{LOF} TDF cells (*MSL1*-KO, *MSL2*-KO and *MSL3*-KO) and control cells (WT and cntr-KO), as assessed by RNA-seq analysis. The average value of three biological replicates is visualized. Colors represent log_2-transformed TPM values for each gene.

d. Quantification of mitotic cells after nocodazole treatment in the indicated TDF cells. Values for *MSL1*-KO cells are shown (MSL^{LOF}). Similar results were obtained with *MSL3*-KO cells. N = 245 and 239 biologically independent cells for cntr-KO and MSL^{LOF}, respectively. P-value from two-tailed Fisher's exact test. N.s.: non-significant (p-value > 0.99).

e. Scanning electron microscopy of mitotic chromosomes from the indicated samples. Scale bar: 1 μm.

f. Immunofluorescence microscopy of chromatin fibers using the indicated antibodies showing comparable CENP-A domains at centromeres in control and MSL^{LOF} (*MSL1*-KO) cells. Scale bar: 5 μm.

g-h. Immunofluorescence microscopy of mitotic (**g**) and interphase (**h**) MSL^{LOF} (*MSL1*-KO) TDF cells using the indicated antibodies. Arrowheads: centromeres in a γH2AX-marked unaligned chromosomes (**g**) or in γH2AX-marked micronuclei (**h**). Scale bar: 10 μm.

i. Quantification of double-positive micronuclei in the indicated samples. N = 267 and 489 biologically independent micronuclei for cntr-KO and MSL^{LOF}, respectively collected from three experiments. P-value from two-tailed Fisher's exact test.

d-f, Data are representative of two independent experiments. Source data are provided.

