NEOPLASIA
www.neoplasia.com

Volume 12 Number 11 November 2010 pp. 877–888 877

Integrative Functional Genomics Analysis of Sustained Polyploidy Phenotypes in Breast Cancer Cells Identifies an Oncogenic Profile for GINS2[1,2]

Juha K. Rantala*, Henrik Edgren†, Laura Lehtinen*, Maija Wolf†, Kristine Kleivi*,‡, Hans Kristian Moen Vollan‡,§,¶, Anna-Riina Aaltola*, Petra Laasola*, Sami Kilpinen†, Petri Saviranta*, Kristiina Iljin*,# and Olli Kallioniemi*,†

*Medical Biotechnology, VTT Technical Research Centre of Finland, Turku, Finland; †Institute for Molecular Medicine Finland (FIMM), Biomedicum 2U, University of Helsinki, Helsinki, Finland; ‡Department of Genetics, Institute for Cancer Research, Oslo University Hospital – Radiumhospitalet, Oslo, Norway; §Department of Breast and Endocrine Surgery, Division of Surgery and Cancer, Oslo University Hospital Ullevål, Oslo, Norway; ¶Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway; #Turku Centre for Biotechnology, University of Turku, Turku, Finland

Abstract

Aneuploidy is among the most obvious differences between normal and cancer cells. However, mechanisms contributing to development and maintenance of aneuploid cell growth are diverse and incompletely understood. Functional genomics analyses have shown that aneuploidy in cancer cells is correlated with diffuse gene expression signatures and aneuploidy can arise by a variety of mechanisms, including cytokinesis failures, DNA endoreplication, and possibly through polyploid intermediate states. To identify molecular processes contributing to development of aneuploidy, we used a cell spot microarray technique to identify genes inducing polyploidy and/or allowing maintenance of polyploid cell growth in breast cancer cells. Of 5760 human genes screened, 177 were found to induce severe DNA content alterations on prolonged transient silencing. Association with response to DNA damage stimulus and DNA repair was found to be the most enriched cellular processes among the candidate genes. Functional validation analysis of these genes highlighted *GINS2* as the highest ranking candidate inducing polyploidy, accumulation of endogenous DNA damage, and impairing cell proliferation on inhibition. The cell growth inhibition and induction of polyploidy by suppression of GINS2 was verified in a panel of breast cancer cell lines. Bioinformatic analysis of published gene expression and DNA copy number studies of clinical breast tumors suggested GINS2 to be associated with the aggressive characteristics of a subgroup of breast cancers *in vivo*. In addition, nuclear GINS2 protein levels distinguished actively proliferating cancer cells suggesting potential use of GINS2 staining as a biomarker of cell proliferation as well as a potential therapeutic target.

Neoplasia (2010) 12, 877–888

Abbreviations: cDNA, complementary DNA; GINS2, GINS complex subunit 2; qRT-PCR, quantitative reverse transcription–polymerase chain reaction; CSMA, cell spot microarray; siRNA, small interfering RNA

Address all correspondence to: Olli Kallioniemi, MD, PhD, or Juha K. Rantala, MSc, Medical Biotechnology, VTT Technical Research Centre of Finland, 20521, Turku, Finland. E-mail: Olli.Kallioniemi@helsinki.fi, Juha.K.Rantala@vtt.fi

[1]This work was partly supported by EU-FP7 projects TIME and GENICA, Academy of Finland Center of Excellence funding, Sigrid Juselius Foundation, and Finnish Cancer Organizations.

[2]This article refers to supplementary materials, which are designated by Tables W1 to W3 and Figures W1 to W6 and are available online at www.neoplasia.com.

Received 13 April 2010; Revised 25 June 2010; Accepted 3 July 2010

DOI 10.1593/neo.10548

Introduction

The dysregulated gene expression patterns supporting growth and survival of cancer cells and maintenance of sustained ploidy changes can involve various cellular processes. Genome-wide expression microarray studies have revealed that the biological and clinical heterogeneity of breast cancers can be partly explained by information embedded within complex, but ordered transcriptional architecture [1] that influences the biochemical and behavioral properties of tumors. These profiles can be used for improved disease subtyping, patient prognosis, and disease treatment [2]. Recent advances in the molecular profiling of tumors have revealed a multitude of genes whose expression levels in primary tumors correlate strongly with the probability of metastasis and disease progression [2–4]. In particular, genes that are involved in the cellular responses to DNA damage, cell cycle, and differentiation have been associated with cancer initiation and metastasis [5–8]. An important link between these processes and a driver of cancer proliferation, invasion, and metastasis is chromosomal instability. Chromosomal instability increases genetic variability helping cancer cells to adapt to different environmental challenges and therapies. By acquiring genomic alterations, cancer cells become independent of normal regulatory cell processes and environmental stimuli. Changes in the chromosomal content and gene expression signatures persist also in cancer-derived model cell lines [9]. Excessive genetic damage is, however, detrimental, and many chemotherapeutic agents are active because of the inability of cancer cells to sustain DNA damage and repair the genetic defects.

Mechanistically aneuploidy can arise by multiple mechanisms, such as nondisjunction of chromosomes during mitosis, chromosomal breakage leading to DNA loss or gain, or by generation of multinucleation and polyploidy. Here, we focused on the mechanisms of polyploidy formation and carried out a RNAi screen to identify genes involved in the regulation of coordinated cell division and maintenance of ploidy level in cancer cells. We used our recently developed cell spot microarray (CSMA) method to analyze RNAi effects causing sustained ploidy changes in MDA-MB-231 breast cancer cells. Silencing of 177 of the 5760 human genes included in the analysis was found to induce impaired cellular effects, causing either sustained polyploidy or polyploidy with subsequent induction of apoptosis. Bioinformatic comparison of clinical gene expression profiling coupled with analysis of clinicopathologic parameters and gene copy number analysis of breast cancer cell lines and primary breast tumors was then applied to study the clinical significance of the candidate genes *in vivo*. DNA replication complex GINS protein 2, GINS2, was identified as the top candidate based on both RNAi effects and clinicopathologic associations. Silencing of GINS2 led to increased fraction of polyploid cells and accumulation of endogenous DNA damage in the MDA-MB-231 cells and inhibition of cell growth and viability in all subsequently analyzed breast cancer cell lines.

Materials and Methods

Preparation of CSMAs

Small interfering RNA (siRNA) libraries for CSMA printing were prepared with a Hamilton STAR liquid handling robot (Hamilton Robotics, Bonaduz, Switzerland) by mixing for each sample 5 μl of 1.67 μM siRNA aliquots in 384-well microplates with 0.8 μl of siLentFect (Bio-Rad, Hercules, CA) transfection reagent and 0.2 μl of OptiMEM I (Gibco, Carlsbad, CA). Solutions were incubated for 20 minutes at room temperature and mixed with 2 μl of growth factor–reduced Matrigel (BD Biosciences, Franklin Lakes, NJ) and 2 μl of ice-cold OptiMEM I supplemented with 65 mM sucrose, snap-frozen at −80°C, and stored at −20°C. A prevalidated siRNA for CD9 was used for validation of the transfection efficacy of MDA-MB-231 cells on CSMAs (SI02777187; Qiagen, Hilden, Germany). A Qiagen druggable genome siRNA library version 1.0 with 2 siRNA constructs against 5760 genes was used for the primary analysis. For validation experiments two prevalidated siRNA constructs (Qiagen) were used to target each candidate (Table W2). Arrays were printed on untreated polystyrene microplates with four large rectangular wells (Nunc, Rochester, NY) using a Genetix QArray2 (Genetix Ltd, New Milton, UK) microarray printer with 200-μm solid tip pins (Point Technologies, Alajuela, Costa Rica). For the CSMA experiments with MDA-MB-231 breast cancer cells, cells were grown to 80% confluence on 10-cm culture dishes and dissociated with HyQtase (HyClone, Logan, UT) treatment for 5 minutes. After dissociation, cells were suspended back to the conditioned culture medium and dispersed on the array wells as a uniform cell suspension. A total of 3×10^6 cells in 4.5 ml of medium were added to each array well and were allowed to adhere at +37°C and 5% CO_2 for 20 minutes. After adhesion, all unadhered cells were washed off from the well, and 4.5 ml of fresh culture medium was added to each well. Cells were then transfected on the arrays for 7 days for primary analysis, with medium changed after 96 hours. In the secondary analyses, cells were transfected on the arrays for 96 hours before immunostaining and analysis.

Cell Culture and Cell Lines

HCC-1937, MCF-10A, MDA-MB-231, and T-47D used in the experiments were obtained from ATCC (Rockville, MD). JIMT-1 breast carcinoma cells were obtained from DSMZ (German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany). HCC-1937 cells were grown in RPMI-1640 (Gibco) supplemented with 10% FBS (Gibco), 10 μg/ml penicillin and streptomycin (Sigma, St Louis, MO), and 2 mM L-glutamine (Sigma). MCF-10A cells were grown in mammary epithelial basal medium (Lonza, Basel, Switzerland) supplemented with SingleQuots additives (Lonza) and 100 ng/ml cholera toxin. MDA-MB-231 cells were grown in low-glucose Dulbecco's modified Eagle medium (Gibco) supplemented with 10% FBS (Gibco), 10 μg/ml penicillin and streptomycin (Sigma), 2 mM L-glutamine (Sigma), and 1% MEM nonessential amino acids (Sigma). JIMT-1 cells were grown in 50:50 mixture of high-glucose Dulbecco's modified Eagle medium (Gibco)–RPMI-1640 (Gibco) supplemented with 10% FBS (Gibco), 10 μg/ml penicillin and streptomycin (Sigma), and 2 mM L-glutamine (Sigma). Cells were maintained in a state of logarithmic growth, and experiments were performed between passages 3 and 20 after first thaw.

