Article

# Whole-genome doubling confers unique genetic vulnerabilities on tumour cells

https://doi.org/10.1038/s41586-020-03133-3

Received: 6 January 2020

Accepted: 17 December 2020

Published online: 27 January 2021

Check for updates

Ryan J. Quinton[1], Amanda DiDomizio[1], Marc A. Vittoria[1], Kristýna Kotýnková[1], Carlos J. Ticas[1], Sheena Patel[1], Yusuke Koga[2], Jasmine Vakhshoorzadeh[1], Nicole Hermance[3], Taruho S. Kuroda[4], Neha Parulekar[2], Alison M. Taylor[5,6,7], Amity L. Manning[3], Joshua D. Campbell[2,6] & Neil J. Ganem[1,2]✉

Whole-genome doubling (WGD) is common in human cancers, occurring early in tumorigenesis and generating genetically unstable tetraploid cells that fuel tumour development[1,2]. Cells that undergo WGD (WGD$^+$ cells) must adapt to accommodate their abnormal tetraploid state; however, the nature of these adaptations, and whether they confer vulnerabilities that can be exploited therapeutically, is unclear. Here, using sequencing data from roughly 10,000 primary human cancer samples and essentiality data from approximately 600 cancer cell lines, we show that WGD gives rise to common genetic traits that are accompanied by unique vulnerabilities. We reveal that WGD$^+$ cells are more dependent than WGD$^-$ cells on signalling from the spindle-assembly checkpoint, DNA-replication factors and proteasome function. We also identify *KIF18A*, which encodes a mitotic kinesin protein, as being specifically required for the viability of WGD$^+$ cells. Although KIF18A is largely dispensable for accurate chromosome segregation during mitosis in WGD$^-$ cells, its loss induces notable mitotic errors in WGD$^+$ cells, ultimately impairing cell viability. Collectively, our results suggest new strategies for specifically targeting WGD$^+$ cancer cells while sparing the normal, non-transformed WGD$^-$ cells that comprise human tissue.

The vast majority of human cells are diploid, and numerous cell-cycle controls exist to help ensure that this state is maintained across successive cell divisions[1]. Despite these controls, errors can occur that result in a doubling of the whole genome, whereby a natively diploid cell transitions to a tetraploid state[1]. Cells that have experienced a WGD event (hereafter WGD$^+$ cells) are oncogenic and can facilitate tumorigenesis[2]. WGD promotes tumorigenesis in at least two ways: first, proliferating WGD$^+$ cells are genomically unstable and rapidly accumulate numerical and structural chromosomal abnormalities[2]; and second, WGD$^+$ cells are better able to buffer against the negative effects of deleterious mutations and ongoing chromosome instability[3–6]. Such traits enable nascent WGD$^+$ tumour cells to proliferate in the presence of otherwise lethal genomic alterations, while simultaneously sampling increased genetic permutations, ultimately enabling phenotypic leaps that give rise to tumours[4,7]. WGD also carries important clinical implications, with recent reports showing its correlation with advanced metastatic disease and a worse overall prognosis[8].

Given the oncogenic potential associated with WGD, tumour-suppression mechanisms exist to limit the proliferation of these unstable cells. WGD$^+$ cells activate both the p53 and the Hippo tumour-suppressor pathways, and are prone to apoptosis, senescence and immune clearance[9–11]. WGD also gives rise to numerous abnormalities in cellular physiology that impair fitness[3,9,12]. To promote tumorigenesis, WGD$^+$ cells must adapt to overcome these barriers[2,9]. Thus, while WGD confers traits that favour tumorigenesis, it also imposes adaptive requirements upon cells that could give rise to unique vulnerabilities[13,14]. Identifying and exploiting these vulnerabilities represents an exciting therapeutic avenue, particularly because WGD is a broadly shared, distinguishing characteristic of many tumours[8].

## Genetic alterations enriched in WGD$^+$ tumours

To understand the genetic differences between WGD$^+$ and WGD$^-$ tumours, we first obtained calls regarding WGD status made by the ABSOLUTE algorithm for roughly 10,000 primary tumour samples, spanning 32 distinct tumour types, from The Cancer Genome Atlas (TCGA). This allowed us to separate tumour samples by whether they had (WGD$^+$) or had not (WGD$^-$) undergone a WGD event[15]. Consistent with previous estimates, we found that approximately 36% of tumours experienced at least one WGD during their evolution[8,16]. We also observed a substantial range in the occurrence of WGD between different tumour subtypes, implying that specific genetic, physiological and/or microenvironmental cues can favour or repress WGD-driven tumorigenesis (Fig. 1a).

Having differentiated WGD$^+$ and WGD$^-$ tumours, we then assessed their mutational burdens. This analysis revealed that WGD$^+$ tumours tend to have a higher total mutational burden[8]. However, we observed several tumour subtypes in which the ploidy-corrected WGD$^-$ tumours

[1]Department of Pharmacology & Experimental Therapeutics, Boston University School of Medicine, Boston, MA, USA. [2]Department of Medicine, Boston University School of Medicine, Boston, MA, USA. [3]Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA, USA. [4]Department of Molecular Pathobiology and Cell Adhesion Biology, Mie University Graduate School of Medicine, Mie, Japan. [5]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [6]Cancer Program, Broad Institute, Cambridge, MA, USA. [7]Department of Pathology and Cell Biology, Columbia University Medical Center, Member, Herbert Irving Comprehensive Cancer Center, New York, NY, USA. ✉e-mail: nganem@bu.edu

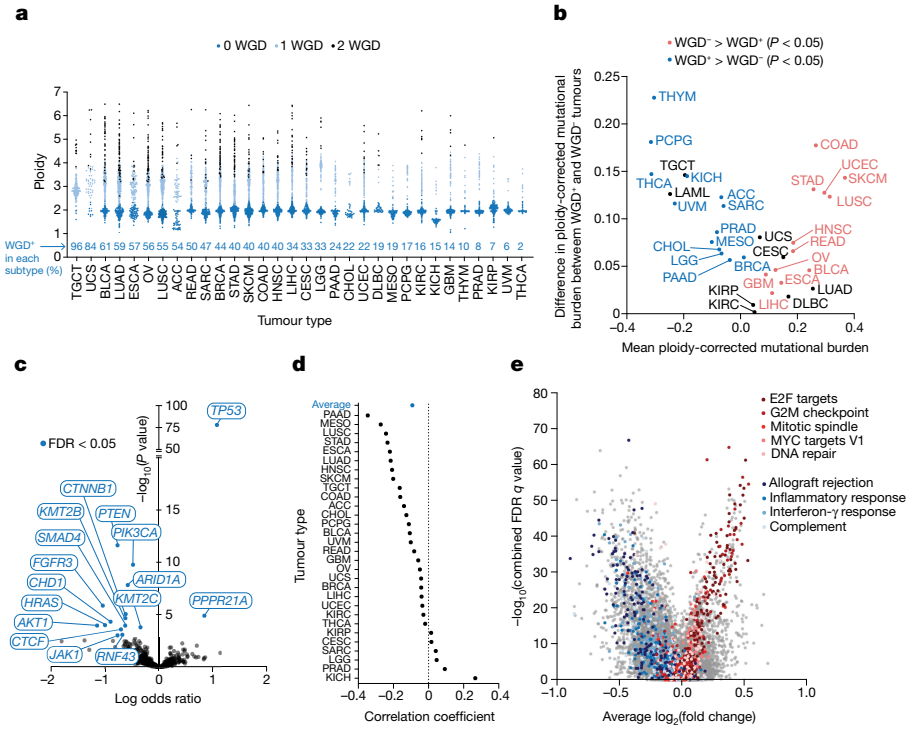



**Fig. 1 | Genetic analysis of WGD$^+$ tumours. a**, Quantification of WGD status and total ploidy of 9,700 primary human solid tumour samples from the TCGA using ABSOLUTE. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell carcinoma; ESCA, oesophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumour; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma. **b**, Mean ploidy-corrected mutational burden in the indicated tumour subtypes, plotted against the difference in the ploidy-corrected mutational burden between WGD$^+$ and WGD$^-$ tumours within each subtype (two-sided Wilcoxon rank-sum test). **c**, Enrichment of mutations in the indicated genes in WGD$^+$ tumours (log odds ratio generated by logistic regression corrected for mutation burden and tumour type). FDR, false discovery rate. **d**, Correlation of leukocyte infiltration and WGD (Pearson's correlation). **e**, Fold changes in gene expression in WGD$^+$ tumours relative to WGD$^-$ tumours, plotted against combined FDR $q$ values across all tumour types, with genes from enriched gene sets highlighted.

had a higher mutational burden than the WGD$^+$ tumours (Extended Data Fig. 1a–c). This tended to occur in subtypes with a high mutational load, characteristic of tumour types that are prone to microsatellite instability (MSI) or exposure to exogenous mutagens. Conversely, in subtypes with a lower mutational burden, it was the WGD$^+$ tumours that had the higher ploidy-corrected mutational burden (Fig. 1b). This supports a recent report[6] predicting that highly mutated tumours, which experience fewer somatic copy-number alterations (SCNAs), encounter selection pressures that disfavour WGD, while tumour types with a lower mutational burden and more SCNAs will favour WGD owing to its capacity to buffer against deleterious mutations in genomic regions of loss of heterozygosity. We also observed that tumours with MSI or mutations in the gene for DNA polymerase ε (*POLE*), which have a very high mutational burden, tend not to experience WGD events, as has been shown in other cohorts[6,8] (Extended Data Fig. 1d).

We next explored the mutational landscape of WGD$^+$ tumours, where we observed a substantial enrichment of mutations in *TP53* and *PPPR21A* (Fig. 1c), consistent with findings from individuals with advanced cancer[8,16]. The positive selection for these mutations is clear: p53 represents a major barrier to the proliferation of WGD$^+$ cells, and so inactivating mutations in *TP53* are favoured in WGD$^+$ cancers. Mutations in *PPP2R1A* promote centrosome clustering, an important adaptation for preventing multipolar cell division and cell death in WGD$^+$ cells with supernumerary centrosomes[3,17]. We also identified mutations that are negatively enriched in WGD$^+$ tumours, implying that these mutations are either less important for, or perhaps incompatible with, driving tumorigenesis in the context of WGD (Fig. 1c). Many of these negatively enriched genes are known to contain microsatellite indels, providing further evidence for the diverging pressures that separate the evolutionary trajectory of MSI and WGD$^+$ tumours[18].

To assess changes in the microenvironment of WGD tumours, we obtained ABSOLUTE purity values (that is, the fraction of non-tumour cells) for TCGA tumour samples[15]. We found that WGD correlates with decreased purity and increased non-immune stromal infiltration (Extended Data Fig. 2a, b). We also assessed the correlation of WGD with TCGA estimates of tumour-infiltrating leukocytes (TILs), and found a negative correlation between WGD and TILs (Fig. 1d). Analysis of gene expression revealed that the most negatively enriched gene sets in WGD$^+$ tumours were inflammatory processes, corroborating that these tumours present with diminished host immune responses, similar to highly aneuploid tumours (Fig. 1e)[19,20]. We further identified that WGD$^+$ tumours overexpress genes important for cellular proliferation, mitotic spindle formation and DNA repair (Fig. 1e, Supplementary Table 1). Genomic characteristics such as mutational burden, purity and TILs interact in poorly understood ways to contribute to the response of patients to immunotherapy, and we hypothesized that WGD may be a contributing factor in this response. Indeed, individuals with WGD$^+$ tumours respond better to drugs that block the immune checkpoint

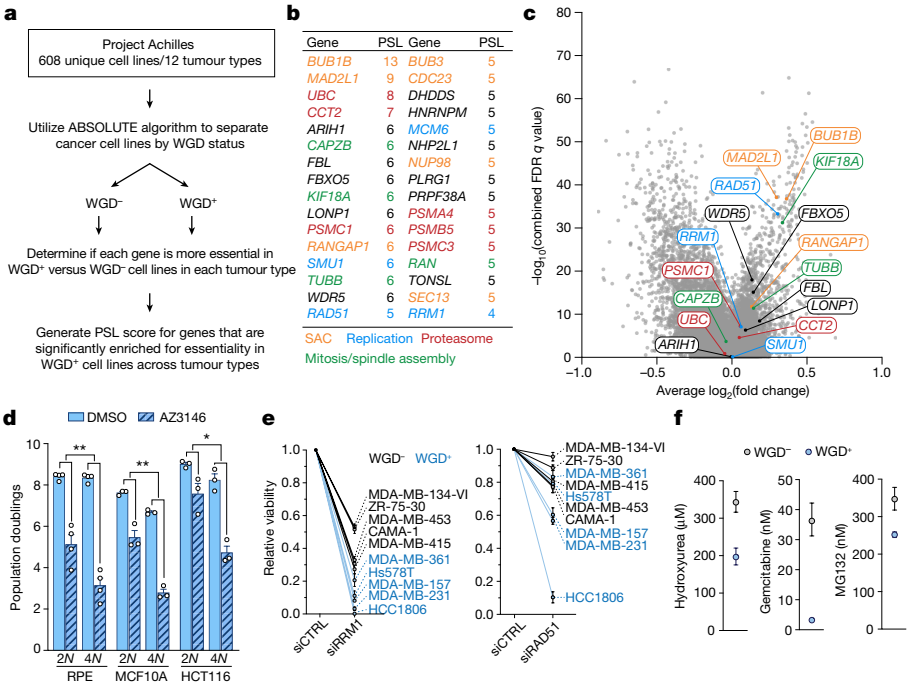



**Fig. 2 | Identification and validation of PSL genes. a**, Workflow used to identify gene essentiality in WGD$^+$ cancer cells from Project Achilles data (see Methods). **b**, Top hits from PSL analysis (text colour indicates genes associated with the indicated pathways). **c**, Fold changes in gene expression in WGD$^+$ tumours relative to WGD$^-$ tumours, plotted against combined FDR *q* values across all tumour types, with select PSL genes highlighted. **d**, Population doublings of the indicated tumour cell types after 8 days of treatment with the small molecule AZ3146 ($n = 3$ independent experiments; two-way analysis of variance (ANOVA) with interaction; graph shows means ± s.e.m.; interaction *P* values = 0.0085, 0.0020, 0.0156, from left to right). **e**, Relative viability of the indicated cell lines 7 days after treatment with the indicated short interfering RNAs (siRNAs) ($n = 3$ independent experiments; graph shows means ± s.e.m.; CTRL, control). **f**, Mean 50% lethal concentration ($LC_{50}$) values of the indicated drug treatments for 5 WGD$^-$ and 5 WGD$^+$ breast-cancer cell lines ($n = 3$ independent experiments; nonlinear regression; graphs show mean $LC_{50}$ ± 95% confidence interval). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

inhibitor PD1, and this applies across numerous tumour subtypes (Extended Data Fig. 2d). Collectively, our data demonstrate key genetic and phenotypic differences between WGD$^+$ and WGD$^-$ tumours, and hint at potential adaptations and vulnerabilities that may inexorably arise following a WGD.