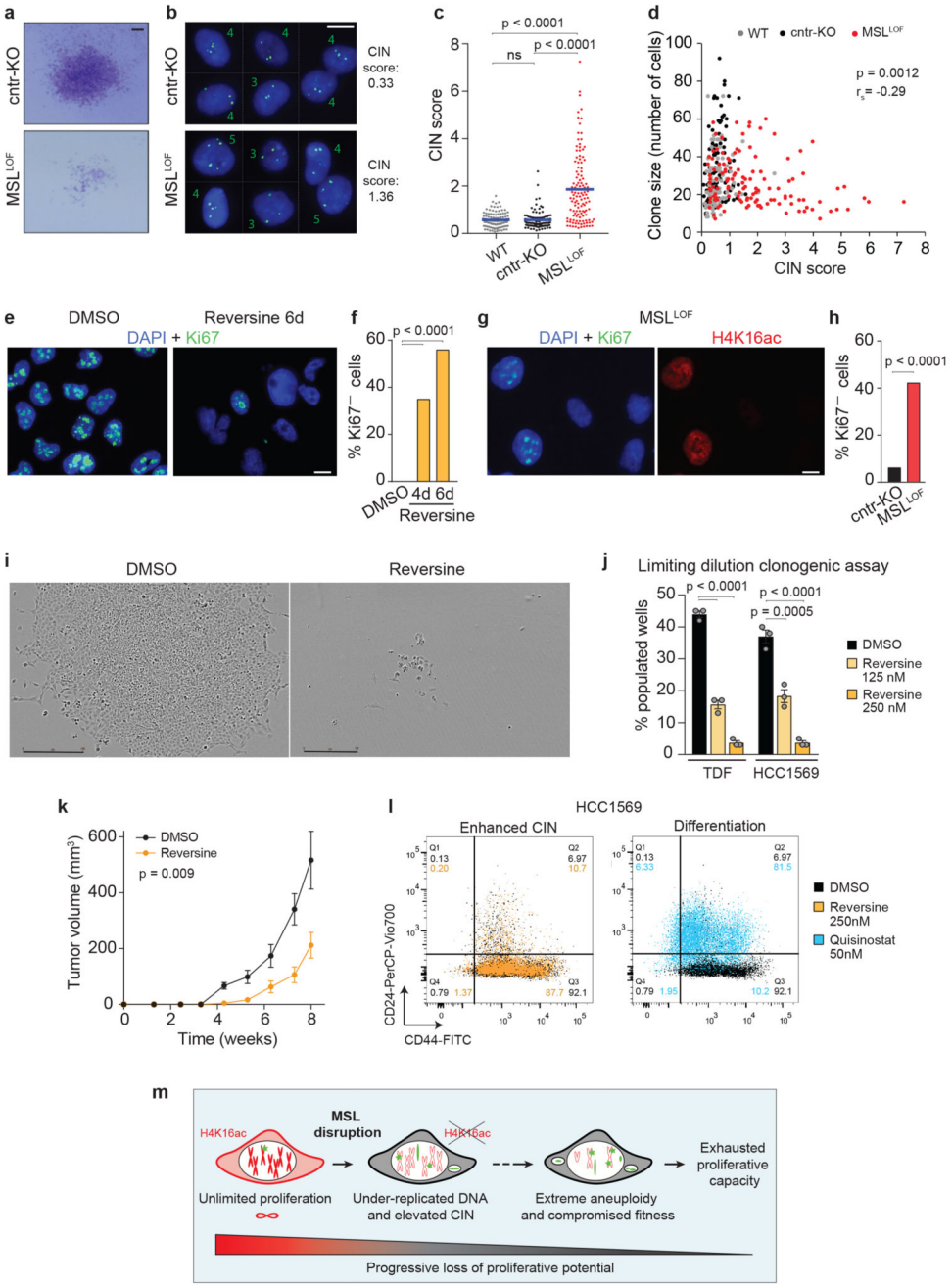


Figure 6. Enhanced chromosomal instability induced by MSL disruption exhausts cancer cell proliferative capacity.

a. Clonogenic assays comparing MSL^{LOF} (*MSL3*-KO) and control TDF cells visualized by crystal violet staining. Scale bar: 1 mm.

b. DNA FISH detecting the pericentromeric region of chromosome 12 in clonal populations grown for 14 d. Cells from distinct locations within the clones are juxtaposed. Scale bar: 10 μm.

Europe PMC Funders Author Manuscripts

c-d. Quantification of CIN score (**c**) and relationship with clone size (**d**). N = 105, 96, 121 biologically independent clones collected from two experiments. Horizontal lines: median values. P-value from one-way ANOVA followed by Tukey's test (**c**) and two-sided Spearman correlation test (MSL^{LOF} cells in **d**). N.s.: non-significant (p-value > 0.99). R_s: correlation coefficient for MSL^{LOF} cells.

e-h. Immunodetection of Ki67 (**e**, **g**) and quantification of non-proliferative Ki67-negative cells (**f**, **h**). N = 100, 102, 130, 1375, 1300 independent cells for DMSO, 4d and 6d reversine, cntr-KO and MSL^{LOF} (*MSL3*-KO) cells collected from two experiments. MEXF-2090 cells are shown in **g**-**h**; TDF and other PDX-derived cells showed similar results. P-value from two-tailed Fisher's exact test. Scale bar: 10 μm.

i-j. Limiting dilution clonogenic assays. Phase contrast images of DMSO- and reversine-treated TDF cells 12 d after plating (**i**) and quantification of populated wells (**j**). The percentage of wells containing at least one clone larger than 20 cells is plotted. Values are average ± SEM from n = 3 biologically independent samples. P-value from one-way ANOVA followed by Dunn's test. Scale bar: 400 μm.

k. Growth kinetics of tumors induced by untreated or reversine-treated TDF cells. Values are average ± SEM from n = 8 biologically independent tumors. P-value from one-tailed Student's t-test at the last time point).

l. Flow cytometry analysis of HCC1569 cells treated with reversine for 6 d to induce CIN or with quisinostat for 3 d to induce differentiation. The overlay between DMSO- and drug-treated cells is shown.

m. Model of how disruption of the MSL complex exhausts the proliferative capacity of cancer cells. Filled and empty red shapes: H4K16ac-marked and -depleted chromosomes. Green stars: focal region of under-replicated chromosomes. Green rods: severely affected chromosomes often lost through micronuclei formation.

Source data are provided.