Bioinformatic Analysis of Expression of Candidate Genes in Clinical Breast Cancer Samples

Analysis of clinical expression profile of the candidate genes was performed with meta-analysis of a previously published gene expression profiling of 251 human breast tumors hybridized to both Affymetrix U133A and U133B human GeneChips ([10], GEO accession number GSE3494) as previously described [11,12]. Briefly, data were preprocessed using R (R Development core team) and the RMA method implemented in the Bioconductor package affy. Affymetrix probes were mapped directly to Ensembl gene IDs during preprocessing. Data for genes appearing on both chip types were combined by calculating their medians from both chips. ERBB2, Ki-67, and PCNA status was

estimated from the expression data themselves. Tumors were classified into defined breast cancer subtypes according to a previously described molecular descriptor [13]. The Mann-Whitney *U* test was used to test whether the medians of two category phenotypes are statistically significantly different. Analysis of relative median expression level of candidate genes in the GeneSapiens transcriptomics database was performed as described previously [12].

Gene Copy Number Analysis

Genome-wide DNA copy number analysis of MDA-MB-231, MCF-7, and T-47D cells was based on a previous analysis (GSE15477) of the cell lines using human genome Comparative Genome Hybridization (CGH) 44A and 44B oligo aCGH microarrays (Agilent Technologies, Palo Alto, CA). Array-based CGH validation analysis of the T-47D cells was performed using Agilent 244K oligonucleotide microarrays according to the direct method of the June 2006, version 4 protocol (Agilent Technologies). Female genomic DNA (Promega, Madison, WI) was used as reference. Briefly, 1 μg of digested and purified sample and reference DNA was labeled with Cy5-dUTP and Cy3-dUTP (Perkin-Elmer, Wellesley, MA), respectively, according to the protocol. Labeled cell line and reference samples were pooled and hybridized onto an array. After hybridization, arrays were washed and scanned with a laser confocal scanner (Agilent Technologies). Signal intensities were extracted using the Feature Extraction software (Agilent Technologies), and the CGH Analytics (Agilent Technologies) was used for data analysis and visualization GEO accession no. GSE22547.

Copy number changes for GINS2 (Agilent ID: A16P20552751) from 178 primary breast cancer cases were extracted from Hu-244A CGH microarrays (Agilent Technologies; unpublished data). The tumors are part of a cohort of 212 primary breast cancer cases sequentially collected at Oslo University Hospital Ullevål, Norway, from 1990 to 1994 with an observation time of 12 to 16 years [14]. The samples were profiled by standard protocol [15] without prelabeling amplification step. Scanned microarray images were read and analyzed with Feature Extraction v9.5 (Agilent Technologies) with protocols (CGH-v4_95_Feb07 and CGH-v4 91 2) for aCGH preprocessing, which included linear normalization. Data were segmented using the PCF (Piecewise Constant Fit) algorithm [16] with settings $K_{\min} = 5$ and $\gamma = 25$. Aberrations were scored with a threshold of 0.3; gain > 0.3 and loss < −0.3. Statistical association of copy number changes for GINS2 and survival were performed in SPSS 16.0 (SPSS, Inc, Chicago, IL).

Immunofluorescence Staining

Immunofluorescence staining of the CSMAs was performed using standard procedures. Cells on arrays were fixed with 2% paraformaldehyde solution for 15 minutes and permeabilized with 0.3% Triton X-100 in PBS for 15 minutes, and the background was blocked with 2% BSA in PBS for 1 hour before staining with primary and secondary antibodies. Before addition of the primary antibody, the arrays were rinsed with dH_2O and air-dried. A PAP-pen (Sigma) was used to line the arrays with a hydrophobic lining to reduce antibody/staining solution consumption. Primary antibodies for CD9 (1:250, rabbit anti-CD9; Santa Cruz Biotechnology, Santa Cruz, CA), cleaved PARP (cPARP; 1:300, mouse anti-cPARP; Cell Signaling Technologies, Danvers, MA), and γ-H2Ax (1:300, rabbit anti–γ-H2Ax; Abcam, Cambridge, MA) were diluted in the blocking buffer and incubated for 60 minutes at room temperature at 80 μl per array. Secondary labeling antibodies goat–antimouse and donkey–antirabbit conjugated with Alexa 488 and 647 dyes (1:300; Molecular Probes, Invitrogen, Carlsbad, CA) were diluted in blocking buffer and incubated for 60 minutes at room temperature. Then, 1 μg/ml 4′,6-diamidino-2-phenylindole (DAPI; Invitrogen) and 0.1 μM phalloidin–Alexa488 were added to secondary antibody solution for DNA and F-actin staining. After secondary labeling, CSMAs plates were rinsed with dH_2O, air-dried, and stored protected from light for imaging.

For immunofluorescence staining of GINS2 and Ki-67, MDA-MB-231 cells were cultured on coverslips and stained using the same protocol as the CSMAs. Primary antibodies for GINS2 (1:200, chicken anti-GINS2; Sigma) and Ki-67 (1:300, rabbit anti-Ki67; Abcam) were diluted in 2% BSA-PBS blocking buffer and incubated for 60 minutes at room temperature. Secondary labeling antibodies goat–anti-chicken and donkey–anti-rabbit conjugated with Alexa488 and 647 dyes (1:300; Molecular Probes, Invitrogen) were diluted in blocking buffer and incubated for 60 minutes at room temperature. Then, 1 μg/ml DAPI (Invitrogen) was added to secondary antibody solution for DNA staining. After secondary labeling, cells were mounted with ProLongGold mounting medium (Invitrogen). Imaging and image-based cytometry analysis was performed with scanR microscope using 20× and 60× objectives.

In the secondary CSMA analysis, the cells were pulse labeled with 5-ethynyl-2′-deoxyuridine (EdU) for 1 hour at +37°C (5 μM; Invitrogen) before fixing and immunofluorescent staining. Alexa488 EdU detection kit (Invitrogen) was used for EdU staining according to the manufacturer's instructions. For the analysis, each array spot was imaged using 20× objective, and nuclear intensities of EdU, cPARP, and γ-H2Ax were measured against the DNA counterstaining with the scanR image analysis software. To evaluate the significance of the signal distribution of the targeting siRNAs, a *z* score against the mean signals, and SD of the control samples was calculated.

Reverse Transcription–Polymerase Chain Reaction and Western Blot Analysis

For quantitative reverse transcription–polymerase chain reaction (qRT-PCR) analysis, the total cellular RNA was isolated using TRIzol reagent (Invitrogen). For complementary DNA (cDNA) synthesis, 200 ng of total RNA was reverse-transcribed with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Carlsbad, CA). The cDNA was diluted 1:10, and TaqMan qRT-PCR analysis was performed with an Applied Biosystems 7900HT instrument, using primers designed for GINS2 by the Universal Probe Library Assay Design Center (Roche, Indianapolis, IN): forward 5′-TCTGGACAAGATC-TACCTCATCG-3′ and reverse 5′-CACTTCCACGGGTAAACCA-3′. The fluorescent TaqMan probes were obtained from Roche Human Probe Library. Results were analyzed using SDS 2.3 and RQ manager software (Applied Biosystems), and the relative expression of messenger RNA (mRNA) was determined using GAPDH as an endogenous control. The data from two separate biological experiments with triplicate samples were combined. For Western blot analysis, aliquots of total cell lysates were fractionated on sodium dodecyl sulfate–polyacrylamide gels and transferred to Whatman Protran nitrocellulose membrane (Whatman, Inc, Piscataway, NJ). The filters were blocked against nonspecific binding using 5% skim milk. Membranes were probed with antibodies overnight at +4°C (GINS2, 1:500; CD9, 1:1000; Santa Cruz Biotechnology). Total protein loading was confirmed by probing the same filter with a specific antibody for tubulin (1:5000; Abcam). Signals were revealed by incubating the filters with secondary antibody Alexa Fluor 680 antimouse IgG (Invitrogen)

and scanning the filters with an Odyssey Licor infrared scanner (LI-COR Biosciences, Lincoln, NE).

Validation Experiments with siRNAs: Cell Viability Assays and Live-Cell Imaging

In validation of target silencing and assaying the effect of GINS2 siRNAs (Qiagen SI02653056, SI02653581, SI04439757, SI04439764) on the growth of breast cancer cells, the cells were cultured on clear-bottom 96-well (2000 cells per well) and 12-well (2×10^5 cells per well) plates and transfected with 10 nM siRNA constructs (Qiagen) using siLentFect (Bio-Rad). Cell viability was assayed with CellTiter-Blue cell viability assay (Promega). About 10 μl of CellTiter-Blue diluted with 10 μl of OptiMEM I (Gibco) medium without supplements was added to each 96-well containing 100 μl of medium and cells. Reagent was incubated at +37°C for 4 hours followed with 2 hours of stabilization at room temperature before analysis. Fluorescence signal (excitation, 560 nm; emission, 590 nm) reflecting the relative number of viable cells per well was measured with EnVision fluorescence plate reader (Perkin-Elmer). Data from four replicate wells were combined for analysis.

Time-lapse imaging of cells transfected on CSMAs was performed with Incucyte HD live-cell imaging microscope using 20× objective (Essen Instruments, Ann Arbor, MI). Images were acquired every 2 hours for 7 days.

Results

RNAi-Induced Polyploidy Assay

To characterize genes involved in the maintenance of coordinated cell division of aneuploid cancer cells, we performed a RNAi screen with human MDA-MB-231 breast cancer cells using our recently developed CSMA technique allowing production of high-density siRNA reverse transfection cell microarrays. In the method, individual targeting siRNAs and negative control siRNAs are printed on an Society for Biomolecular Sciences–sized microplate with a hydrophobic polystyrene surface in an arrayed configuration for the lipid based reverse transfection of siRNAs to the cells [17]. To test efficacy of the MDA-MB-231 transfection on the CSMAs before the screening, we performed a Western blot and antibody-based immunofluorescent analysis of the cells transfected on CSMAs for 48 and 72 hours, respectively, using a prevalidated siRNA for CD9. On the basis of microscopic immunofluorescent quantitation of whole-cell area CD9 staining against DNA counterstaining of the cells after 72 hours of transfection, a mean CSMA spot level silencing efficacy of 75% in 384 random-order printed CD9 spots was achieved in comparison to an equal number of control siRNA spots (CD9 siRNA coefficient of variation 15.17% *vs* control siRNA coefficient of variation 14.11%, $P < .0001$; Figure 1*B*). In addition, CD9 silencing was also found to be accompanied with a 40% reduction in number of cells per spot (31 ± 12 *vs* 58 ± 6, $P < .0001$; Figure 1*B*). With Western blot analysis of cells transfected on identical arrays with 384 replicate spots (Ø 200 μm) of the CD9 siRNA and the control siRNA, more than 80% silencing efficacy was measured with normalization against β-tubulin (Figure 1*B*). Verifying the microscopy detected reduction of cell numbers on CSMA spots after silencing of CD9, in the Western blot analysis, the total protein amount detected by probing for tubulin was decreased by 25% after 48 hours of transfection.