## WGD confers unique genetic vulnerabilities

We examined whether WGD confers unique genetic dependencies on tumour cells by obtaining ABSOLUTE WGD calls on cancer cell lines from Project Achilles. This is a comprehensive catalogue that quantifies the essentiality of approximately 20,000 genes across roughly 600 cell lines following gene depletion through both CRISPR and RNA interference (RNAi) (Supplementary Table 2)[21]. We used Project Achilles scores to identify genes that are enriched for essentiality in WGD$^+$ cell lines relative to WGD$^-$ cell lines (so-called ploidy-specific lethal (PSL) genes[14]) (Fig. 2a, b, Extended Data Fig. 2c, Supplementary Tables 3, 4). We mapped these PSL genes against the gene-expression signature of WGD$^+$ samples in the TCGA, and found several PSL genes to be greatly overexpressed, reinforcing their importance in the progression of WGD$^+$ tumours (Fig. 2c).

To validate these PSL genes, we first generated three isogenically matched diploid (WGD$^-$ or 2*N*) and tetraploid (WGD$^+$ or 4*N*) cell lines as previously described[3]. These lines included the non-transformed epithelial cell lines RPE-1 and MCF10A, as well as the colon-cancer cell line HCT116. Growth of these tetraploid cell lines over multiple passages selects for genomically stable variants that have lost their extra centrosomes (Extended Data Fig. 2e–h)[3,22]. The development of these isogenic lines enabled us to directly compare cellular dependencies in cells that differ only by WGD status.

We first validated *BUB1B* and *MAD2L1*, the two strongest PSL gene hits from our analysis. These genes encode proteins that are essential to the function of the spindle-assembly checkpoint (SAC), which delays the onset of anaphase until all chromosomes have attached to the mitotic spindle, promoting the accurate partitioning of genomic content during mitosis[23]. Increasing the number of chromosomes prolongs the time needed to achieve full chromosome attachment and alignment[24], suggesting that premature anaphase induced by disruption of the SAC should give rise to chromosome-segregation errors at elevated rates in tetraploid cells. Using live-cell imaging, we found that tetraploid cells indeed require more time to attach and align chromosomes relative to diploid cells in all three cell lines tested (Extended Data Fig. 3a). Consequently, we found that inhibiting the SAC using the small molecule AZ3146, which inhibits the MPS1 kinase and abrogates the SAC in a manner similar to depletion of MAD2 or BUBR1, leads to a substantial increase in chromosome-segregation defects and micronuclei formation in tetraploid cells relative to diploids (Extended Data Fig. 3b). The presence of micronuclei and errors in chromosome segregation impair cell fitness; concordantly, assays of population doubling confirmed that tetraploid cells are much more sensitive to SAC inhibition than diploids (Fig. 2d). These data corroborate previous work[25] and validate our methodology for analysing PSL genes.

The identification of several genes involved in DNA replication as PSL hits suggests that WGD$^+$ cells may also be more vulnerable to challenges to DNA replication than WGD$^-$ cells. We first validated that reductions in the levels of *RRM1* and *RAD51* (two PSL genes known to mitigate the DNA damage associated with replication stress) preferentially impair the viability of tetraploid cells (Extended Data Fig. 3c–f). As an orthogonal approach, we also treated isogenic diploid and tetraploid cells with

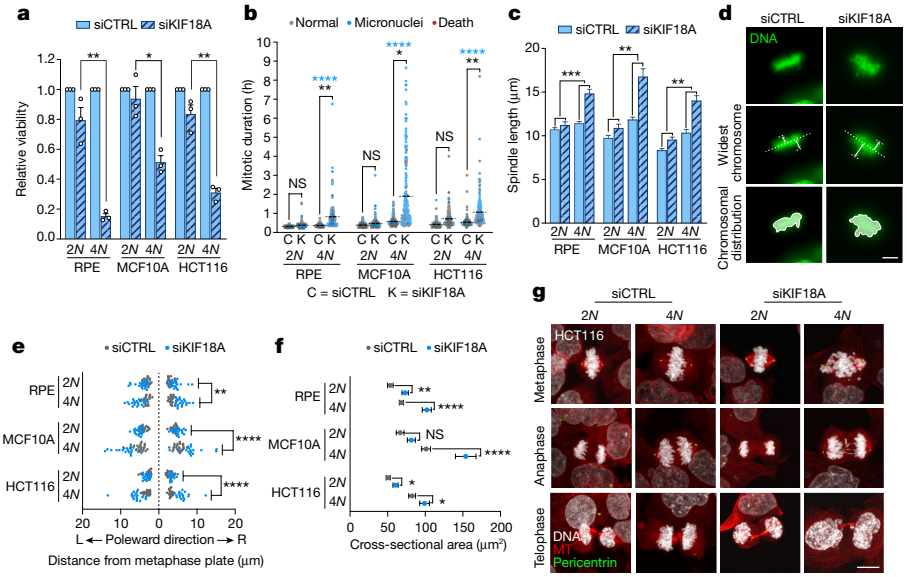



**Fig. 3 | Depletion of KIF18A impairs the mitotic fidelity of WGD$^+$ cells.** **a**, Relative viability of the indicated cell lines 8 days after transfection with the indicated siRNAs ($n = 3$ independent experiments; each condition normalized to respective control; two-sided Student's unpaired $t$-test; graph shows means ± s.e.m.; $P$ values = 0.0019, 0.0109, 0.0017, from left to right; CTRL, control). **b**, Duration of mitosis and cell fate after treatment with the indicated siRNAs ($n = 200$ cells per condition; black asterisks indicate $P$ values from two-sided Student's $t$-tests comparing mean mitotic duration; blue asterisks indicate $P$ values from two-sided Fisher's exact tests comparing the fraction of mitoses that give rise to micronuclei; dotted lines represent mean mitotic duration; NS, not significant). **c**, Measurement of spindle length (centrosome-to-centrosome) after transfection with the indicated siRNAs ($n = 20$ cells per condition; two-way ANOVA with interaction; graph shows means ± s.e.m.; scale bar, 10 μm; interaction $P$ values = 0.0001, 0.0011, 0.0032, from left to right). **d**, Images showing the measurement of chromosome oscillations immediately before anaphase by assessing the widest chromosome oscillations in each poleward direction and the cross-sectional area of all the chromosomes. Scale bar, 10 μm. **e**, Widest chromosome oscillations in each poleward direction immediately before anaphase ($n = 20$ cells per condition from 2 independent experiments; two-way ANOVA with interaction; interaction $P = 0.0025$, $< 0.0001$, $< 0.0001$, from top to bottom). **f**, Two-dimensional cross-sectional area of the entire body of chromosomes immediately before anaphase ($n = 20$ cells per condition; two-sided Student's unpaired $t$-test; graph shows means ± s.e.m.; $P$ values = 0.0012, $< 0.0001$, 0.0525, $< 0.0001$, 0.0318, 0.0432, from top to bottom). **g**, Representative confocal images showing phases of mitosis in the indicated cell lines 48 h after transfection with the indicated siRNAs. Scale bar, 10 μm. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

hydroxyurea or gemcitabine, which inhibit the activity of ribonucleotide reductase (RRM1) and induce replication stress. We observed that tetraploid cell lines show an increased sensitivity to these inhibitors (Extended Data Fig. 4a, b). We also confirmed this result in a panel of ten breast-cancer cell lines (five WGD$^+$ and five WGD$^-$) (Fig. 2e, f, Extended Data Fig. 4c, d, f, g). These data reveal that WGD$^+$ tumour cells are more dependent than WGD$^-$ tumour cells on specific DNA-replication factors, perhaps in compensation for the increased replication stress induced by tetraploidy[26,27].

We also identified several PSL genes that encode regulators of the proteasome, suggesting that WGD confers vulnerability to disruptions in protein turnover. Indeed, we found that WGD$^+$ cells are more sensitive than WGD$^-$ cells to the proteasome inhibitor MG132 (Fig. 2f, Extended Data Figs. 4e, 5a). This vulnerability may be due to the highly aneuploid nature of WGD$^+$ cells, as aneuploidy has been shown to induce proteotoxic stress[28]. Supporting this view, we found that tetraploid RPE-1 cells, which maintain a euploid number of chromosomes (92) (Extended Data Fig. 2h), were the only tetraploid cell line that was not more sensitive to MG132 relative to diploids (Extended Data Fig. 5a).

## WGD confers dependence on KIF18A

Our analysis identified the gene *KIF18A*, which encodes a mitotic kinesin protein, as an important PSL hit (Fig. 2b). KIF18A functions to suppress chromosomal oscillations at the metaphase plate by regulating microtubule dynamics to facilitate proper alignment and distribution of chromosomes during mitosis[29–32]. Notably, unlike the aforementioned genes which regulate essential cellular processes such as SAC function, DNA replication and proteasome activity, *KIF18A* is a non-essential gene in normal diploid cells, as attested by the fact that transgenic *KIF18A*-knockout mice survive to adulthood[33]. Further, *KIF18A* is commonly overexpressed in WGD$^+$ tumours (Fig. 2c), and its overexpression correlates with a worse prognosis for individuals with cancer (Extended Data Fig. 5b).

We validated *KIF18A* as a PSL gene by confirming that depletion of KIF18A greatly impairs the viability of tetraploid but not diploid cells (Fig. 3a, Extended Data Fig. 5c). To understand the mechanism underlying this reduction in viability, we used live-cell imaging to monitor mitotic progression following KIF18A depletion. We observed that depletion of KIF18A had no effect on mitotic duration in diploid cells, but led to significantly prolonged mitoses in tetraploid cells (Fig. 3b). We also observed that although diploid cells lacking KIF18A exhibited subtle defects in chromosome alignment at anaphase onset, chromosome segregation proceeded relatively normally with no substantial increase in the formation of micronuclei (Fig. 3b, g). By contrast, tetraploid cells depleted of KIF18A exhibited marked increases in chromosome misalignment, lagging chromosomes in anaphase, and micronuclei formation (Fig. 3b, g, Extended Data Fig. 5d–f, Supplementary Videos 1–4). We also observed that micronuclei in tetraploid cells depleted of KIF18A were more prone to rupture of the nuclear envelope than were micronuclei in diploid cells, thus exposing the chromosomal contents within the micronuclei to the cytosolic environment and inducing both DNA damage and activation of the cGAS–STING pathway (Extended Data Fig. 6a)[34–37].

We speculated that the mitotic delays and defects in chromosome segregation observed following KIF18A loss may be induced by changes in spindle morphology in tetraploid cells. To accommodate their doubled chromosome content, tetraploid cells assemble larger mitotic

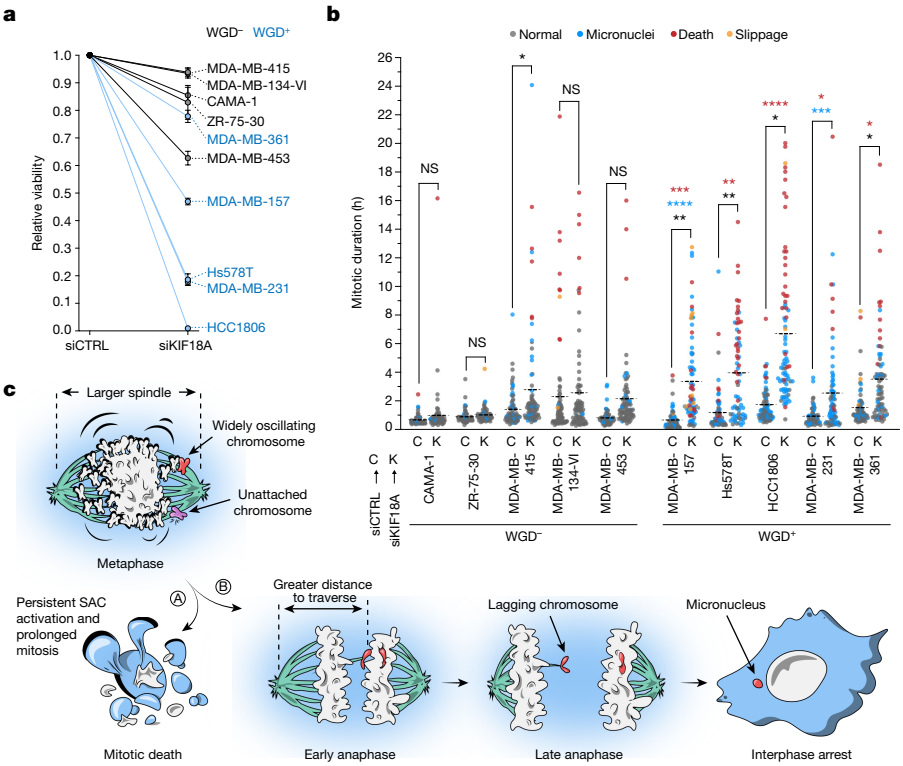



**Fig. 4 | WGD confers dependence on KIF18A in a panel of breast-cancer cell lines. a**, Relative viability of the indicated cell lines 8 days after transfection with the indicated siRNAs ($n = 3$ independent experiments; graph shows means ± s.e.m.). **b**, Mitotic duration and fate following transfection with the indicated siRNAs ($n = 80$ cells per condition across 2 independent experiments; dotted lines represent mean mitotic duration; black asterisks indicate *P* values from two-sided Student's unpaired *t*-tests comparing mean mitotic duration; blue asterisks indicate *P* values from two-sided Fisher's exact tests comparing the fraction of mitoses that give rise to micronuclei; red asterisks indicate *P* values from two-sided Fisher's exact tests comparing the fraction of cells that die in mitosis). **c**, Depletion of KIF18A impairs the viability of WGD$^+$ cells through two distinct mechanisms. A, Widely oscillating chromosomes fail to properly attach to microtubules, thus activating the SAC and leading to prolonged mitosis and death. B, Larger spindles and wider oscillations increase the distance that some chromosomes must traverse in anaphase, leading to lagging chromosomes, formation of micronuclei and cellular arrest. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

spindles (Fig. 3c)[3]. Depletion of KIF18A led to an additional increase in spindle size, and this effect was substantially more notable in tetraploid cells than in diploids (Fig. 3c).

We also measured the magnitude of chromosome oscillations immediately before the onset of anaphase in diploid and tetraploid cells by assessing the chromosomes with the widest oscillation in each poleward direction, as well as the overall efficiency of chromosome alignment by measuring the total two-dimensional area occupied by the entire body of chromosomes (Fig. 3d). These analyses revealed that the magnitude of chromosomal oscillations is much greater in tetraploid cells relative to diploid cells following KIF18A depletion (Fig. 3e, f). One consequence of the presence of hyperoscillating chromosomes in tetraploid cells depleted of KIF18A is that they have a propensity to lose their attachment to the mitotic spindle and to activate the SAC, thus explaining the mitotic delays we observed (Extended Data Fig. 6b)[38,39]. A second consequence is that severely misaligned chromosomes must traverse a substantially greater distance during anaphase in tetraploid cells than in diploid cells, thus explaining the observed increase in lagging chromosomes and micronuclei.

We used long-term live-cell imaging to track the fates of isogenic diploid and tetraploid cells depleted of KIF18A. Our analysis revealed that although most such diploid cells undergo normal cell-cycle progression, isogenic tetraploid cells are prone to arresting in interphase following abnormal mitosis, concomitant with activation of the p53 pathway (Extended Data Fig. 6d, e). Thus, our data reveal that loss of KIF18A in WGD$^+$ cells predisposes cells to lagging chromosomes, micronuclei formation, micronuclei rupture and proliferative arrest. In support this mechanism, we found that cellular proliferation is required for the loss of KIF18A to drive our observed viability defects (Extended Data Fig. 6f).

We sought to further validate the ploidy-specific lethality of KIF18A across a panel of breast-cancer cell lines. Supporting our pan-cancer analysis of gene expression (Fig. 2c), we found that KIF18A protein levels are typically elevated in WGD$^+$ cells (Extended Data Fig. 7a). Knockdown of KIF18A from all ten cell lines (Extended Data Fig. 7b) confirmed that WGD$^+$ lines experience a substantially greater reduction in viability relative to WGD$^-$ cell lines (Fig. 4a, Extended Data Fig. 7c–e, Supplementary Videos 5, 6). Live-cell imaging revealed that WGD$^+$ breast-cancer cells exhibited increased spindle lengths and chromosome hyperoscillations relative to WGD$^-$ cells after loss of KIF18A (Extended Data Figs. 7f, g, 8a), thus promoting chromosome detachment, SAC activation and prolonged mitosis (Fig. 4b and Extended Data Fig. 6c). Notably, we observed that a large fraction of WGD$^+$ cells were never able to satisfy the SAC, and exhibited a markedly prolonged mitotic arrest before ultimately undergoing mitotic cell death (Fig. 4b). WGD$^+$ cells depleted of KIF18A that achieved anaphase exhibited substantial increases in both anaphase lagging chromosomes and micronuclei relative to the WGD$^-$ cell lines, similar to what was observed in the isogenic tetraploid models (Fig. 4c, Extended Data Fig. 8b). However, in contrast to the p53-proficient isogenic tetraploid cells, WGD$^+$ breast-cancer cell lines depleted of KIF18A were not prone to cell-cycle arrest following abnormal mitosis (Extended Data Fig. 8c). Instead, a fraction of these cells died in interphase after experiencing catastrophic mitoses resulting in micronuclei formation, while most initiated a second round of

mitosis without KIF18A, where they were just as, or even more, prone to mitotic cell death (Extended Data Fig. 8d).

WGD is a discrete event with the capacity to fix a cell along an evolutionary path towards tumorigenesis. It gives rise to supernumerary centrosomes and a markedly increased chromosomal burden, and the resultant chromosomal instability leads to substantial deviations away from euploidy towards highly aneuploid states. Indeed, WGD$^+$ cells comprise the vast majority of the highly aneuploid cells observed in human cancers (Fig. 1a, Extended Data Fig. 8e, f, Supplementary Table 5). Our data reveal that loss of KIF18A impairs mitotic fidelity and cell viability in WGD$^+$ cancer cells.

Although many factors may promote KIF18A dependency, our data show that the enhanced reliance of WGD$^+$ cells on KIF18A is predominantly due to the added chromosomal burden that inexorably follows a WGD. Two key observations support this view. First, we observe viability defects in euploid tetraploid RPE-1 cells depleted of KIF18A, despite the fact that these cells possess a normal number of centrosomes and are chromosomally stable (Fig. 3a)[3,9]. Second, we find that the presence of only one or two additional chromosomes in otherwise diploid cells is not sufficient to impart sensitivity to loss of KIF18A (Extended Data Fig. 8g). This suggests that aneuploidy per se is insufficient to drive the dependency on KIF18A that we observe; instead, notable increases in chromosome number or the pronounced chromosomal instability and high levels of aneuploidy that arise predominantly through WGD are required (Extended Data Fig. 8e, f). Our results are consistent with recent studies that have demonstrated dependency on KIF18A in highly aneuploid and chromosomally unstable cancer cell lines[40,41]. Together, our data highlight KIF18A as an attractive therapeutic target whose inhibition may enable the specific targeting of WGD$^+$ tumours while sparing the normal diploid cells that comprise human tissue. Supporting this view, it has been shown that mice lacking *KIF18A* are protected from tumorigenesis, and that depletion of KIF18A from the WGD$^+$ breast-cancer cell line MDA-MB-231 impairs tumour growth in vivo[42,43].

Herein, we have comprehensively catalogued genomic characteristics unique to WGD$^+$ tumours, and generated a list of PSL genes that highlight vulnerabilities that can arise with a WGD event. Notably, many of these dependencies are conserved in budding yeast, highlighting the broad relevance of this phenomenon[13]. This work serves to underscore the importance and untapped potential of exploring and targeting WGD in human cancers.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-03133-3.


1. Lens, S. M. A. & Medema, R. H. Cytokinesis defects and cancer. *Nat. Rev. Cancer* **19**, 32–45 (2019).
2. Fujiwara, T. et al. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043–1047 (2005).
3. Ganem, N. J., Godinho, S. A. & Pellman, D. A mechanism linking extra centrosomes to chromosomal instability. *Nature* **460**, 278–282 (2009).
4. Thompson, D. A., Desai, M. M. & Murray, A. W. Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. *Curr. Biol.* **16**, 1581–1590 (2006).
5. Dewhurst, S. M. et al. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discov.* **4**, 175–185 (2014).
6. López, S. et al. Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. *Nat. Genet.* **52**, 283–293 (2020).
7. Selmecki, A. M. et al. Polyploidy can drive rapid adaptation in yeast. *Nature* **519**, 349–352 (2015).
8. Bielski, C. M. et al. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat. Genet.* **50**, 1189–1195 (2018).
9. Ganem, N. J. et al. Cytokinesis failure triggers hippo tumor suppressor pathway activation. *Cell* **158**, 833–848 (2014).
10. Senovilla, L. et al. An immunosurveillance mechanism controls cancer cell ploidy. *Science* **337**, 1678–1684 (2012).
11. Andreassen, P. R., Lohez, O. D., Lacroix, F. B. & Margolis, R. L. Tetraploid state induces p53-dependent arrest of nontransformed mammalian cells in G1. *Mol. Biol. Cell* **12**, 1315–1328 (2001).
12. Ben-David, U. & Amon, A. Context is everything: aneuploidy in cancer. *Nat. Rev. Genet.* **21**, 44–62 (2019). 10
13. Storchová, Z. et al. Genome-wide genetic analysis of polyploidy in yeast. *Nature* **443**, 541–547 (2006).
14. Lin, H. et al. Polyploids require Bik1 for kinetochore-microtubule attachment. *J. Cell Biol.* **155**, 1173–1184 (2001).
15. Carter, S. L. et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **30**, 413–421 (2012).
16. Zack, T. I. et al. Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013).
17. Antao, N. V., Marcet-Ortega, M., Cifani, P., Kentsis, A. & Foley, E. A. A cancer-associated missense mutation in PP2A-Aα increases centrosome clustering during mitosis. *iScience* **19**, 74–82 (2019).
18. Maruvka, Y. E. et al. Analysis of somatic microsatellite indels identifies driver events in human tumors. *Nat. Biotechnol.* **35**, 951–959 (2017).
19. Taylor, A. M. et al. Genomic and functional approaches to understanding cancer aneuploidy. *Cancer Cell* **33**, 676–689 (2018).
20. Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017).
21. Broad, D. *DepMap Achilles 18Q3 Public* https://depmap.org/portal/ (2018).
22. Godinho, S. A. et al. Oncogene-like induction of cellular invasion from centrosome amplification. *Nature* **510**, 167–171 (2014).
23. Musacchio, A. The molecular biology of spindle assembly checkpoint signaling dynamics. *Curr. Biol.* **25**, R1002–R1018 (2015).
24. Yang, Z., Lončarek, J., Khodjakov, A. & Rieder, C. L. Extra centrosomes and/or chromosomes prolong mitosis in human cells. *Nat. Cell Biol.* **10**, 748–751 (2008).
25. Jemaà, M. et al. Whole-genome duplication increases tumor cell sensitivity to MPS1 inhibition. *Oncotarget* **7**, 885–901 (2016).
26. Wangsa, D. et al. Near-tetraploid cancer cells show chromosome instability triggered by replication stress and exhibit enhanced invasiveness. *FASEB J.* **32**, 3502–3517 (2018).
27. Zheng, L. et al. Polyploid cells rewire DNA damage response networks to overcome replication stress-induced barriers for tumour progression. *Nat. Commun.* **3**, 815 (2012).
28. Santaguida, S. & Amon, A. Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nat. Rev. Mol. Cell Biol.* **16**, 473–485 (2015).
29. Stumpff, J., von Dassow, G., Wagenbach, M., Asbury, C. & Wordeman, L. The kinesin-8 motor Kif18A suppresses kinetochore movements to control mitotic chromosome alignment. *Dev. Cell* **14**, 252–262 (2008).
30. Stumpff, J., Wagenbach, M., Franck, A., Asbury, C. L. & Wordeman, L. Kif18A and chromokinesins confine centromere movements via microtubule growth suppression and spatial control of kinetochore tension. *Dev. Cell* **22**, 1017–1029 (2012).
31. Fonseca, C. L. et al. Mitotic chromosome alignment ensures mitotic fidelity by promoting interchromosomal compaction during anaphase. *J. Cell Biol.* **218**, 1148–1163 (2019).
32. Mayr, M. I. et al. The human kinesin Kif18A is a motile microtubule depolymerase essential for chromosome congression. *Curr. Biol.* **17**, 488–498 (2007).
33. Czechanski, A. et al. Kif18a is specifically required for mitotic progression during germ line development. *Dev. Biol.* **402**, 253–262 (2015).
34. Hatch, E. M., Fischer, A. H., Deerinck, T. J. & Hetzer, M. W. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell* **154**, 47–60 (2013).
35. Mackenzie, K. J. et al. cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**, 461–465 (2017).
36. Harding, S. M. et al. Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature* **548**, 466–470 (2017).
37. Zhang, C.-Z. et al. Chromothripsis from DNA damage in micronuclei. *Nature* **522**, 179–184 (2015).
38. Janssen, L. M. E. et al. Loss of Kif18A results in spindle assembly checkpoint activation at microtubule-attached kinetochores. *Curr. Biol.* **28**, 2685–2696 (2018).
39. Edzuka, T. & Goshima, G. *Drosophila* kinesin-8 stabilizes the kinetochore-microtubule interaction. *J. Cell Biol.* **218**, 474–488 (2019).
40. Cohen-Sharir, Y. et al. Aneuploidy renders cancer cells vulnerable to mitotic checkpoint inhibition. *Nature* https://doi.org/10.1038/s41586-020-03114-6 (2021).
41. Marquis, C. et al. Chromosomally unstable tumor cells specifically require KIF18A for proliferation. *Nat. Comm.* (in the press).
42. Zhu, H. et al. Targeted deletion of Kif18a protects from colitis-associated colorectal (CAC) tumors in mice through impairing Akt phosphorylation. *Biochem. Biophys. Res. Commun.* **438**, 97–102 (2013).
43. Zhang, C. et al. Kif18A is involved in human breast carcinogenesis. *Carcinogenesis* **31**, 1676–1684 (2010).

# Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### WGD, purity and ploidy calls

TCGA samples were analysed previously using the ABSOLUTE algorithm[15]. ABSOLUTE takes data on gene copy numbers and mutation to estimate sample purity, ploidy and number of whole-genome doublings. ABSOLUTE calls for TCGA samples are available in ref. [19]. Briefly, the algorithm infers from sequencing data what fraction of a tumour sample is composed of tumour cells versus non-tumour cells (purity), as well as the ploidy of a tumour sample from analysis of copy number ratios across the entire genome. WGD status is inferred from the ploidy distribution within a tumour type, from information on homologous copy numbers across the genome, and from the presence of duplicated mutations.

### Ploidy-corrected mutational burden

To compare the ploidy-corrected mutational burden of $WGD^+$ and $WGD^-$ TCGA samples, we divided the non-synonymous mutations per megabase ($log_{10}$-transformed)[44] of each sample by their ploidy as defined by ABSOLUTE. We performed a linear regression using the lm function in R version 3.2.3. The formula was: mutational burden as a function of WGD + tumour type + MSI/*POLE* status. We applied a Wilcoxon rank-sum test to analyse the total mutational burden and ploidy-corrected mutational burden between $WGD^+$ and $WGD^-$ samples within each subtype.