For the RNAi analysis of induction of polyploidy, we focused on the identification of sustained polyploidy phenotypes in the MDA-MB-231 cells after 7 days of transfection. MDA-MB-231 is an aneuploid cell line with a modal chromosomal number of 54 (including 22 numerical changes) and has multiple structurally rearranged chromosomes [9]. In the CSMA analysis, we used a druggable genome siRNA library consisting of two individual siRNA constructs targeting 5760 human genes. Individual targeting siRNAs and negative control siRNAs were printed on the microplate surface in a randomized order, resulting in an array with a total of 15,552 cell spots, each associated with a single siRNA or control reagent. Cells were seeded onto the CSMA and allowed to transfect for 7 days (Figure 1*A*). The culture medium was changed after 4 days to prevent cell starvation. After transfection, cells were fixed and stained with a DNA binding dye (DAPI) and with fluorescently labeled phalloidin for detection of nuclear and cellular shape and size (Figure 1*C*).

To analyze the array and detect polyploid cells, we used automated fluorescence microscopy imaging of the array using 20× objective and visual nuclear morphology–based identification of polyploid cells after automated quantification of total cell number on the array spots. We calculated the fraction of polyploid cells per spot (polyploid cell index, or PCI) by dividing the number of detected polyploid cells by the total number of cells per spot. The identified phenotypes were further classified as sustained or apoptotic polyploidy on basis of the nuclear morphology (symmetric round nucleus *vs* fragmented nucleus; Figure 2*A*). On the basis of the prevailing phenotype, the candidate genes inducing a greater than 20% penetrance of polyploid cells per spot (PCI = 0.2) were divided into these two categories (Figure 2*A*). With the threshold of 20% polyploid cells per spot and two independent analyses of the resulting phenotypes, silencing of 177 genes were considered positive (3% hit rate), with 28 genes scoring with both of the used siRNAs. On the basis of the phenotypic stratification, 134 of the candidate genes were considered sustained polyploidy phenotypes and 43 as apoptotic polyploidy (Table W1). The PCI threshold used was determined so that the spontaneously occurring polyploid cell fraction within the MDA-MB-231 cell population was clearly lower. Analysis of the mean PCI in the group of 375 control siRNA positions was 0.022 (2.2%). Besides automated analysis of cell numbers, a total of 15,552 composite images produced in the screen were visually inspected for the presence of polyploidy in two independent analyses.

Data on cell numbers per spot for all array positions (targeting and control siRNAs) approximately fitted a normal distribution, but for siRNAs inducing polyploidy, a significantly altered cell number distribution was identified (Figure 2*B*). On the basis of visual validation of cell numbers in these positions, we found that in spots with a PCI ≥ 0.2, the mean cell number distinctly decreased from 45 cells per the segmented 200-μm spot area in typical positions to a mean of 23 cells per spot.

Functional Annotation

To understand the underlying biological processes inducing polyploidy in the analysis, we performed Gene Ontology (GO) categorization of the genes whose inhibition was associated with increased polyploidy in MDA-MB-231 cells. Gene GO term enrichment analysis of the 177 genes using DAVID2008 Functional Annotation Tool [18,19], focusing on associations with biological processes and cellular component, indicated that response to DNA damage stimulus (21 genes), cell cycle (34 genes), DNA repair, response to stress, DNA replication, and cell cycle process were among the main ($P < .0001$) biological processes linked to the polyploidy phenotypes in our cell array screen (Figure 2*C* and Table W2). On basis of the cellular component associations, nuclear part (34 genes), microtubule (12 genes), and spindle were the most enriched terms associated with the gene products ($P < .0001$).

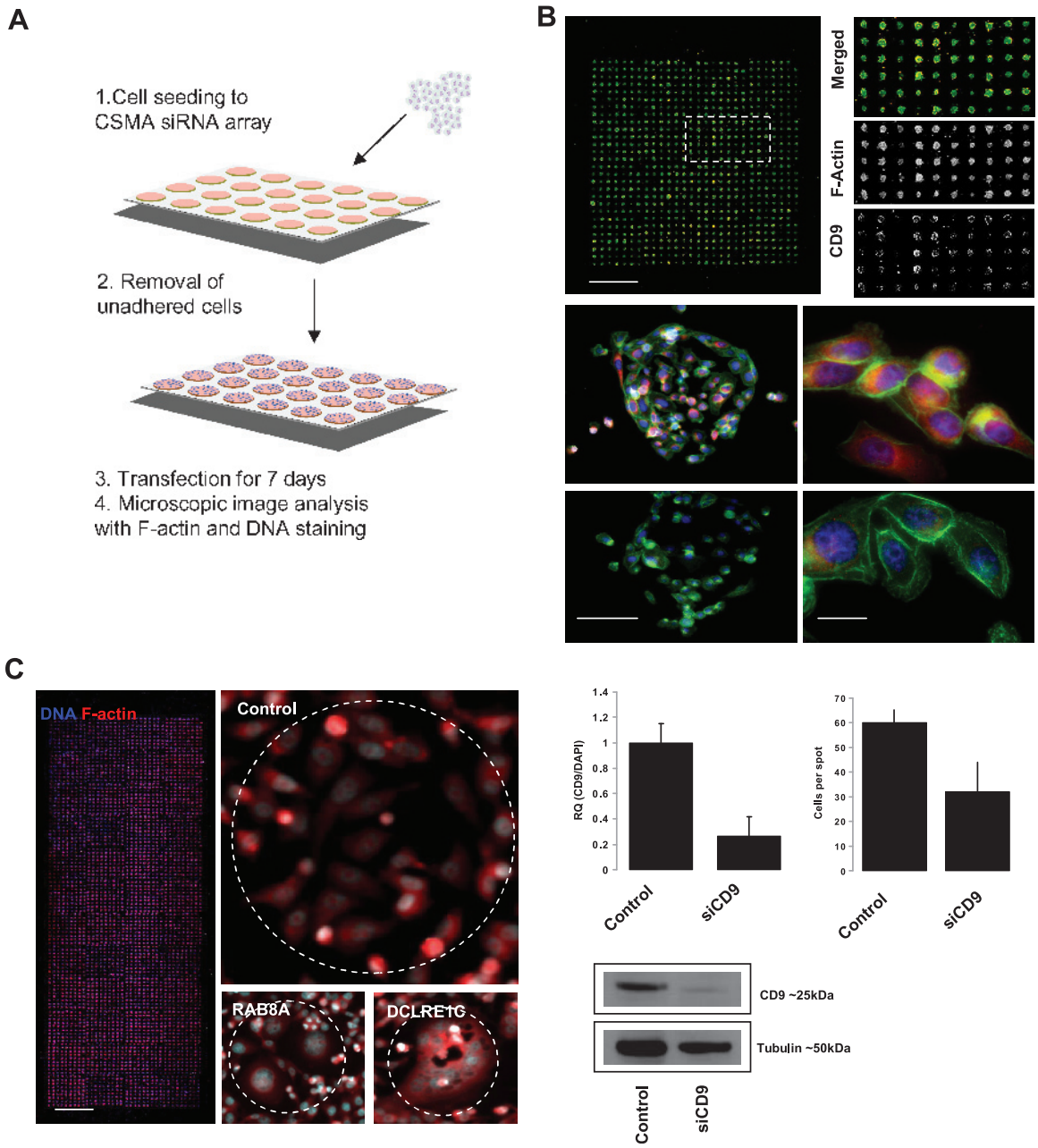


Figure 1. CSMA RNAi analysis. (A) Schematic of procedure for cell seeding and assaying MDA-MB-231 cells on CSMAs. (B) A color composite image of laser microarray scanned view of a CSMA with MDA-MB-231 cells cultured for 72 hours in 384 random-order printed replicate spots of CD9 and negative control siRNA spots (F-actin = green, CD9 = red). Scale bar, 4 mm. Objective: 20× microscopic image of a negative control and CD9 siRNA spot. Scale bar, 100 μm. Objective: 63× microscopic images of a negative control and CD9 siRNA–transfected MDA-MB-231 cells stained for DNA (blue), F-actin (green), and CD9 (red). Scale bar, 10 μm. Quantification of immunofluorescence analysis of CD9 silencing and cell numbers on CSMA spots after 72 hours of CD9 silencing. Lower panel: Immunoblot analysis of CD9 and α-tubulin protein levels in MDA-MB-231 cells after 48 hours of siRNA knockdown on CSMAs with an equal number of replicate spots of control siRNA or siCD9. (C) Low-resolution microarray-scanned fluorescence image of a CSMA with 3888 spots stained for DNA (blue) and F-actin (red). Scale bar, 5 mm. Microscopic images of MDA-MB-231 cultured for 72 hours on CSMA spots demonstrate detection of changed cell morphology and DNA content on the basis of F-actin and DNA staining. Scale bar, 100 μm.

The two different gene sets whose knockdown was leading to either apoptotic or sustained polyploidy had similar GO term associations. The 134 genes, whose knockdown induced sustained polyploidy, fell into GO categories such as DNA repair ($P = 3.2e{-}7$, $n = 14$ genes), cell cycle ($P = 3.8e{-}6$, $n = 23$), developmental process ($P = 6.5e{-}4$, $n = 44$), and regulation of apoptosis ($P = 5.0e{-}3$, $n = 12$). The 43 genes whose silencing was correlated with apoptotic polyploid cells were associated with GO categories such as DNA metabolic process ($P = 1.2e{-}3$, $n = 9$), DNA replication ($P = 3.6e{-}3$, $n = 5$), and cell cycle ($P = 5.7e{-}5$, $n = 11$).

Validation Analysis

Supporting the current understanding that the impaired cellular responses to DNA damage are among the most important molecular

processes contributing to development and maintenance of aneuploidy in cancer cells, these biological processes were also most enriched among the identified candidate genes in our CSMA analysis. To confirm the primary screening results and to gain more insights to the underlying biological processes inducing polyploidy, we performed a secondary CSMA screening of 30 of the candidate genes associated with GO term GO:0006974 (response to DNA damage stimulus) and/or GO:0006281 (DNA repair) (Table W3). A siRNA library with two functionally prevalidated siRNA constructs for each target gene was acquired for the analysis (Table W3).

To confirm the polyploidy induction by inhibition of these genes, MDA-MB-231 cells transfected for 96 hours on the secondary CSMA with three technical replicate spots of each siRNA and 304 replicates of a negative control siRNA were analyzed using the same phenotypic stratification as in the primary analysis. Of the 30 candidates, 27 scored positive for induction polyploidy with both two siRNAs with at least two-thirds of the analyzed siRNA replicate spots (Figure 3*A*). From the candidates, *DKC1* and *LIG3* scored negative for induction of polyploidy with all analyzed replicates and *MSH4* scored positive with one-sixth of all the siRNA replicates.

In addition to the visual phenotypic polyploidy analysis, we examined the effects of siRNA-mediated knockdown of the candidate genes on cell proliferation, induction of apoptosis, and formation of endogenous DNA damage with a multiplexed high-content immunofluorescence assay. To evaluate the effect of candidate gene silencing on cell proliferation, we applied a fluorescence-based EdU incorporation assay for detection of active DNA synthesis as measure of cell proliferation. To detect apoptotic cells, we used an antibody against cleaved PARP to measure induction of apoptosis and an antibody against γ-H2Ax to expose the potential roles of the genes in the maintenance of DNA integrity (Figure 3, *A* and *B*). siRNAs against many of the candidate genes (16/30) significantly (z score < −2) altered the proliferation of

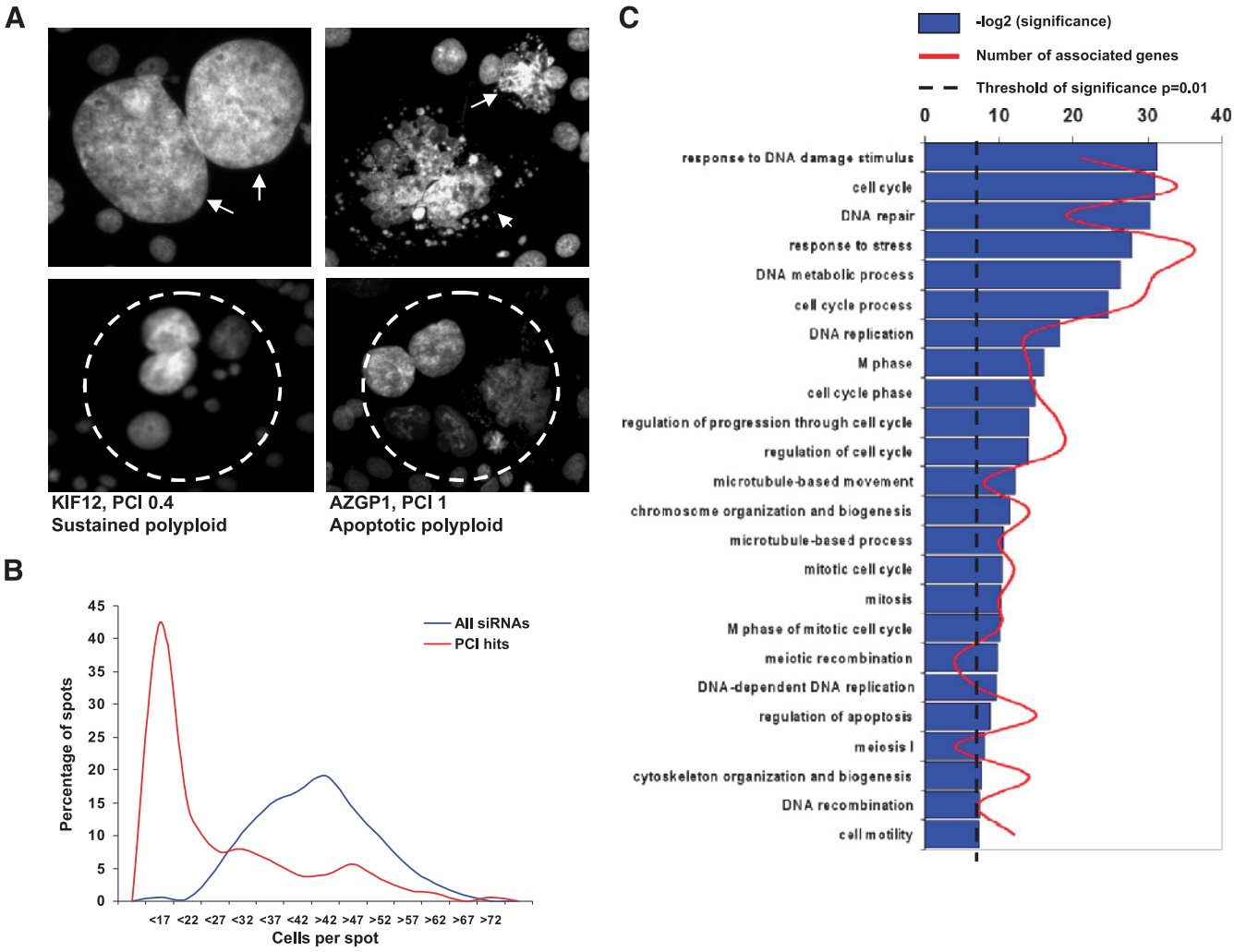
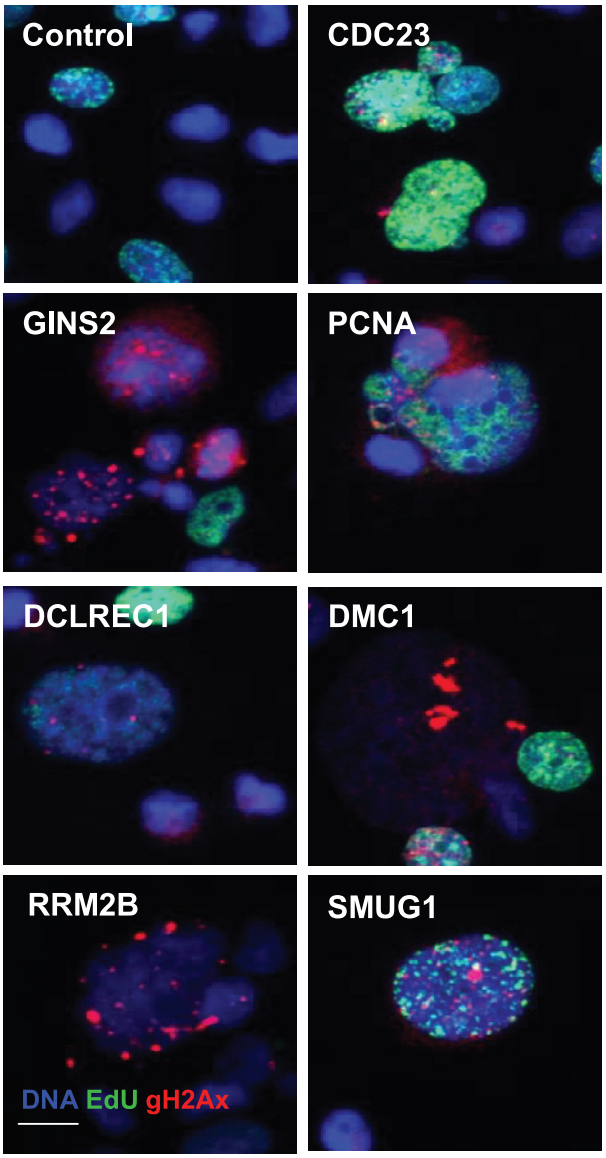


Figure 2. Analysis of polyploidy induction in MDA-MB-231 cells. (A) Images of representative nuclear phenotypes used for stratification of identified polyploid cell phenotype. Round, smooth polyploid cells were considered sustained polyploidy, and fragmented nucleus containing cells were considered apoptotic polyploid cells. Images of MDA-MB-231 cells after 7 days of siRNA knockdown on KIF12 and AZGP1 siRNA spots. Positions that gave an elevated polyploid cell index (PCI ≥ 0.2) from the primary screen were divided into categories of sustained (KIF12) or apoptotic (AZGP1) polyploidy phenotypes. (B) A line graph of cell number distribution across all siRNA spots on the CSMA (blue) and siRNA hits inducing polyploidy (red). A distinct reduction of cell number on spots was associated with the induced polyploidy phenotypes. (C) Gene ontology analysis of the candidate genes. Number of genes associated with the enriched GO categories and significance of enrichment ($-\log_2 P$ value). Threshold of significance is $P = .01$.