### Mutations in $WGD^+$ tumours

To identify the frequencies of gene mutations associated with WGD status, we applied logistic regression to 631 genes that were found to be driver genes in one or more tumour types by MutSig2CV (ref. [45]). The formula for the logistic regression model was: mutation status as a function of WGD + mutation burden + tumour type, where 'mutation burden' was the number of non-synonymous mutations per megabase (log10-transformed)[1] and WGD status was defined by ABSOLUTE calls retrieved from http://api.gdc.cancer.gov/data/4f277128-f793-4354-a13d-30cc7fe9f6b5. The maf file from the TCGA PanCanAtlas MC3 project[46] was used to derive the mutation status for each gene in each tumour retrieved from https://api.gdc.cancer.gov/data/1c8cfe5f-e52d-41ba-94da-f15ea1337efc. This file was filtered to include only those variants with 'PASS', 'wga' or 'native_wga_mix' in the 'FILTER' column. Variants with 'Frame_Shift_Del', 'Frame_Shift_Ins', 'In_Frame_Del', 'In_Frame_Ins', 'Missense_Mutation', 'Nonsense_Mutation', 'Nonstop_Mutation', 'Translation_Start_Site', 'Splice_Site', 'De_novo_Start_InFrame', 'De_novo_Start_OutOfFrame', 'Stop_Codon_Del', 'Stop_Codon_Ins', 'Start_Codon_Del' or 'Start_Codon_Ins' in the 'Variant_Classification' column were considered non-synonymous. An FDR correction was applied to the *P* values for the WGD term to control for multiple hypothesis testing.

### Leukocyte infiltrate and stromal calls

Estimates of leukocyte fraction in the TCGA samples were generated using a mixture model of DNA methylation in pure leukocytes versus normal tissue. More details and all calls can be found in ref. [47]. Stromal calls were made by subtracting the leukocyte fraction from ABSOLUTE purity estimates described above. Pearson's correlation coefficients were calculated after removing MSI/*POLE* mutant samples from the data set and comparing TIL/stromal/purity scores against the binary calls of WGD status for each analysis.

### Immunotherapy response

Responses of individuals with cancer to immunotherapy were determined according to response and evaluation criteria in solid tumours (RECIST), where complete and partial responders were considered as responders and those with progressive disease were considered non-responders. For each tumour subtype, the fraction of responders and non-responders was determined by ABSOLUTE for people with $WGD^-$ and $WGD^+$ tumours. For the composite analysis, all tumour types were combined and the difference in responders was assessed using a Fisher's exact test[48,49].

### Analysis of gene expression

Data on gene expression and copy numbers from TCGA samples were obtained from the PanCanAtlas project (https://gdc.cancer.gov/about-data/publications/pancanatlas). RNA-seqV2 data were used for expression analysis (http://api.gdc.cancer.gov/data/3586c0da-64d0-4b74-a449-5ff4d9136611). Expression values were $log_2$-transformed after adding a pseudo-count of 1. Copy number ratios were obtained for each gene by running GISTIC2.0 on the PanCan segmentation file (http://api.gdc.cancer.gov/data/00a32f7a-c85f-4f86-850d-be53973cbc4d). Analysis was limited to primary tumours across all cancer types. *P* values for WGD were corrected for multiple hypothesis testing with the Benjamini–Hochberg false discovery rate (FDR).

To identify gene-expression profiles associated with WGD status, we applied the following linear model to each gene within each tumour type: expression as a function of WGD + purity + CN_Local, where 'purity' is the ABSOLUTE-estimated purity for each tumour, and 'CN_Local' is the $log_2$-transformed copy number ratio for that gene in each tumour estimated by GISTIC2.0 (ref. [50]).

Note that the CN_Local variable was different for each gene (as each gene has a different copy number profile), while the WGD and purity variables were the same for all genes. The Benjamini–Hochberg method was used to correct *P* values from the WGD term for multiple hypothesis testing. Genes were considered significantly associated with WGD status if they had an FDR *q* value of less than 0.05. Genes upregulated in more than ten tumour types were analysed with hypeR (ref. [51]) using the MSigDB Hallmark gene sets to identify biological categories enriched among these genes. Similarly, genes downregulated in more than ten tumour types were analysed with hypeR in the same fashion. To generate a volcano plot across tumour types, the coefficient for WGD was averaged and the FDR-corrected *q* values were combined using Fisher's method.

### Analysis of PSL scores

*Thresholded analysis*. Genes were assigned a binary classification (essential or non-essential) on the basis of cutoffs established by Project Achilles. In the database, a score of −1 is assigned to a gene when its depletion in a given cell line results in a viability defect equal to the depletion of a curated list of gold standard common-essential genes[52,53]. On the basis of this scoring system, we defined any gene with a score of −1 or less for a given cell line as essential. We then compared the fraction of cell lines in the $WGD^-$ and $WGD^+$ groups where a gene was essential. When a gene was essential in a significantly greater fraction of $WGD^+$ cell lines than $WGD^-$ cell lines (Fisher's exact test, $P < 0.1$) in a specific tumour subtype, it was considered a 'hit' in this analysis (Extended Data Fig. 3a).

*Non-thresholded analysis*. Within each tumour type, the median essentiality scores for each gene in the $WGD^-$ and $WGD^+$ cell lines were identified. When a gene showed a statistically significant enrichment in its median essentiality score in the $WGD^+$ compared with the $WGD^-$ cell lines (Wilcoxon's test, $P < 0.05$), and also had an essentiality score of −0.5 or less in the $WGD^+$ cell lines, it was considered at 'hit' in this analysis (Extended Data Fig. 3a).

*Final PSL score*. We used the thresholded analysis with the Fisher's exact test and non-thresholded analysis with the Wilcoxon's rank-sum test in each individual tumour type ($n = 12$) as well as in a combined pan-cancer analysis. These analyses were also performed separately for the CRISPR and RNAi data sets. Only genes that had measurable data in

95% of total cell lines were analysed. The final PSL score for each gene was the total number of instances a gene was found to be a hit across all analyses (Fig. 2b, Supplementary Tables 3, 4). As a result, some hits may have come entirely from either the CRISPR or the RNAi data sets, such as *KIF18A*, which was enriched for essentiality only in the CRISPR data set, probably because of insufficient knockdown in the RNAi data set.

### Cox proportional-hazards regression

To identify hazard ratios for progression-free survival and overall survival based on expression of *KIF18A* in samples in the TCGA, we separated the expression of *KIF18A* into tertiles within each subtype and compared outcomes in the upper tertile against the bottom two tertiles using a Cox proportional-hazards model. An FDR correction was applied to the *P* values to control for multiple hypothesis testing. The formula used was: outcome as a function of *KIF18A* expression.

### Cell culture

All breast-cancer cell lines were authenticated by ATCC and used at early passage numbers. Isogenic tetraploid cell lines were generated as described[6]. hTERT-RPE-1 cells were cultured in Dulbecco's modified Eagle medium (DMEM) with F12 medium (HyClone) supplemented with 10% fetal bovine serum (FBS; ThermoFisher) plus 50 IU ml$^{-1}$ penicillin and 50 μg ml$^{-1}$ streptomycin (ThermoFisher). HCT116, CAMA-1, MDA-MB-415, MDA-MB-453, MDA-MB-134-VI, MDA-MB-157, Hs578T, MDA-MB-231 and MDA-MB-361 cells were cultured in high-glucose DMEM (Gibco) supplemented with 10% FBS plus 50 IU ml$^{-1}$ penicillin and 50 μg ml$^{-1}$ streptomycin. ZR-75-30 and HCC1806 cells were cultured in RPMI medium (Gibco) supplemented with 10% FBS plus 50 IU ml$^{-1}$ penicillin and 50 μg ml$^{-1}$ streptomycin. MCF10A cells were cultured in DMEM/F12 (HyClone) supplemented with 5% horse serum (ThermoFisher), 20 ng ml$^{-1}$ epidermal growth factor (EGF; ThermoFisher), 500 ng ml$^{-1}$ hydrocortisone (ThermoFisher), 100 ng ml$^{-1}$ cholera toxin (Sigma) and 10 μg ml$^{-1}$ insulin (ThermoFisher), plus 50 IU ml$^{-1}$ penicillin and 50 μg ml$^{-1}$ streptomycin. RPE-1, HCT116 and MCF10A 2*N* and 4*N* cell lines were tested for mycoplasma contamination and confirmed to be negative.

### siRNA transfections

siRNA transfections using lipofectamine RNAiMAX (Invitrogen) were performed according to the manufacturer's instructions. The final concentration of KIF18A or control siRNA in the cell culture medium was 10 nM, excepting MCF10A KIF18A siRNA transfections, which were performed at a final concentration of 1 nM, and RRM1/RAD51 siRNA transfections, which were performed at a final concentration of 50 pM with control siRNA adjusted accordingly.

### siRNA sequences

The following sequences were used: non-targeting control (Dharmacon), 5′-UGGUUUACAUGUCGACUAA-3′; KIF18A (silencer select s37882; Ambion)[31], 5′-UCUCGAUUCUGGAACAAGCAG-3′; RAD51 (silencer select s11735; Ambion), 5′-UGAUUAGUGAUUACCACUGCT-3′; RRM1 (on-target plus SMARTpool; Dharmacon), 5′-UAUGAGGGCUCUCCAGUUA-3′, 5′-UGAGAGAGGUGCUUUCAUU-3′, 5′-UGGAAGACCUCUAUAACUA-3′, 5′-CUACUAAGCACCCUGACUA-3′.

### Inducible short hairpin (sh)RNA

We infected cells with a SMARTvector inducible lentiviral shRNA (Horizon) targeting KIF18A, and selected cells with puromycin (Santa Cruz Biotechnology) at 2 μg ml$^{-1}$. Cells were induced with doxycycline (Sigma) at 1 μg ml$^{-1}$ for 7 days and viability was assessed. The shRNA sequence was 5′-CGATGACACACATATAACACT-3′.

### Inducible CRISPR–Cas9

We infected cells with pCW-Cas9 plasmid (Addgene catalogue number 50661) and selected cells with puromycin at 2 μg ml$^{-1}$. To improve knockout efficiency, cells were then infected with two distinct KIF18A single guide (sg)RNA plasmids. Each sgRNA sequence was cloned into its own lenti–sgRNA–blast plasmid (Addgene catalogue number 104993), and these plasmids were co-packaged into lentivirus and used to infect cells, which were then selected with blasticidin (Sigma) at 5 μg ml$^{-1}$. The sequences for both KIF18A-targeting sgRNAs are available in ref. [38].

### Cell viability

All assays of cell viability were done using CellTiter-Glo (Promega) and performed according to the manufacturer's instructions. Viability was assessed via luminescence and analysed using BMG Labtech Optima v2.0R2 software.

### Drug treatments

AZ3146 (Tocris) was used at a concentration of 1 μM in HCT116 cells, 2 μM in MCF10A cells and 4 μM in RPE-1 cells. These concentrations were experimentally determined to be the minimum concentration required to inhibit the SAC in each respective cell line. MG132 (Selleck Chemicals) was used at the indicated concentrations (Fig. 2f).

### Antibodies

Antibodies were used at the following dilutions: rabbit polyclonal anti-KIF18A, 1:1,000 (Bethyl, catalogue number A301-080A); rabbit monoclonal anti-RRM1, 1:1,000 (Cell Signaling Technology, catalogue number 8637); rabbit polyclonal anti-RAD51, 1:1,000 (Santa Cruz Biotechnology, catalogue number sc-8349); rabbit monoclonal anti-cGAS, 1:250 (Cell Signaling Technology, catalogue number 15102); mouse monoclonal anti-phospho-histone H2A.X (Ser 139), 1:250 (Sigma-Aldrich, catalogue number 05-636-I); mouse monoclonal anti-p53, 1:2,000 (Santa Cruz Biotechnology, catalogue number sc-126); rabbit monoclonal anti-p21, 1:500 (Cell Signaling Technology, catalogue number 2947); rabbit monoclonal anti-Cas9, 1:1,000 (Active Motif, catalogue number 61978); rabbit monoclonal anti-GAPDH, 1:10,000 (Cell Signaling Technology, catalogue number 2118); mouse monoclonal anti-vinculin, 1:5,000 (Abcam, catalogue number ab18058); mouse monoclonal anti-tubulin (clone DM1A), 1:500 (immunofluorescence), 1:10,000 (western blot) (Sigma-Aldrich, catalogue number 05-829); rabbit polyclonal anti-pericentrin, 1:250 (Abcam, catalogue number ab4448).

### Population doubling assay

We seeded 10,000 cells in a 10-cm dish with AZ3146 at the indicated concentrations (Fig. 2d). Fresh drug was added every 3 days. After 8 days, cells were counted, and population doublings were calculated using the formula $\mathrm{PD}=\frac{\log[(N_{\mathrm{final}})/N_{\mathrm{initial}})]}{\log(2)}$.

### Live-cell imaging

Cells that stably expressed histone H2B fused with green fluorescent protein (GFP) were grown on glass-bottomed 12-well tissue culture dishes (Cellvis), and treated with drugs or transfected with siRNAs of interest. At 24 h post-treatment, imaging was performed on a Nikon TE2000-E2 inverted microscope equipped with the Nikon Perfect Focus system. The microscope was enclosed within a temperature- and atmosphere-controlled environment at 37 °C and 5% humidified $CO_2$. Fluorescent images were captured every 3 min with a ×20 0.5 NA Plan Fluor objective at multiple points for 72 h. Captured images were analysed for mitotic defects using NIS Elements software.

### Measurements of chromosome alignment

Live-cell imaging was used to track cells expressing H2B–GFP to the frame immediately preceding anaphase, and the distance from the metaphase plate to the widest oscillating chromosomes in each poleward direction was measured manually. We also measured the total

chromosomal distribution immediately before anaphase by recording the area of automatically generated regions of interest (ROIs) based on fluorescence intensity using NIS Elements software.