A



B

Gene	Polyploidy	EdU	cleaved PARP	gamma-H2Ax
BARD1	BARD1		BARD1	BARD1
CDC23	CDC23			
CRY2	CRY2			
DCLRE1C	DCLRE1C	DCLRE1C		
DDIT3	DDIT3			
DKC1		DKC1	DKC1	
DMC1	DMC1		DMC1	DMC1
FANCD2	FANCD2	FANCD2	FANCD2	FANCD2
GINS2	GINS2	GINS2	GINS2	GINS2
HDAC3	HDAC3	HDAC3	HDAC3	HDAC3
HDAC4	HDAC4	HDAC4	HDAC4	HDAC4
HDAC6	HDAC6			
HUS1	HUS1	HUS1		
LIG1	LIG1	LIG1		
LIG3		LIG3		
MCM3	MCM3		MCM3	MCM3
MNAT1	MNAT1			
MSH4		MSH4		
PCNA	PCNA	PCNA	PCNA	
PMS1	PMS1	PMS1		
PMS2	PMS2	PMS2		
POLB	POLB	POLB	POLB	POLB
PSMC4	PSMC4		PSMC4	PSMC4
RAD50	RAD50			
RNF14	RNF14			RNF14
RRM2B	RRM2B	RRM2B		
RSF1	RSF1			
SIRT6	SIRT6			
SMUG1	SMUG1		SMUG1	SMUG1
TYMS	TYMS	TYMS		

Figure 3. Functional validation of a subset of primary screening candidates. (A) Representative magnified images of polyploid phenotypes induced by the candidate gene knockdowns. Cells were stained for DNA (blue), γ-H2Ax (red), cleaved PARP (magenta), and EdU (green) for analysis of nuclear morphology, maintenance of DNA integrity, induction of apoptosis, and cell proliferation, respectively. (B) Summary of the assay results: polyploidy induction, EdU incorporation (cell proliferation), induction of apoptosis, and γ-H2Ax assay. Candidate genes scored in each assay with both used siRNAs are marked in red.

the cells as measured with the decrease in the EdU incorporation compared with the control siRNA. Inhibition of multiple genes from the list also showed significant functional defects in the other assays; 12 of 30 genes were associated with induction of apoptosis (cPARP staining z score > 2) and 11 of 30 of the candidates affected maintenance of DNA integrity of the cells (γ-H2Ax formation z score > 2). In summary, 5 of the 30 genes scored with both two siRNAs in all the four assays, and 20 in two or more assays (Figure 3*B*). Thus, we concluded that our gene list is highly enriched with genes that affect the maintenance of ploidy and participate in cellular processes such as the DNA damage responses or DNA repair mechanisms contributing to the active cell growth of aneuploid cancer cells.

Clinical Significance

To study gene expression patterns of the validated polyploidy-inducing genes in clinical breast cancers and to evaluate the correlation between the gene expression data and clinicopathologic profiles in breast cancer samples, we applied meta-analysis of the genes included in the validation experiments in a previously published breast cancer gene expression analysis [10]. Transcript profiles of 251 primary breast tumors were assessed in comparison with clinicopathologic variables: TP53 mutation, Ki-67, PCNA, ERBB2, estrogen receptor, progesterone receptor, and lymph node status; tumor grade; and patient survival (Figure W1). In addition, tumors were divided into previously

defined cancer subtypes: normal, luminal A, luminal B, basal type, and ERBB2-positive [16].

By using unsupervised hierarchical clustering of gene expression data of the candidate genes, a subgroup of the basal-type tumors associated with the aggressive clinicopathologic characteristics formed a separate cluster from the rest. The most variably expressed genes distinguishing the aggressive (basal) type of the tumors, included 33% (10/30) of the genes: *PCNA*, *BARD1*, *LIG1*, *MCM3*, *DKC1*, *TYMS*, *GINS2*, *POLB*, *RNF14*, and *RAD50* (Figure 4). Many of the genes highly expressed in these tumors have been previously associated with breast cancer and shown to play a role in cell cycle and cell proliferation, such as *BARD1*, *LIG1*, *MCM3*, and *PCNA*. However, also a less well-characterized candidate gene, *GINS2* was included among the gene set displaying the highest association with the aggressive characteristics of the analyzed breast cancers.

Effect of GINS2 on Cell Proliferation and Cell Cycle Progression

GINS2 was among the identified 28 double siRNA hits inducing a sustained polyploidy phenotype in the primary analysis. In the secondary validation experiments, *GINS2* was found to have a significant effect on cell proliferation, survival, and maintenance of genomic integrity in addition to the induction of polyploidy. Within the bioinformatics profiling, *GINS2* was identified to associate with the clinicopathologic characteristics for aggressive basal-type breast cancers and in analysis of the GeneSapiens database [12]; it showed an elevated expression (≥three-fold, $P = .008$, t test = 3.59, $N = 768$) in most breast cancer samples in comparison to normal breast tissues. Moreover, *GINS2* is located at chromosomal region 16q24, which has been previously identified to be frequently upregulated in clinical breast cancers [20]. We validated the *GINS2* mRNA down-regulation in response to RNAi

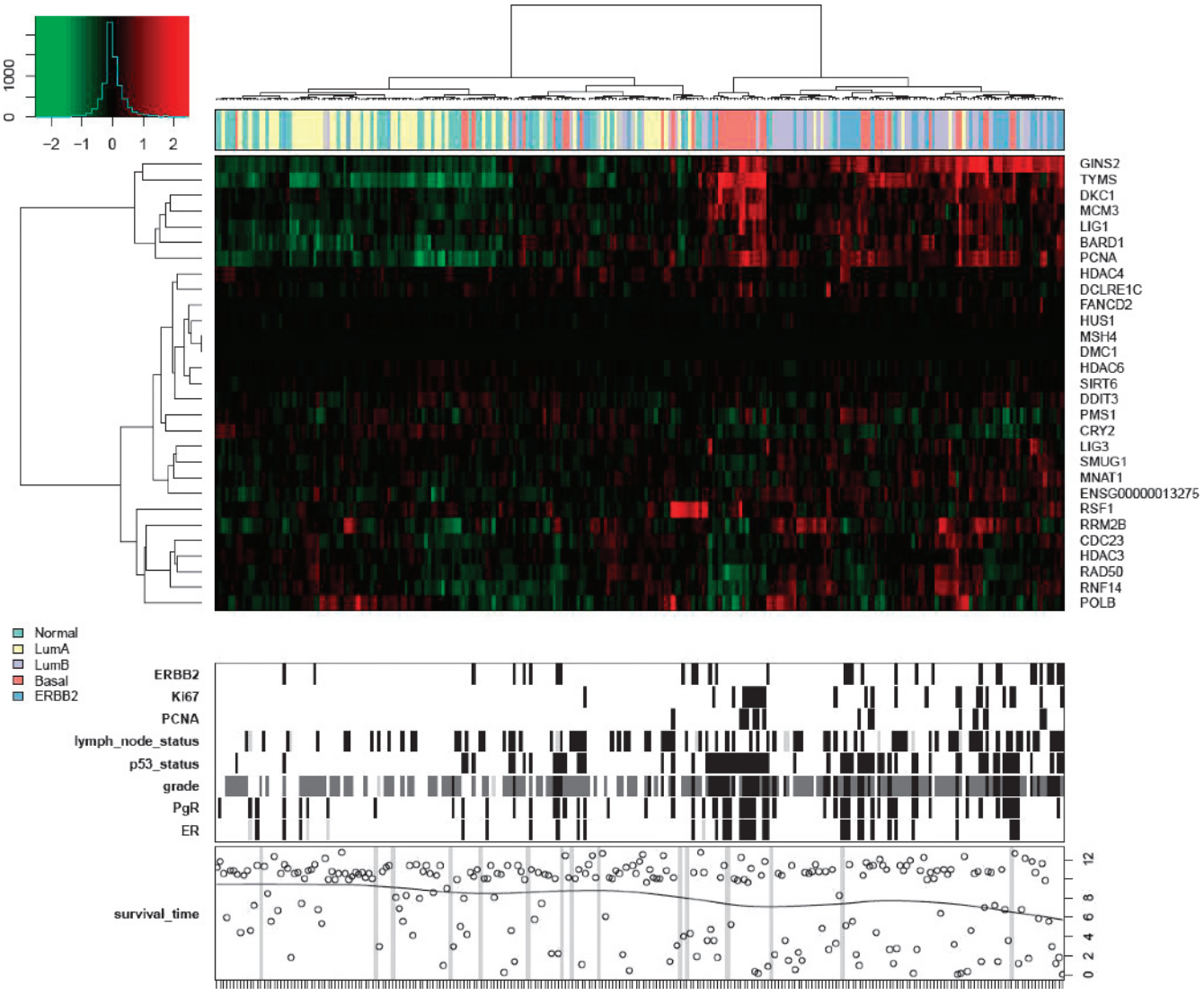


Figure 4. *In silico* transcriptomics analysis of the RNAi hits in clinical breast cancers. Unsupervised hierarchical clustering of the expression level of the 30 candidate primary RNAi hits included in the validation experiments in 251 breast tumors. Each cell in the cluster shows the $\log_2$ expression ratio for the particular gene in separate tumor samples divided by the median expression of that gene in all samples. Red indicates expression above the median; green, below the median. Upper panel: tumor-type classification of each sample. Lower panel: sample status for clinicopathologic parameters: high ERBB2 expression, high Ki-67 expression, high PCNA expression, lymph node positivity, presence of p53 mutation, tumor grade, PgR positivity (black bars), and patient survival.