### Analysis of cell fate

Live-cell imaging was used to track cells treated with control siRNA to obtain the average cell-cycle time for each cell line. Cells treated with siKIF18A were called as 'arrested/delayed' if they spent greater than three standard deviations above the mean cell-cycle time of control cells in interphase.

### Immunofluorescence microscopy

Cells were plated on glass cover slips and then washed in microtubule-stabilizing buffer (MTSB) (4 M glycerol, 100 mM PIPES, pH 6.9, 1 mM EGTA, 5 mM $MgCl_2$) for 1 min, extracted in MTSB plus 0.5% Triton for 2 min, and washed again in MTSB for 2 min. Cells were then fixed in 1% electron-microscope-grade glutaraldehyde for 10 min. Glutaraldehyde was quenched by washing twice in $NaBH_4$ in water for 12 min each. Cells were then blocked for 30 min in Tris-buffered saline (TBS)–bovine serum albumin (BSA) (10 mM Tris, pH 7.5, 150 mM NaCl, 5% BSA, 0.2% sodium azide), and incubated with primary antibodies diluted in TBS–BSA for 60 min in a humidified chamber. Primary antibodies were visualized using species-specific fluorescent secondary antibodies (Molecular Probes) and DNA was detected with 2.5 μg $ml^{-1}$ Hoechst. Confocal immunofluorescence images were collected at 405 nm, 488 nm and 561 nm on a Nikon Ti-E inverted microscope with a C2+ laser scanning head. A series of 0.5-μm optical sections were acquired using a ×60 objective lens. Images presented in figures are maximum intensity projections of entire *z*-stacks.

### Spindle length

Spindles were measured using immunofluorescence microscopy. Cells were stained for tubulin/centrosomes, and spindle length was assessed by measuring the distance from centrosome to centrosome of cells in metaphase using NIS Elements software.

### Western botting

Cells were rinsed with ice-cold 1× phosphate-buffered saline (PBS) (Boston Bioproducts) and lysed immediately with cell lysis buffer (2% w/v SDS, 10% glycerol, 60 mM Tris-HCl) supplemented with 1× HALT protease and phosphatase dual-inhibitor cocktail (ThermoFisher). Cell lysates were then sonicated for 15 s at 20 kHz, and sample buffer (Boston Bioproducts) was added to a final concentration of 1×, after which protein samples were incubated at 95 °C for 5 min.

Cell lysates were resolved via SDS–PAGE (resolving/separating gel: 7.5% acrylamide, 375 mM Tris-HCl (pH 8.8), 0.1% SDS, 0.25% ammonium persulfate, 0.15% tetramethylethylenediamine; stacking gel: 4% acrylamide, 125 mM Tris-HCl (pH 6.8), 0.1% SDS, 0.5% ammonium persulfate, 0.3% tetramethylethylenediamine) in SDS–PAGE running buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS). Samples were passed through the stacking gel layer at 130 V for 15 min, followed by resolution of samples at 230 V for 25 min. Samples were transferred to 0.45-μm Immobilon polyvinylidene difluoride (PVDF) membranes (EMD Millipore) using a wet-tank transfer system (Bio-Rad) in Towbin transfer buffer (25 mM Tris-HCl, 192 mM glycine, 10% methanol) for 16 h at 30 mA at 4 °C. Following transfer, membranes were blocked in TBS with 0.5% Tween-20 (10 mM Tris-HCl, 150 mM NaCl, 0.5% Tween-20) containing 5% non-fat dried milk (NFDM) for 1 h and then incubated overnight at 4 °C with primary antibodies diluted in 1% NFDM TBS with 0.5% Tween-20 solution. Following incubation, membranes were rinsed in TBS with 0.5% Tween-20 solution for 30 min with vigorous shaking, and then incubated with secondary antibodies diluted in 1% NFDM TBS with 0.5% Tween-20 solution for 1 h at room temperature, followed again by 30 min of rinsing with vigorous shaking in TBSwith 0.5% Tween-20 solution.

Primary antibodies were detected using horseradish-peroxidase-conjugated species-specific secondary antibodies (1:5,000, Cell Signaling Technology) and Clarity ECL blotting substrate (Bio-Rad) or Clarity Max ECL blotting substrate (Bio-Rad). Imaging of blots were performed using the ChemiDoc XRS+ imaging system (Bio-Rad), and quantitative densitometry was performed using Bio-Rad ImageLab software.

### Aneuploidy scores

Aneuploidy scores were calculated for 998 cell lines in the Cancer Cell Line Encyclopedia (CCLE) according to previously published methods[19,54]. Briefly, we used ABSOLUTE to determine the total copy number of chromosome arms in the genome, and designated each arm as either amplified or deleted depending on whether the longest altered segment covered more than 80% of the chromosome arm. Each arm was assigned −1 if lost, +1 if gained, 0 if unchanged (with less than 20% of the arm affected), or 'not called' if intermediate-sized copy number alterations occurred. The total aneuploidy score for each cell line was then calculated as the sum total of altered arms, with a range of 0 (no arm alterations) to 39 (all arms altered: long and short arms for each non-acrocentric autosomal chromosome, and long arms only for chromosomes 13, 14, 15, 21 and 22).

### TCGA abbreviations

ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; ESCA, oesophageal carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumours; THYM, thymoma; THCA, thyroid carcinoma; UCS, uterine carcinosarcoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

TCGA data used here are publicly available. All other data supporting this study are available within the article or supplementary figures or are available from the authors upon request.

## Code availability

All code used here is available at https://github.com/campbio/Manuscripts/tree/master/Quinton_WGD_2020.


44. Ellrott, K. et al. Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines. *Cell Syst*. **6**, 271–281 (2018).
45. Knijnenburg, T. A. et al. Genomic and molecular landscape of DNA damage repair deficiency across The Cancer Genome Atlas. *Cell Rep*. **23**, 239–254 (2018).
46. Bailey, M. H. et al. Comprehensive characterization of cancer driver genes and mutations. *Cell* **173**, 371–385 (2018).
47. Thorsson, V. et al. The immune landscape of cancer. *Immunity* **48**, 812–830 (2018).
48. Miao, D. et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat. Genet*. **50**, 1271–1281 (2018).
49. Liu, D. et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat. Med*. **25**, 1916–1927 (2019).

50. Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhim, R. & Getz, G. GISTIC20 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
51. Federico, A. & Monti, S. hypeR: an R package for geneset enrichment workflows. *Bioinformatics* **36**, 1307–1308 (2020).
52. McFarland, J. M. et al. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat. Commun.* **9**, 4610 (2018).
53. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
54. Ben-David, U. et al. Patient-derived xenografts undergo mouse-specific tumor evolution. *Nat. Genet.* **49**, 1567–1575 (2017).

**Acknowledgements** The findings published here are in part based upon data generated by the TCGA Research Network (https://www.cancer.gov/tcga). We thank J. Weinberg for statistical advice; J. Stumpff for technical advice, reagents, and sharing unpublished data; D. Pellman for sharing reagents; E. Van Allen and D. Liu for technical advice; and S. Carter for sharing ABSOLUTE data for cell lines in the CCLE. R.J.Q. is supported by a Canadian Institutes of Health Research Doctoral Foreign Study Award (152266). N.J.G. is a member of the Shamim and Ashraf Dahod Breast Cancer Research Laboratories and is supported by National Institutes of Health (NIH) grants CA154531 and GM117150, the Karin Grunebaum Foundation, the Smith Family Awards Program, the Melanoma Research Alliance and the Searle Scholars Program. This work was also supported in part by a pilot grant from the American Cancer Society (ACS) and the Boston University Clinical and Translational Science Institute Bioinformatics Group, who are in turn supported by a grant from the NIH/National Center for Advancing Translational Sciences (NCATS; 1UL1TR001430).

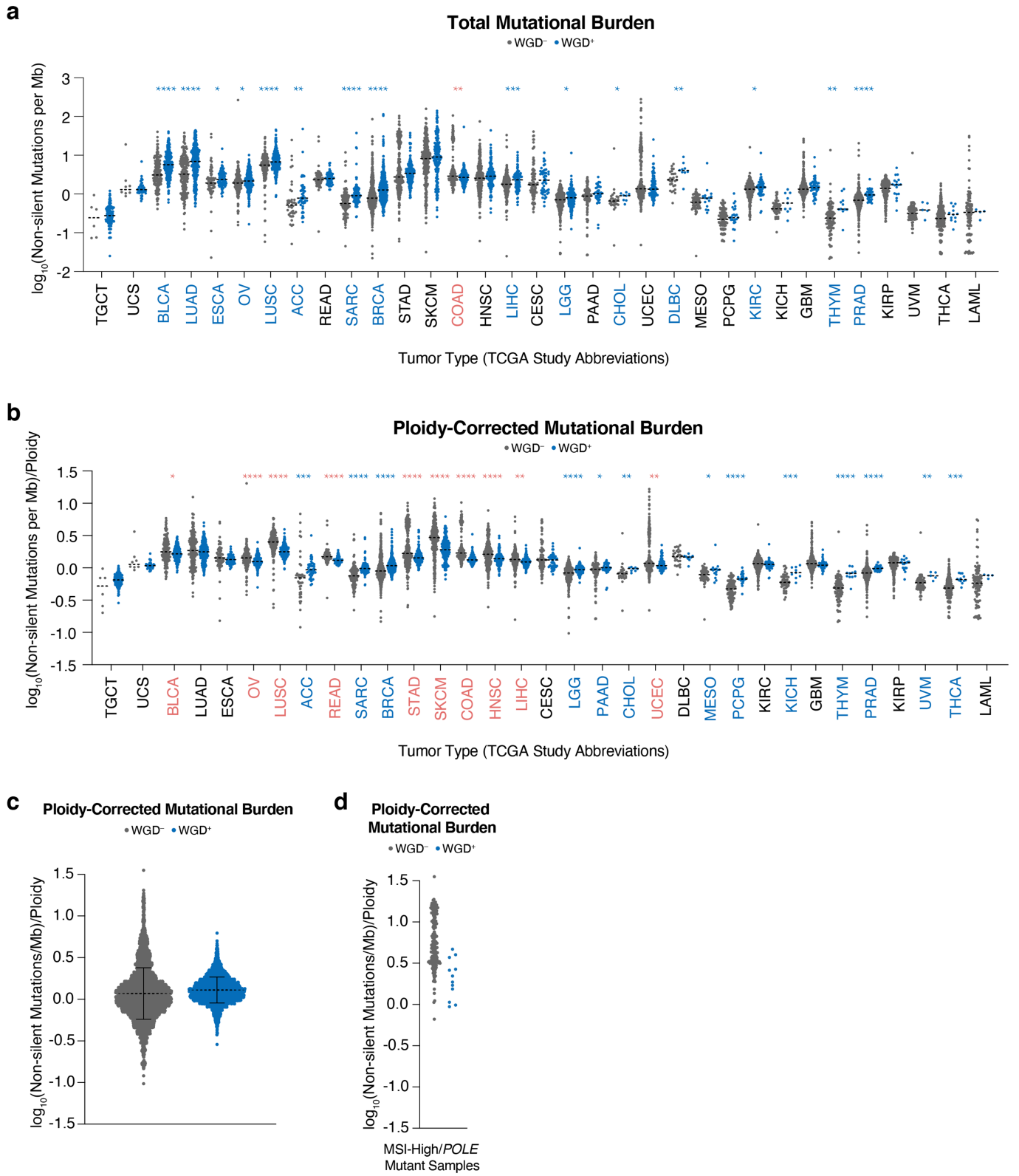
**Author contributions** R.J.Q. and N.J.G. designed the experiments and wrote the manuscript. R.J.Q. performed most of the cell biological assays and imaging analysis. A.D., C.J.T. and S.P. assisted R.J.Q. with the cell biological assays. K.K. and M.A.V. analysed images. T.S.K. generated the isogenic diploid and tetraploid HCT-116 cells. J.V. generated cell lines. N.H. and A.L.M. performed animal studies. R.J.Q., A.M.T., Y.K., N.P. and J.D.C. performed the computational analyses. All authors edited the manuscript.

**Competing interests** The authors declare no competing interests.

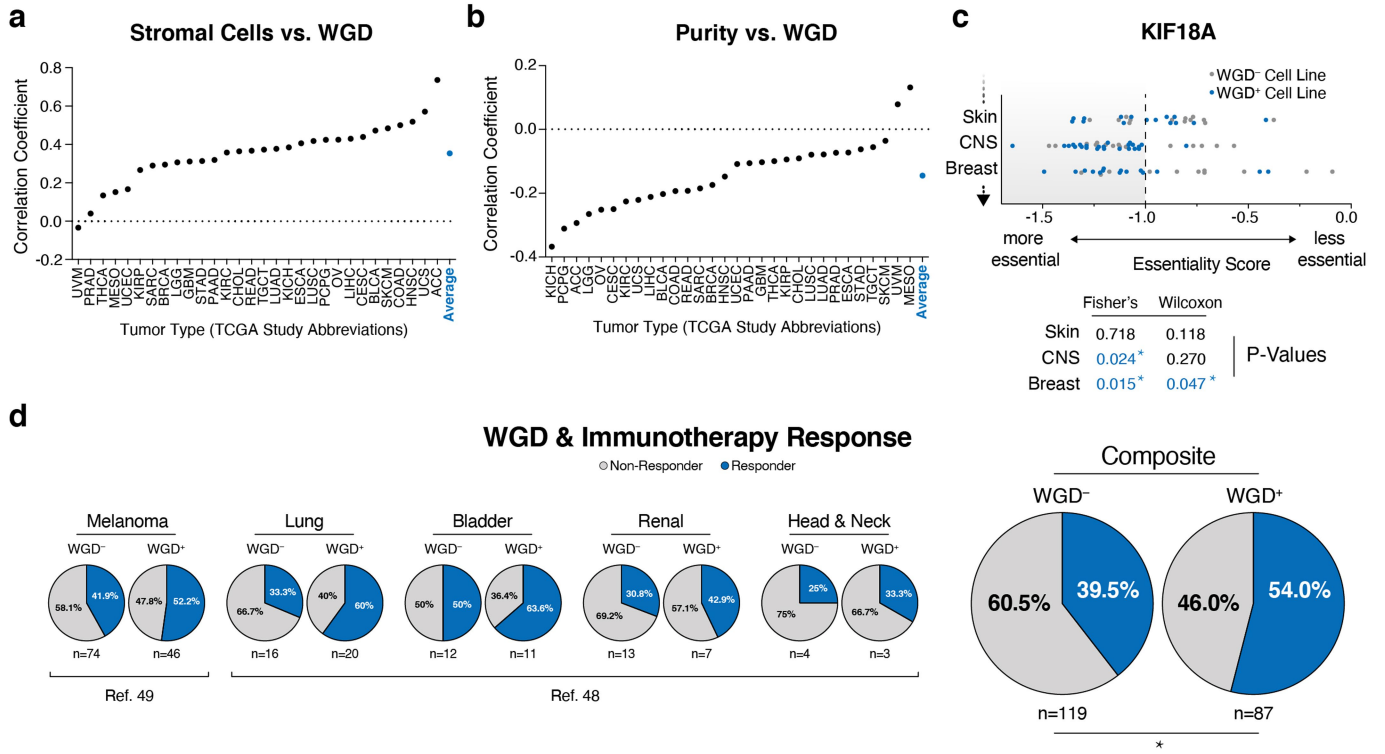
**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41586-020-03133-3.

**Correspondence and requests for materials** should be addressed to N.J.G.



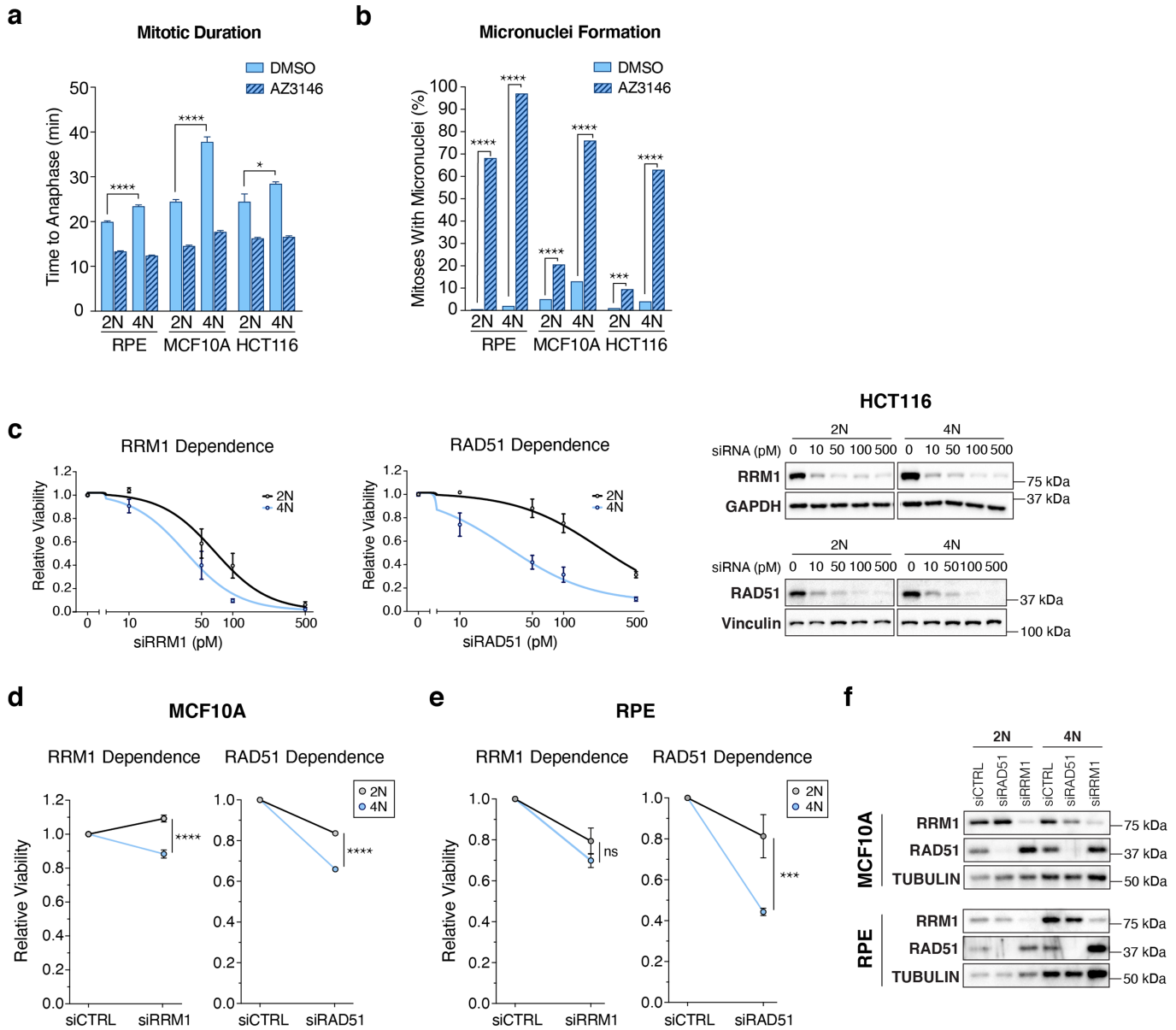

**Extended Data Fig. 1 | Mutational burden in WGD$^+$ and WGD$^-$ tumours.** **a**, Total mutational burden in the indicated tumour subtypes across 9,240 TCGA samples (horizontal dotted lines show medians; two-sided Wilcoxon rank-sum test; red asterisks indicate higher burden in WGD$^-$ samples and blue asterisks indicate higher burden in WGD$^+$ samples). **b**, Ploidy-corrected mutational burden in the indicated tumour subtypes across 9,240 TCGA samples (dotted lines show medians; two-sided Wilcoxon rank-sum test; red asterisks indicate higher burden in WGD$^-$ samples and blue asterisks indicate higher burden in WGD$^+$ samples). **c**, Ploidy-corrected mutational burden in WGD$^+$ and WGD$^-$ samples in the TCGA ($n$ = 9,414 samples; dotted line shows means ± s.d.). **d**, Ploidy-corrected mutational burden of TCGA WGD$^+$ and WGD$^-$ samples with MSI/*POLE* mutations ($n$ = 174 samples). *$P$ < 0.05, **$P$ < 0.01, ***$P$ < 0.001, ****$P$ < 0.0001.



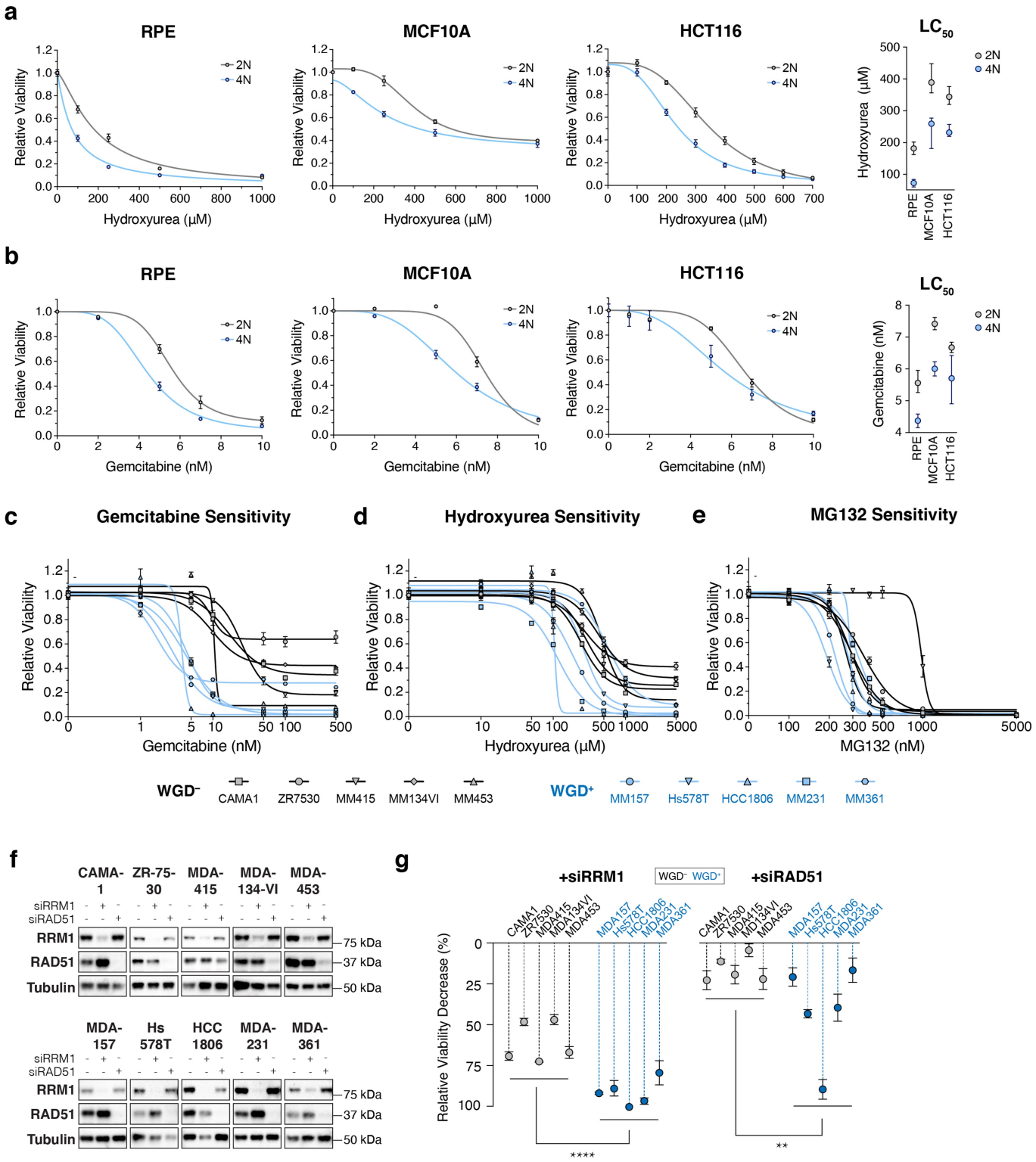

h

| | RPE | | MCF10A | |
|---|---|---|---|---|
| Cell Line | 2N | 4N | 2N | 4N |
| Missegregation Rate | ~0.02 | ~0.02 | N/A | N/A |
| FACS Profile | Stable | Stable | Stable | Stable |
| Modal Chromosome No. (Range) | 46 (45-47) | 92 (91-92) | 47 (40-48) | 94 (79-107) |
| Reference | Ref. 3 | | Ref. 22 | |

**Extended Data Fig. 2 | Characteristics of WGD⁺ cells. a**, Correlation of the fraction of stromal cells in the indicated tumour subtypes with WGD (Pearson's correlation). **b**, Correlation of purity with WGD (Pearson's correlation). **c**, Illustration of our PSL analysis using gene essentiality scores for *KIF18A* in the Project Achilles CRISPR data set. Asterisked *P* values in blue represent instances where the cutoff for enrichment in WGD⁺ cell lines was met in either our thresholded (two-sided Fisher's exact) or our non-thresholded (two-sided Wilcoxon's) analyses (see Methods). CNS, central nervous system. **d**, Fractions of individuals whose cancers responded or did not respond to blockade of PD1, according to WGD status (two-sided Fisher's exact test; $P = 0.0351$). See refs. [48,49]. **e**, Rate of chromosome missegregation in HCT116 cells ($n = 3{,}107$ 2*N* cells, 2,594 4*N* cells; shown are means ± s.d.). **f**, DNA fluorescence-activated cell sorting (FACS) profile of diploid and tetraploid HCT116 cells at 40 and 70 days of culture. **g**, Karyotype of diploid and tetraploid HCT-116 cells with the modal chromosome number and range ($n = 20$ karyotypes analysed per condition). **h**, Previously published data[3,22] showing the stability of isogenic diploid and tetraploid RPE and MCF10A cell lines. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.




**Extended Data Fig. 3 | Validation of WGD$^+$ vulnerabilities in isogenic 2*N*/4*N* cells. a**, Mitotic duration of the indicated cells following the indicated treatments ($n = 200$ cells; two-sided Student's unpaired *t*-test; shown are means ± s.e.m.; $P < 0.0001, < 0.0001, 0.0265$, respectively). **b**, Fractions of mitoses that generate micronuclei following the indicated treatments ($n = 200$ cells; two-sided Student's unpaired *t*-test; $P < 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.0002, < 0.0001$, respectively). **c**, Left, relative viability of 2*N* and 4*N* HCT116 cells seven days after treatment with the indicated siRNAs at the indicated concentrations. Right, western blot showing protein knockdown 48 h after treatment with siRNA ($n = 3$ independent experiments; shown are means ± s.e.m. at each dose; for gel source data, see Supplementary Fig. 1). **d**, Relative viability of 2*N* and 4*N* MCF10A cells seven days after treatment with the indicated siRNAs at 50 pM concentration ($n = 3$ independent experiments; one-sided Student's unpaired *t*-test; means ± s.e.m.; $P < 0.0001, < 0.0001$). **e**, Relative viability of 2*N* and 4*N* RPE cells five days after treatment with the indicated siRNAs at 50 pM concentration ($n = 3$ independent experiments; one-sided Student's unpaired *t*-test; means ± s.e.m.; $P = 0.090, 0.0007$, respectively). **f**, Representative western blot showing knockdown of the indicated proteins 48 h after treatment with the indicated siRNAs ($n = 3$ independent experiments; for gel source data, see Supplementary Fig. 1). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