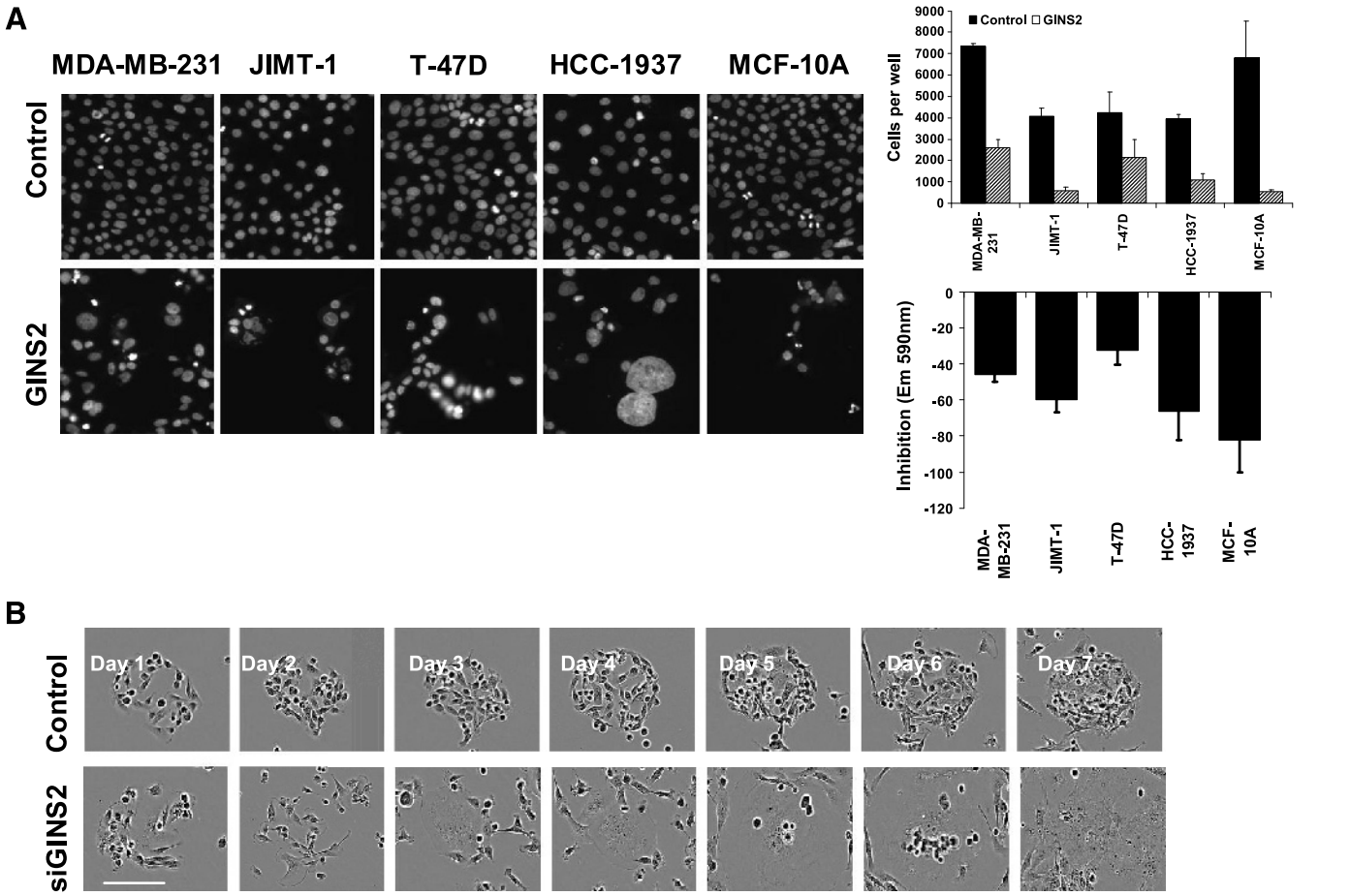


Figure 5. Functional profiling of GINS2. (A) Images of MDA-MB-231, JIMT-1, T-47D, HCC-1937, and MCF-10A cells transfected with GINS2 and negative control siRNA for 120 hours and stained with a DNA binding dye (DAPI). Image-based cytometry analysis and an enzymatic cell viability assay were used to confirm the growth-inhibitory effects of GINS2 inhibition. Error bars indicate SD of two replicate transfections with two siRNA constructs. (B) Time-lapse microscopic image series of MDA-MB-231 cells transfected on CSMA spots for 7 days with GINS2 and control siRNA. Scale bar, 100 μm. (C) Immunofluorescence staining of GINS2 and Ki-67 in MDA-MB-231 cells with 60× magnification. Scale bar, 10 μm. (D) Kaplan-Meier survival plot for disease-specific survival of 178 primary breast cancer samples, classified according to GINS2 copy number status.

for both two siRNAs used in the primary screen and the two additional siRNA constructs used for validation screening in MDA-MB-231 cells by means of qRT-PCR. In addition, Western blot analysis of cells transfected on CSMAs for 5 and 7 days (Figure W2) was used to verify the endured silencing efficacy on CSMA spots. In addition to the verified 75% GINS2 protein level inhibition, a reduction of up to 48% (7 days of transfection) in cell numbers on the CSMAs was detected as decrease in the measured total protein levels based on probing for tubulin. We then compared the mRNA expression level of GINS2 in 15 breast cancer cell lines (Figure W2). From the analyzed cell lines, T-47D cells displayed the highest expression level for GINS2. Results from aCGH analysis of the corresponding cell line indicated that T-47D cells have a heterozygous amplification of GINS2 locus, providing a direct mechanism for the up-regulation through increased DNA copy number (Figure W2). In addition, amplification of GINS2 locus was found on five additional breast cancer cell lines: LY2, BT-20, HCC-1937, HCC-1954, and MDA-MB-468 [16].

To confirm the CSMA results, we compared the GINS2 RNAi effect in four breast cancer cell lines, namely, MDA-MB-231, T-47D, HCC-1937, and JIMT-1, and in a nonmalignant breast epithelial cell line MCF-10A (Figure 5*A*). Reduced cell viability was confirmed using an enzymatic cell viability assay of cells transfected for 120 hours (with two of the validated GINS2 siRNAs), followed by with image-based cytometry analysis of cell cycle effects. Cell viability assays and the automated microscopic image analysis revealed a drastic reduction in the number of cells in response to GINS2 inhibition (Figure 5*A*). Cell cycle analysis of the cells revealed the decrease in the number of cells at G_1-S phases and the increase in the G_2-M and G_0/sub-G_1 phases and polyploid cell fraction (Figure W3). To further clarify the effect of GINS2 knockdown on cellular morphology and cell cycle, we examined MDA-MB-231 cells transfected for 7 days on GINS2 CSMA spots using time-lapse microscopy. A distinct formation of polyploid cells and apoptotic cells was detected on GINS2 inhibition, whereas in control position, cells retained a tight active cell growth through the time lapse (Figure 5*B*).

Clinical Significance of GINS2

To investigate the biological and clinicopathologic significance of the GINS2 copy number levels in breast carcinogenesis and possible association with clinical outcome, we compared the copy number status of GINS2 in an ongoing analysis of clinical primary breast tumors. In previously published gene copy number analyses of clinical

breast cancers, the chromosomal region 16q24 has been identified as frequently lost in human breast tumors [21] but amplified and highly expressed in a subgroup of breast tumors [20]. Within comparison of the aCGH profiles of 178 tumors included in the mined analysis, 6 (3.4%) of 178 tumors showed copy number gain of GINS2, whereas 47 (26.4%) of 178 tumors had lost the same region (Figure W5). Patients with a tumor that had gained a copy of GINS2 had a significant poorer survival than the rest (P = .02) using Kaplan-Meier estimates and log-rank comparisons (Figure 5*D*). Data showed a trend that copy number gain of GINS2 was associated with high grade as measured by Kruskal-Wallis test (P = .57).

To examine the subcellular localization of endogenous GINS2 protein in cancer cells, we performed immunocytochemical analysis of MDA-MB-231 cells using a polyclonal chicken antibody against GINS2 (Figure W4). GINS2 was not detected in the nucleus and only weakly around the nuclear membrane in cytoplasm during G_0-G_1. From G_2 to M phases, GINS2 was detected abundantly surrounding the chromatin during anaphase to telophase (Figure 5*C*). Polyploid cells detected within the parental MDA-MB-231 population stained also strongly positive for GINS2 in the given cell cycle phases.

The distinct staining of actively proliferating cells with GINS2 antibody suggested that GINS2 could potentially be used as a biomarker to determine the growth fraction of a given cancer cell population. To evaluate whether GINS2 staining is comparable with currently used proliferation markers, we performed dual staining of MDA-MB-231 cells for Ki-67 and GINS2. Nuclear staining pattern of the two proteins had a significant correlation coefficient (r = 0.84). The only difference between the staining profiles for these two proteins was the staining of G_1 cells for Ki-67 and not for GINS2 (Figure W6*A*). Because staining of cells not in the active cell cycle phases has been considered as a weakness for clinical use of Ki-67, our results indicate that GINS2 could possibly serve as a better marker for cells destined for cell division (G_2-M) and not detecting nonproliferating cells (G_0-G_1).

To evaluate the correlation of these two genes in clinical cancers, we compared the coexpression pattern of Ki-67 and GINS2 in the GeneSapiens database in 3767 gene expression analyses representing