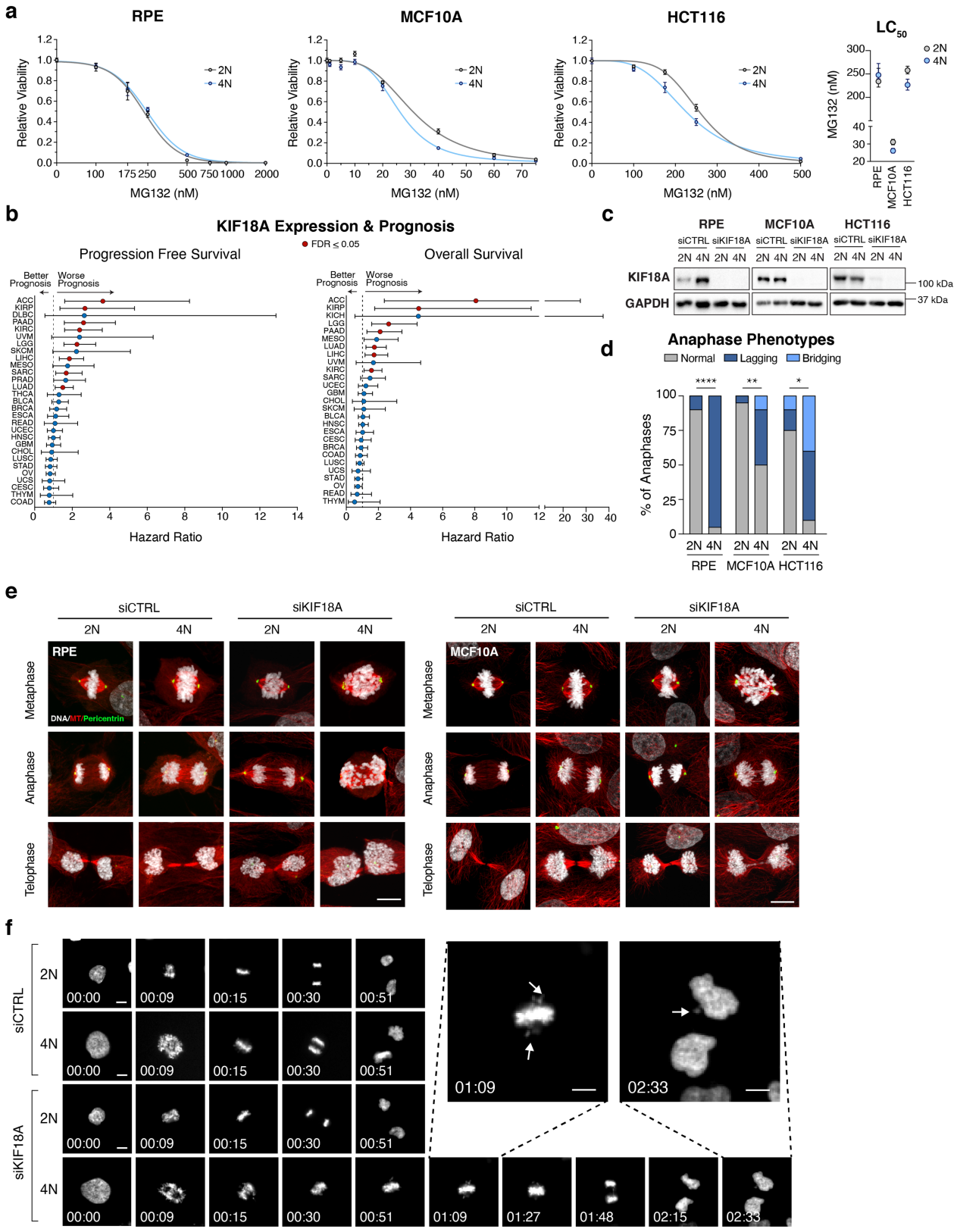






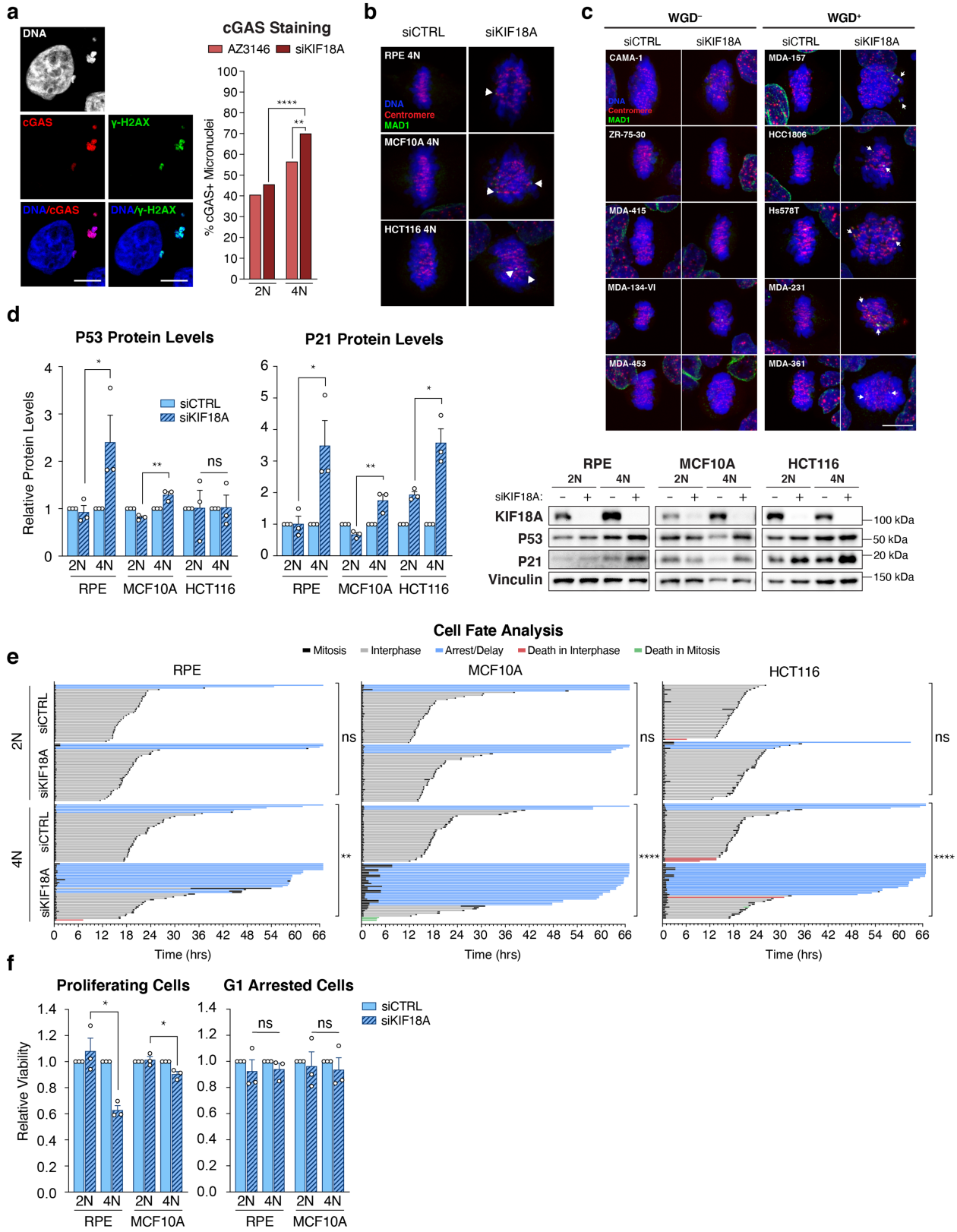



**Extended Data Fig. 4 | Validation of WGD$^+$ vulnerabilities in breast-cancer cells. a**, **b**, Dose–response curves for the indicated cell lines following the indicated treatments after seven days, with accompanying $LC_{50}$ values ($n = 3$ independent experiments; nonlinear regression with variable slope; graphs show mean relative viability ± s.e.m. at each dose or mean $LC_{50}$ ± 95% confidence interval). **c**–**e**, Dose–response curves for 5 WGD$^-$ and 5 WGD$^+$ breast cancer cell lines seven days after treatment with the indicated drugs and concentrations ($n = 3$ independent experiments; nonlinear regression with variable slope; means ± s.e.m. at each dose). **f**, Representative western blot showing knockdown of the indicated proteins in breast-cancer cell lines 48 h after treatment with the indicated siRNAs ($n = 3$ independent experiments; for gel source see Supplementary Fig. 1). **g**, Relative viability decrease in WGD$^+$ and WGD$^-$ breast-cancer cell lines seven days after treatment with the indicated siRNAs (two-sided Wilcoxon's rank-sum test; means ± s.e.m.; $P < 0.0001$ and $P = 0.0027$, respectively). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.



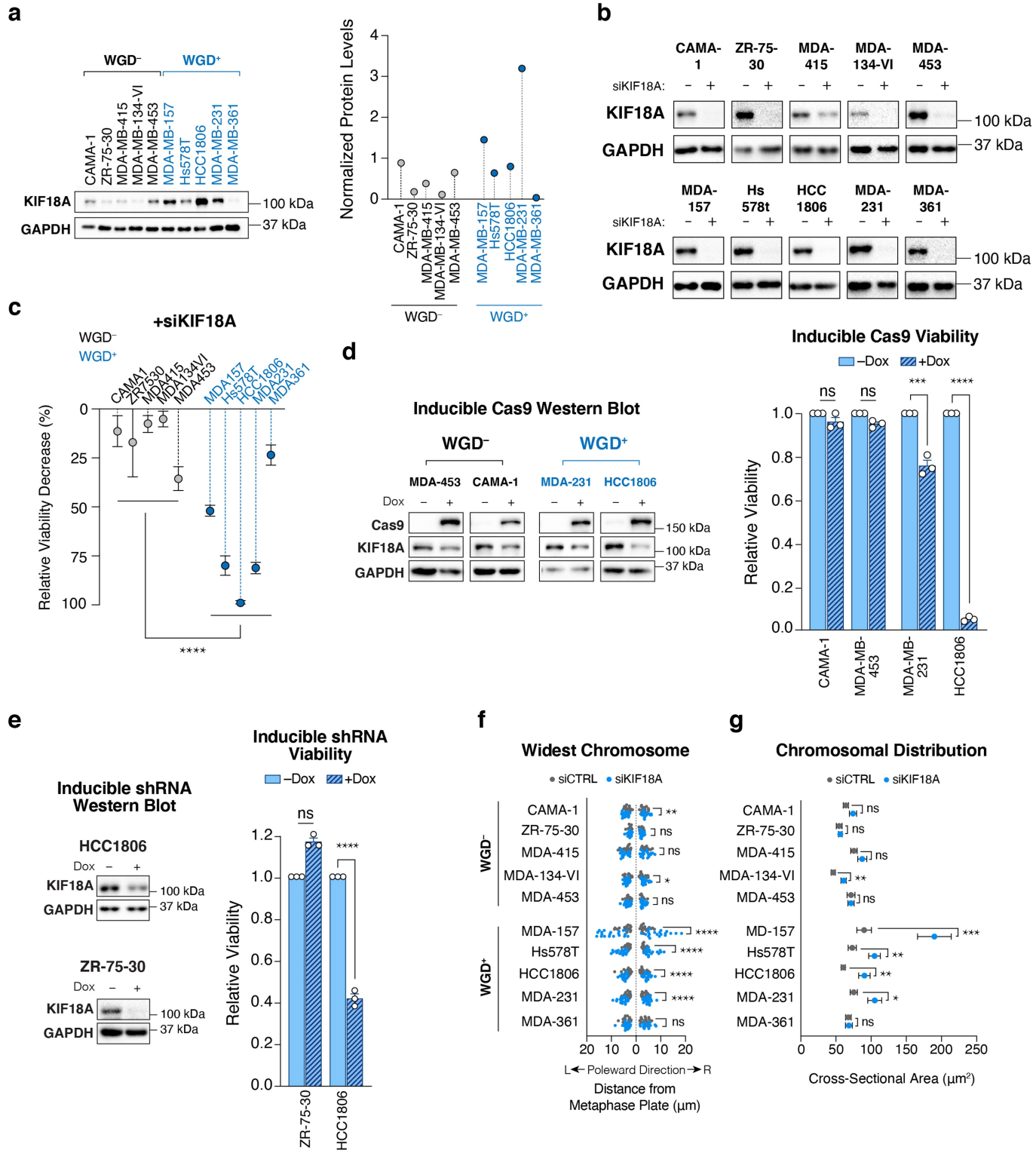

**Extended Data Fig. 5** | See next page for caption.

**Extended Data Fig. 5 | Mitotic fidelity in WGD$^+$ cells following KIF18A depletion. a**, Dose–response curves for the indicated cell lines following a seven-day treatment with MG132, with accompanying $LC_{50}$ values ($n = 3$ independent experiments; nonlinear regression with variable slope; graphs show mean relative viability ± s.e.m. at each dose and mean $LC_{50}$ ± 95% confidence interval). **b**, Progression-free survival and overall survival for individuals in the upper tertile of tumour *KIF18A* expression in the TCGA (Cox's proportional-hazards regression; graph shows hazard ratios ± 95% confidence interval). **c**, Representative western blot showing KIF18A levels following transfection with the indicated siRNAs in the indicated cell lines ($n = 3$ independent experiments; for gel source data, see Supplementary Fig. 1). **d**, Anaphase phenotypes following depletion of KIF18A ($n = 20$ cells per condition; asterisks indicate *P* values from two-sided Fisher's exact tests comparing the fraction of anaphases with lagging chromosomes; $P < 0.0001$, 0.0033, 0.0187, respectively). **e**, Representative confocal images showing phases of mitosis in the indicated cell lines 48 h after transfection with the indicated siRNAs (representative images from two independent experiments; scale bar, 10 µm). **f**, Representative still images from 2*N* and 4*N* MCF10A cells progressing through mitosis after transfection with the indicated siRNAs. Chromosomes labelled with H2B–GFP are shown in white. Arrows in enlarged images show oscillating chromosomes during metaphase and the generation of a micronucleus (shown are h:min; scale bar, 10 µm). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.