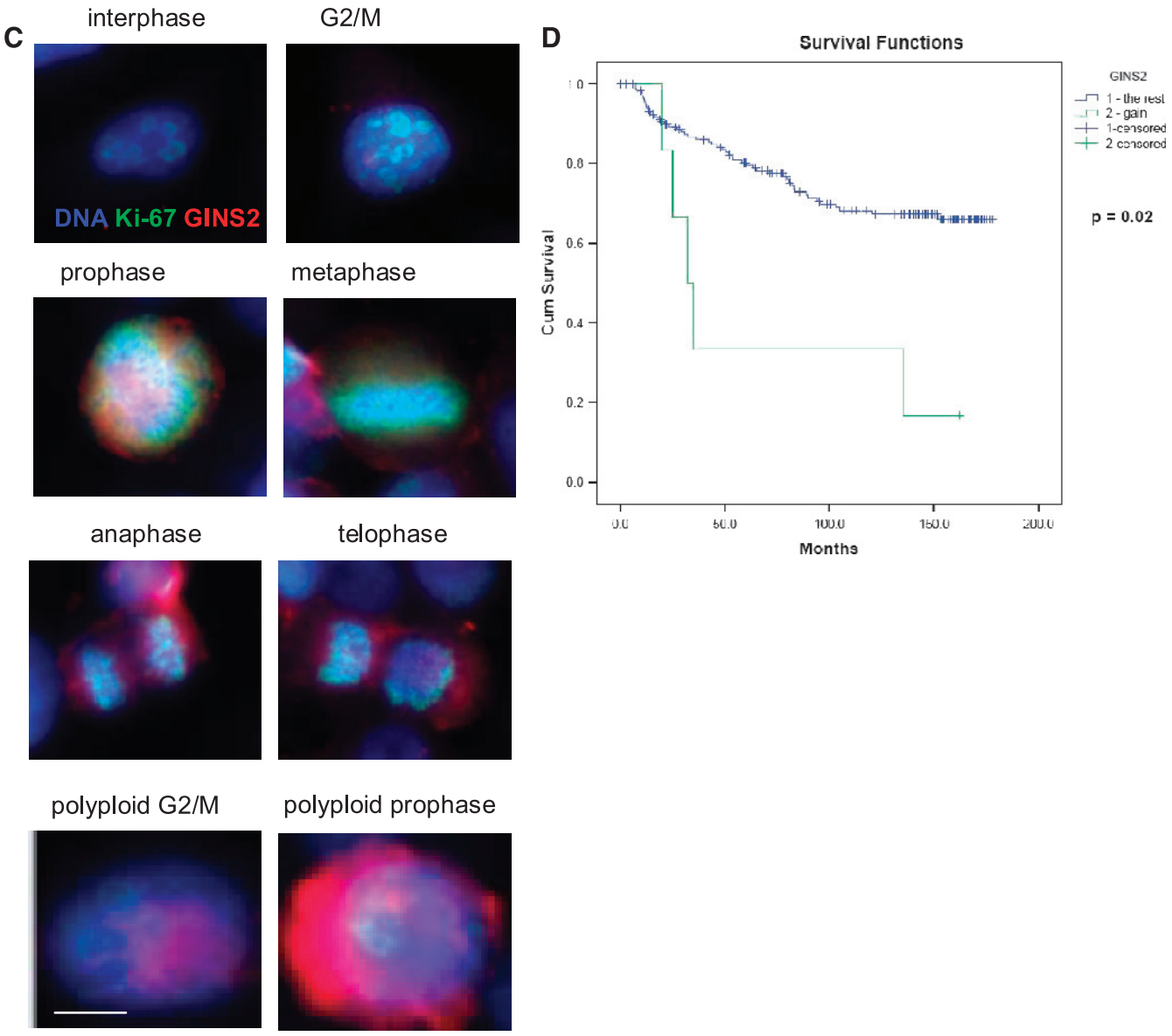


Figure 5. (*continued*).

40 different cancer types. The mRNA coexpression correlation across all the samples was $r = 0.627$ ($P = .01$; Figure W6*B*), indicating a general role for GINS2 in the proliferation of several different cancer types.

Discussion

Aneuploidy and chromosomal instability are common conditions for most epithelial cancer cells, but the relationships between cellular functions allowing growth and survival of aneuploid cells are not clear. Moreover, it is not completely understood whether there is a ploidy-sensing checkpoint in eukaryotic cells and whether cell division of aneuploid cells requires variable expression of specified genes. To address this question, we used RNA interference analysis to identify genes affecting ploidy regulation of cancer cells, followed by GO categorization and *in silico* transcriptomics analysis to identify genes and cellular processes associated with the induction of polyploidy in breast cancer cells.

A collective molecular portrait descriptive of numerical chromosomal heterogeneity in cancer cells has been described to include up-regulation of genes that are associated with increased cell motility and migration, epithelial-mesenchymal transition, and cell cycle processes. Genes whose expression has been shown to correlate negatively with DNA content heterogeneity in cancer cells on the other hand have been associated with, for example, nucleic acid metabolism, regulation of transcription, DNA replication, response to DNA damage stimulus, DNA repair, chromosome organization, and DNA replication initiation [22]. In this study, GO analysis of the distribution of 177 genes identified to functionally correlate with increased polyploidy on silencing indicated an association with highly similar functional categories. DNA damage stimulus, cell cycle processes, DNA repair, response to stress, and DNA replication were the main biological processes associated with the genes increasing the numerical heterogeneity of the chromosomal content. Moreover, within the identified gene list, several genes previously associated with chromosomal heterogeneity and polyploidy were identified. This supports the already-existing assumptions that a compromised expression of genes related to cell cycle processes, DNA damage response, DNA replication, and chromosome condensation is associated with a higher level of chromosomal instability in cancer cells [23–27].

GINS2

GINS2 (GINS complex subunit 2 [Psf2 homolog]) is a member of the tetrameric complex termed *GINS*, composed of GINS1, GINS2, GINS3, and GINS4, which most likely serves as the replicative helicase, unwinding duplex DNA ahead of moving replication forks [28–30]. In studies on mice and yeast, the GINS complex has been shown to associate with the minichromosome maintenance (MCM) 2 to 7 complex and with CDC45, and this complex (CDC45-MCM-2-7-GINS) regulates both the initiation and the progression of DNA replication [31–33]. Later, the GINS complex has been shown to be involved in DNA replication in humans as well [34–38]. However, recent studies suggest that GINS1/2 is also associated with response to replication and DNA damage stress [39–41].

Several recent reports have suggested a role for GINS components in cancer cells. For example, GINS components were found to be overexpressed in aggressive melanoma [41], and GINS1 was identified as an estrogen-regulated target in MCF-7 human breast carcinoma cells [42]. As it has been reported that DNA replication–associated proteins have diverse functions in different cells, for example, in determining centrosome copy numbers, in different phases of development and disease avoidance, GINS has been suggested to have a function in cell division, more precisely in choromosome segregation [43]; however, the role of its components in mammalian cells is not yet clear.

In this study, we found a high level of GINS2 expression in human breast cancers and several breast carcinoma cell lines. We showed also that GINS2 locus is amplified in T-47D breast cancer cell line and a subgroup of clinical breast tumors. Our comparative analysis of different cell types showed that GINS2 is more highly expressed in cancer cells than in nonmalignant breast epithelial cells. Moreover, GINS2 knockdown resulted in growth inhibition and induction of polyploidy in breast cancer cells by the suppression of M-phase progression, indicating that GINS2 impacts, in addition to DNA replication initiation essential for S-phase progression (in GINS complex), cell division, and probably chromosome segregation in human breast carcinoma cells.

Because several prereplicative complex proteins and cell division–related proteins are overexpressed in cancer and reported as useful tumor markers [44,45], we compared the GINS2 mRNA and protein expression with Ki-67 proliferation marker expression. On the basis of the findings, we suggest that GINS2 is a putative biomarker for diagnosis and analysis of progression in breast cancer as well as in several other cancers. As prereplicative complex proteins have also been suggested to have significant therapeutic value [44,46–48], GINS2 might also have use as a drug target because the RNAi results indicate that its inhibition reduces cell proliferation.

In summary, this study provides novel insights into the genetic targets involved in the maintenance of sustained ploidy and causal consequences of the loss-of-function of these genes. Our results can be used as a starting point for formation of hypotheses on events that may impact primary tumor development and suggest regulatory cell cycle components that could be on the basis for dysregulation of cell cycle checkpoints leading to increased gains and losses of chromosomes in cancer cells. The results also provide a foundation for investigating the causative relationships of the identified candidate molecules and the ploidy phenotype, cancer molecular markers, as well as gene expression signatures of clinical cancers.

Acknowledgments

The authors thank R. Mäkelä for his excellent technical assistance.

References

[1] Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, and Chinnaiyan AM (2004). Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* **101**, 9309–9314.

[2] van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536.

[3] Kluger HM, Kluger Y, Gilmore-Hebert M, DiVito K, Chang JT, Rodov S, Mironenko O, Kacinski BM, Perkins AS, and Sapi E (2004). cDNA microarray analysis of invasive and tumorigenic phenotypes in a breast cancer model. *Lab Invest* **84**, 320–331.

[4] Chen JJ, Peck K, Hong TM, Yang SC, Sher YP, Shih JY, Wu R, Cheng JL, Roffler SR, Wu CW, et al. (2001). Global analysis of gene expression in invasion by a lung cancer model. *Cancer Res* **61**, 5223–5230.

[5] Gupta PB, Kuperwasser C, Brunet JP, Ramaswamy S, Kuo WL, Gray JW, Naber SP, and Weinberg RA (2005). The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nat Genet* **37**, 1047–1054.

[6] Gupta GP and Massague J (2006). Cancer metastasis: building a framework. *Cell* **127**, 679–695.

[7] Huber MA, Kraut N, and Beug H (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol* **17**, 548–558.

[8] Kopfstein L and Christofori G (2006). Metastasis: cell-autonomous mechanisms *versus* contributions by the tumor microenvironment. *Cell Mol Life Sci* **63**, 449–468.

[9] Roschke AV, Tonon G, Gehlhaus KS, McTyre N, Bussey KJ, Lababidi S, Scudiero DA, Weinstein JN, and Kirsch IR (2003). Karyotypic complexity of the NCI-60 drug-screening panel. *Cancer Res* **63**(24), 8634–8647.

[10] Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET, et al. (2005). An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* **102**(38), 13550–13555.

[11] Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M, et al. (2008). Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**(9), R139.

[12] Muggerud AA, Edgren H, Wolf M, Kleivi K, Dejeux E, Tost J, Sørlie T, and Kallioniemi O (2009). Data integration from two microarray platforms identifies bi-allelic genetic inactivation of RIC8A in a breast cancer cell line. *BMC Med Genomics* **2**, 26.

[13] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, et al. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* **27**(8), 1160–1167.

[14] Langerød A, Zhao H, Borgan Ø, Nesland JM, Bukholm IR, Ikdahl T, Kåresen R, Børresen-Dale AL, and Jeffrey SS (2009). TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res* **9**(3), R30.

[15] Barrett MT, Scheffer A, Ben-Dor A, Sampas N, Lipson D, Kincaid R, Tsang P, Curry B, Baird K, Meltzer PS, et al. (2004). Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci USA* **101**(51), 17765–17770.

[16] Baumbusch LO, Aarøe J, Johansen FE, Hicks J, Sun H, Bruhn L, Gunderson K, Naume B, Kristensen VN, Liestøl K, et al. (2008). Comparison of the Agilent, ROMA/NimbleGen and Illumina platforms for classification of copy number alterations in human breast tumors. *BMC Genomics* **9**, 379.

[17] Mousses S, Caplen NJ, Cornelison R, Weaver D, Basik M, Hautaniemi S, Elkahloun AG, Lotufo RA, Choudary A, Dougherty ER, et al. (2003). RNAi microarray analysis in cultured mammalian cells. *Genome Res* **13**(10), 2341–2347.

[18] Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, and Lempicki RA (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* **4**(5), P3.

[19] Huang DW, Sherman BT, and Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**(1), 44–57.