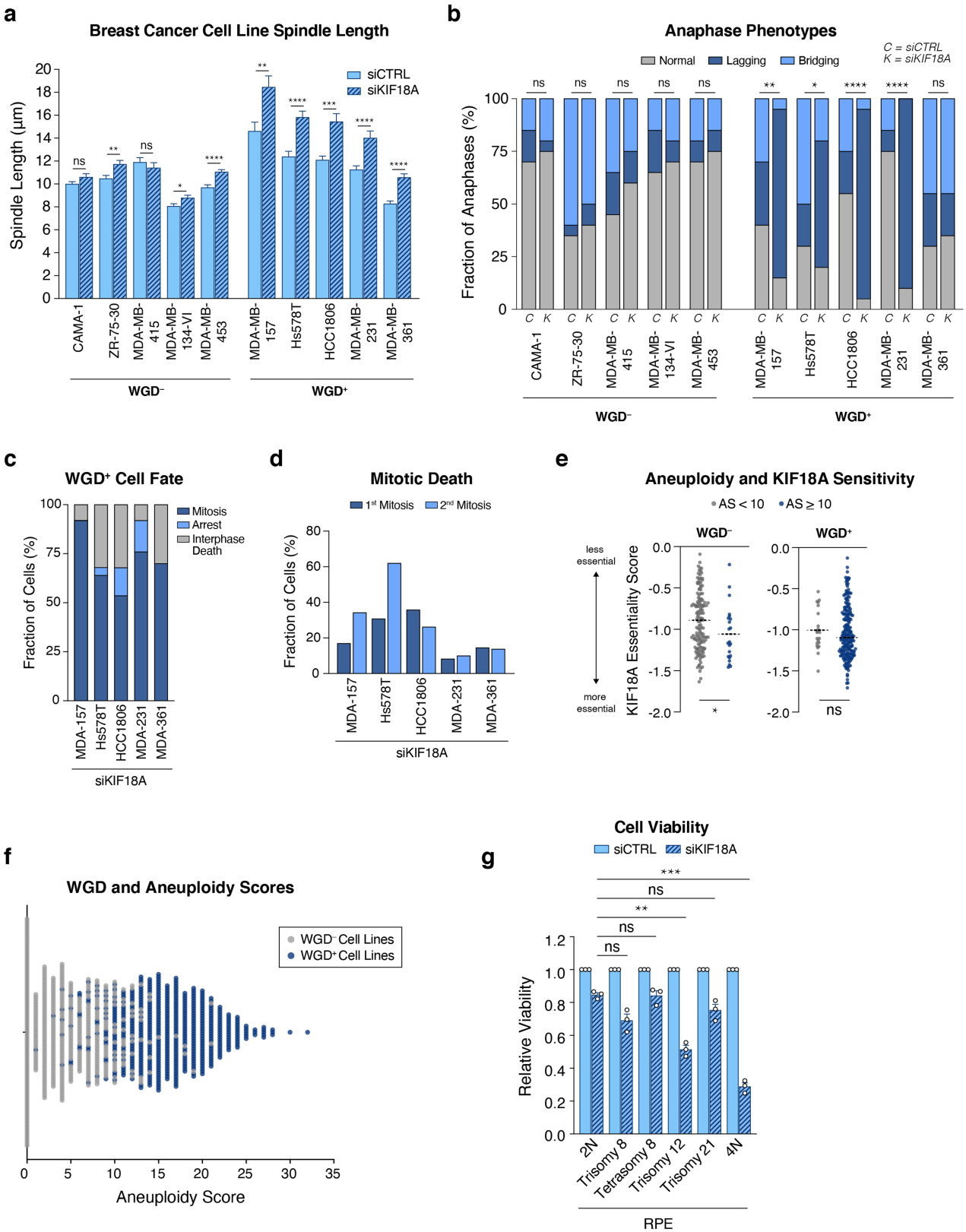

**Extended Data Fig. 6** | See next page for caption.

**Extended Data Fig. 6 | Analysis of cell fate in WGD$^+$ cells following KIF18A depletion. a**, Left, representative images of a 4*N* MCF10A cell four days after transfection with siKIF18A and stained for DNA, cGAS and γ-H2AX. Right, fraction of micronuclei in 2*N* and 4*N* MCF10A cells with the indicated treatments that stained positive for cGAS ($n = 200$ micronuclei per condition; two-sided Fisher's exact test; scale bar, 10 µm; $P < 0.0001$, $P = 0.0069$, respectively). **b**, **c**, Representative confocal images of the indicated cell lines 48 h after transfection with the indicated siRNAs. Arrowheads highlight MAD1-positive kinetochores in misaligned chromosomes (scale bar, 10 µm; representative images from two independent experiments). **d**, Right, representative western blot showing levels of the indicated proteins after treatment with the indicated siRNAs, and left, graphs showing relative protein levels normalized to loading control ($n = 3$ independent experiments; one-sided Student's unpaired *t*-test; means ± s.e.m.; $P = 0.0337$, 0.0030, 0.0674, 0.0421, 0.0067, 0.0227, respectively; for gel source data, see Supplementary Fig. 1). **e**, Cell fates of the indicated lines, tracked for 3 days beginning 18 h after transfection with the indicated siRNAs ($n = 40$ cells per condition; two-sided Fisher's exact test comparing the fraction of cells arresting/delaying in interphase relative to the control group; $P = 0.0016$, $< 0.0001$, $< 0.0001$, respectively). **f**, Relative viability of the indicated cell lines four days after transfection with the indicated siRNAs ($n = 3$ independent experiments; two-sided Student's unpaired *t*-test; means ± s.e.m.; $P = 0.0132$, 0.0310, 0.8808, 0.8615, respectively). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.




**Extended Data Fig. 7** | See next page for caption.

**Extended Data Fig. 7 | PSL effect of KIF18A depletion. a**, Bottom left, western blot showing endogenous KIF18A levels in the indicated cell lines; top and right, graphs showing respective protein levels normalized to the GAPDH loading control (representative blot from three independent experiments; for gel source data, see Supplementary Fig. 1). **b**, Representative western blot showing KIF18A levels 48 h after transfection with the indicated siRNAs ($n = 3$ independent experiments; for gel source data, see Supplementary Fig. 1). **c**, Relative viability decrease in $WGD^+$ and $WGD^-$ breast-cancer cell lines seven days after treatment with the indicated siRNAs ($n = 3$ independent experiments; two-sided Wilcoxon's rank-sum test; means ± s.e.m.; $P < 0.0001$). **d**, Right, relative viability seven days after induction of Cas9 in cells with sgRNA targeting KIF18A. Left, western blot showing protein depletion 72 h after induction ($n = 3$ independent experiments; means ± s.e.m.; two-sided Student's unpaired *t*-test; $P = 0.0007, < 0.0001$, respectively; for gel source data, see Supplementary Fig. 1). **e**, Relative viability seven days after induction of shRNA targeting KIF18A, with a western blot showing protein depletion 120 h after induction ($n = 3$ independent experiments; means ± s.e.m.; one-sided Student's unpaired *t*-test; $P < 0.0001$; for gel source data, see Supplementary Fig. 1). **f**, Chromosomes with the widest oscillations in each poleward direction immediately before anaphase ($n = 20$ cells per condition; two-sided Student's unpaired *t*-test; $P = 0.0022, 0.1781, 0.1487, 0.0136, 0.0820, < 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.4132$, respectively). **g**, Two-dimensional cross-sectional area of the entire body of chromosomes immediately before anaphase ($n = 20$ cells per condition; two-sided Student's unpaired *t*-test; $P = 0.1178, 0.7545, 0.1440, 0.0034, 0.9989, 0.0005, 0.0033, 0.0012, 0.0110, 0.9089$, respectively; means ± s.e.m.). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.




**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | Effects of KIF18 depletion in aneuploid cells.** **a**, Measurement of spindle length (centrosome-to-centrosome) after transfection with the indicated siRNAs ($n = 20$ cells per condition; two-sided Student's unpaired *t*-test; shown are means ± s.e.m.). **b**, Anaphase phenotypes following depletion of KIF18A ($n = 20$ cells per condition; asterisks indicate *P* values from two-sided Fisher's exact tests comparing the fraction of anaphases with lagging chromosomes). **c**, Fractions of cells in each cell line undergoing the indicated fates after completing a KIF18A-deficient mitosis that resulted in the formation of micronuclei ($n = 25$ cells per condition). **d**, Fractions of cells in each cell line that experience mitotic death in their first and second mitoses following KIF18A depletion ($n = 25$ cells per condition). **e**, KIF18A essentiality scores for $WGD^-$ and $WGD^+$ cell lines, segregated into 'highly aneuploid' (aneuploidy score (AS) score of 10 or more) and 'non-highly aneuploid' (AS < 10) categories on the basis of aneuploidy scores (see Methods) (dotted lines show means; two-sided Wilcoxon's rank-sum test; $P = 0.02583$, 0.3682, respectively). **f**, Aneuploidy scores and WGD status for 998 cancer cell lines in the CCLE. **g**, Relative viability of the indicated cell lines seven days after transfection with the indicated siRNAs ($n = 3$ independent experiments; each condition normalized to respective control; one-way ANOVA with Dunnett's post hoc test; means ± s.e.m.; $P = 0.1676$, > 0.9999, 0.0040, 0.2698, 0.0007, respectively). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

nature research

Corresponding author(s): Neil Ganem

Last updated by author(s): Dec 8, 2020

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Imaging Software: NIS Elements-AR 4.13.04. Immunoblot Imaging Software: ImageLab v5.2.1 v11. Luminescence analyzing software: BMG Labtech Optima v2.20R2 |
| Data analysis | Graphpad Prism 9.0. R version 3.2.3. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

TCGA data used in the study are publicly available. The authors declare that all other data supporting this study are available within the article or supplementary figures, acessible by the included link, or are available from the authors upon request. https://github.com/campbio/Manuscripts/tree/master/Quinton_WGD_2020

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | Sample sizes were not predetermined based on statistical methods, but were chosen according to the standards of the field (at least three independent biological replicates for each condition), which gave sufficient statistics for the effect sizes of interest. |
| Data exclusions | No data were excluded for this study. |
| Replication | Where statistical comparisons are being made, experiments were performed using at least three independent biological replicates to ensure reproducibility. All experiments were consistently reproducible. |
| Randomization | Experimental groups were not randomized. All experiments were performed with appropriate positive and negative controls. The experiments performed herein were performed using cell culture or analysis of previously collected patient data. Randomization is not applicable to these experimental modalities. |
| Blinding | Investigators were blinded to group allocation for breast cancer cell lines studies, as WGD status was confirmed after data collection. Investigators were blinded for all live cell imaging studies. Other experiments were not blinded because experimental manipulations required knowledge of cell lines, but all experiments were performed and analyzed using unbiased methodology. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

| | |
|---|---|
| Antibodies used | Rabbit polyclonal anti-KIF18A (Bethyl Cat # A301-080A)<br>Rabbit monoclonal anti-RRM1 (Cell Signaling Technology Cat # 8637)<br>Rabbit polyclonal anti-RAD51 (Santa Cruz Biotechnology Cat # sc-8349)<br>Rabbit monoclonal anti-cGAS (Cell Signaling Technology Cat # 15102)<br>Mouse monoclonal anti-phospho-histone H2A.X (Ser 139) (Sigma-Aldrich Cat # 05-636-I)<br>Rabbit monoclonal anti-GAPDH (Cell Signaling Technology Cat # 2118)<br>Mouse monoclonal anti-VINCULIN (Abcam Cat # ab18058)<br>Mouse monoclonal anti-Tubulin (clone DM1A) (Sigma-Aldrich Cat # 05-829)<br>Rabbit polyclonal anti-PERICENTRIN (Abcam Cat # ab4448) |
| Validation | All antibodies are commercially available and validated by their manufacturer.<br>Bethyl Validation Statement: By analyzing western blots of immunoprecipitates, in conjunction with a western blot of the whole-cell lysate, we can verify the mobility of the target protein. Using multiple dilutions and a broad spectrum of whole-cell lysates, our scientists can verify selective binding as well as the antibody's specificity, reproducibility, and sensitivity. At the conclusion of this two-phase process, only antibodies exhibiting the following characteristics qualify to be released: Specific recognition of the target protein, Selective recognition of the target protein, Acceptable sensitivity, and Reproducibility.<br>Cell Signaling Validation Statement: ST™ antibodies are produced in-house and validated extensively according to a rigorous protocol. Validation Steps Include: Examination of several cell lines and/or tissues of known expression levels allows accurate determination of species cross-reactivity and verifies specificity. Treatment of cell lines with growth factors, chemical activators |

or inhibitors, which induce or inhibit target expression, verifies specificity. Phosphatase treatment confirms phospho-specificity. The use of siRNA transfection or knockout cell lines verifies target specificity. Side-by-side comparison of lots to ensures lot-to-lot consistency. Optimal dilutions and buffers are predetermined, positive and negative cell extracts are specified, and detailed protocols are already optimized, saving valuable time and reagents.
Sigma-Aldrich Validation Statement: We are a global company with manufacturing facilities worldwide. Our development and manufacturing processes are subject to rigorous quality control and quality assurance measures, and each of our antibody products is supplied with a comprehensive Certificate of Analysis and Product Information Sheet.
Abcam Validation Statement: Antibodies are validated in western blot using lysates from cells or tissues that we have identified to express the protein of interest. Once we have determined the right lysates to use, western blots are run and the band size is checked for the expected molecular weight. We will always run several controls in the same western blot experiment, including positive lysate and negative lysate. When possible, we also include knock-out (KO) cell lines as a true negative control for our western blots. We are always increasing the number of KO-validated antibodies we provide. In addition, we run old stock alongside our new stock. If we know the old stock works well, this also acts as a suitable positive control. If the western blot result gives a clear clean band and we are happy with the result from the control lanes, these antibodies will be passed and added to the catalog.

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | ATCC: RPE, HCT-116, MCF10A, CAMA-1, MDA-MB-415, MDA-MB-134-VI, MDA-MB-157, Hs578T, MDA-MB-231, MDA-MB-361, ZR-75-30, MDA-MB-453, and HCC180 |
| Authentication | All cell lines were purchased directly from ATCC, but we did not independently authenticate them. |
| Mycoplasma contamination | RPE, HCT, and MCF10A cell lines tested negative for mycoplasma. All breast cancer cell lines were passaged initially in media containing Plasmocin (anti-mycoplasmic agent) after purchase from ATCC but were not tested for mycoplasma. |
| Commonly misidentified lines (See ICLAC register) | None of the cell lines used are commonly misidentified. |