[20] Thomassen M, Tan Q, and Kruse TA (2009). Gene expression meta-analysis identifies chromosomal regions and candidate genes involved in breast cancer metastasis. *Breast Cancer Res Treat* **113**(2), 239–249.

[21] Adeleide J, Finetti P, Bekhouche I, Repellini L, Geneix J, Sircoulomb F, Charafe-Jauffret E, Cervera N, Desplans D, Parzy D, et al. (2007). Integrated profiling of basal and luminal breast cancers. *Cancer Res* **67**(24), 11565–11575.

[22] Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, et al. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**(6), 515–527.

[23] Roschke AV, Glebov OK, Lababidi S, Gehlhaus KS, Weinstein JN, and Kirsch IR (2008). Chromosomal instability is associated with higher expression of genes implicated in epithelial-mesenchymal transition, cancer invasiveness, and metastasis and with lower expression of genes involved in cell cycle checkpoints, DNA repair, and chromatin maintenance. *Neoplasia* **10**(11), 1222–1230.

[24] Furuya T, Uchiyama T, Murakami T, Adachi A, Kawauchi S, Oga A, Hirano T, and Sasaki K (2000). Relationship between chromosomal instability and intratumoral regional DNA ploidy heterogeneity in primary gastric cancers. *Clin Cancer Res* **6**(7), 2815–2820.

[25] Risques RA, Moreno V, Ribas M, Marcuello E, Capella G, and Peinado MA (2003). Genetic pathways and genome-wide determinants of clinical outcome in colorectal cancer. *Cancer Res* **63**(21), 7206–7214.

[26] Kronenwett U, Huwendiek S, Ostring C, Portwood N, Roblick UJ, Pawitan Y, Alaiya A, Sennerstam R, Zetterberg A, and Auer G (2004). Improved grading of breast adenocarcinomas based on genomic instability. *Cancer Res* **64**(3), 904–909.

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

[28] Walther A, Houlston R, and Tomlinson I (2008). Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut* **57**(7), 941–950.

[29] MacNeill SA (2010). Structure and function of the GINS complex, a key component of the eukaryotic replisome. *Biochem J* **425**(3), 489–500.

[30] Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, and Araki H (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* **17**, 1153–1165.

[31] Bauerschmidt C, Pollok S, Kremmer E, Nasheuer HP, and Grosse F (2007). Interactions of human Cdc45 with the Mcm2-7 complex, the GINS complex, and DNA polymerases delta and epsilon during S phase. *Genes Cells* **12**, 745–758.

[32] Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, and Labib K (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* **8**, 358–366.

[33] Kubota Y, Takase Y, Komori Y, Hashimoto Y, Arata T, Kamimura Y, Araki H, and Takisawa H (2003). A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes Dev* **17**, 1141–1152.

[34] Kanemaki M, Sanchez-Diaz A, Gambus A, and Labib K (2003). Functional proteomic identification of DNA replication proteins by induced proteolysis *in vivo*. *Nature* **423**, 720–724.

[35] Moyer SE, Lewis PW, and Botchan MR (2006). Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci USA* **103**, 10236–10241.

[36] Pacek M, Tutter AV, Kubota Y, Takisawa H, and Walter JC (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell* **21**, 581–587.

[37] Chang YP, Wang G, Bermudez V, Hurwitz J, and Chen XS (2007). Crystal structure of the GINS complex and functional insights into its role in DNA replication. *Proc Natl Acad Sci USA* **104**, 12685–12690.

[38] De Falco M, Ferrari E, De Felice M, Rossi M, Hubscher U, and Pisani FM (2007). The human GINS complex binds to and specifically stimulates human DNA polymerase α-primase. *EMBO Rep* **8**, 99–103.

[39] Boskovic J, Coloma J, Aparicio T, Zhou M, Robinson CV, Méndez J, and Montoya G (2007). Molecular architecture of the human GINS complex. *EMBO Rep* **8**, 678–684.

[40] Barkley LR, Song IY, Zou Y, and Vaziri C (2009). Reduced expression of GINS complex members induces hallmarks of pre-malignancy in primary untransformed human cells. *Cell Cycle* **8**, 1577–1588.

[41] Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER III, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, et al. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**(5828), 1160–1166.

[42] Ryu B, Kim DS, Deluca AM, and Alani RM (2007). Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PLoS One* **2**(7), e594.

[43] Hayashi R, Arauchi T, Tategu M, Goto Y, and Yoshida K (2006). A combined computational and experimental study on the structure-regulation relationships of putative mammalian DNA replication initiator GINS. *Genomics Proteomics Bioinformatics* **4**, 156–164.

[44] Huang HK, Bailis JM, Leverson JD, Gomez EB, Forsburg SL, and Hunter T (2005). Suppressors of Bir1p (Survivin) identify roles for the chromosomal passenger protein Pic1p (INCENP) and the replication initiation factor Psf2p in chromosome segregation. *Mol Cell Biol* **25**, 9000–9015.

[45] Lau E, Tsuji T, Guo L, Lu SH, and Jiang W (2007). The role of pre-replicative complex (pre-RC) components in oncogenesis. *FASEB J* **21**, 3786–3794.

[46] Giaginis C, Georgiadou M, Dimakopoulou K, Tsourouflis G, Gatzidou E, Kouraklis G, and Theocharis S (2009). Clinical significance of MCM-2 and MCM-5 expression in colon cancer: association with clinicopathological parameters and tumor proliferative capacity. *Dig Dis Sci* **54**, 282–291.

[47] Xi Y, Nakajima G, Schmitz JC, Chu E, and Ju J (2006). Multi-level gene expression profiles affected by thymidylate synthase and 5-fluorouracil in colon cancer. *BMC Genomics* **7**, 68.

[48] Gavin EJ, Song B, Wang Y, Xi Y, and Ju J (2008). Reduction of Orc6 expression sensitizes human colon cancer cells to 5-fluorouracil and cisplatin. *PLoS One* **3**, e4054.

Table W1. Top 177 Genes Inducing Polyploidy in MDA-MB-231 Cells When Inhibited with RNAi.

Gene Name	Description	siRNA IDs	Multinucleated	Apoptotic
ADAM29	ADAM metallopeptidase domain 29	Q003485		1
ADAMTS9	ADAM metallopeptidase with thrombospondin type 1 motif, 9	Q004438		1
ALS2	Amyotrophic lateral sclerosis 2 (juvenile)	Q004524	1	
ANXA5	Annexin A5	Q007210	1	
APOC1	Apolipoprotein C-I	Q007215		1
AQP3	Aquaporin 3 (Gill blood group)	Q000150 | Q018226	1	
ARTN	Artemin	Q002910	1	
ASL	Argininosuccinate lyase	Q007240 | Q025316		1
ATRN	Attractin	Q002734	1	
ATXN3	Ataxin 3	Q001569	1	
AZGP1	alpha-2-Glycoprotein 1, zinc binding	Q007258 | Q025334		1
AZIN1	Antizyme inhibitor 1	Q004137	1	
BACE2	beta-Site APP-cleaving enzyme 2	Q003749		1
BARD1	BRCA1-associated RING domain 1	Q000228	1	
BDH2	3-Hydroxybutyrate dehydrogenase, type 2	Q004418	1	
BECN1	Beclin 1 (coiled-coil, myosin-like BCL2 interacting protein)	Q002795		1
BLVRA	Biliverdin reductase A	Q007268		1
BSCL2	Bernardinelli-Seip congenital lipodystrophy 2 (seipin)	Q003844		1
BTNL9	Butyrophilin-like 9	Q005311	1	
CAV1	Caveolin 1, caveolae protein, 22 kDa	Q000350	1	
CCL22	Chemokine (C-C motif) ligand 22	Q002184	1	
CCNL2	Cyclin L2	Q004818	1	
CCT6A	Chaperonin containing TCP1, subunit 6A (zeta 1)	Q007310	1	
CDC23	Cell division cycle 23 homolog (*Saccharomyces cerevisiae*)	Q002800		1
CDH11	Cadherin 11, type 2, OB-cadherin (osteoblast)	Q000439	1	
CHRNA5	Cholinergic receptor, nicotinic, alpha 5	Q000476	1	
CLCA2	Chloride channel, calcium-activated, family member 2	Q003092		1
CLDN4	Claudin 4	Q000544	1	
CLN3	Ceroid-lipofuscinosis, neuronal 3, juvenile (Batten, Spielmeyer-Vogt disease)	Q000506	1	
CRY2	Cryptochrome 2 (photolyase-like)	Q000566	1	
CTSD	Cathepsin D	Q000596	1	
CXADR	Coxsackie virus and adenovirus receptor	Q000607	1	
DAD1	Defender against cell death 1	Q000648	1	
DCC	Deleted in colorectal carcinoma	Q000658	1	
DCLRE1C	DNA cross link repair 1C (PSO2 homolog, *S. cerevisiae*)	Q004633		1
DCXR	Dicarbonyl/L-xylulose reductase	Q004055		1
DDIT3	DNA damage–inducible transcript 3	Q000669	1	
DIO1	Deiodinase, iodothyronine, type I	Q000688	1	
DKC1	Dyskeratosis congenita 1, dyskerin	Q000691	1	
DMC1	DMC1 dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (yeast)	Q003505	1	
EED	Embryonic ectoderm development	Q002810	1	
EEF1A1	Eukaryotic translation elongation factor 1 alpha 1	Q000746	1	
EIF2C1	Eukaryotic translation initiation factor 2C, 1	Q003840	1	
ENTPD6	Ectonucleoside triphosphate diphosphohydrolase 6 (putative function)	Q007314	1	
EPRS	Glutamyl-prolyl-tRNA synthetase	Q000788	1	
EVI2A	Ecotropic viral integration site 2A	Q000812 | Q018888	1	
FANCD2	Fanconi anemia, complementation group D2	Q000836 | Q018912	1	
FANCE	Fanconi anemia, complementation group E	Q000837	1	
FBN1	Fibrillin 1	Q000848	1	
FEM1B	Fem-1 homolog b (*Caenorhabditis elegans*)	Q003223	1	
FKBP2	FK506 binding protein 2, 13 kDa	Q000878	1	
FLOT2	Flotillin 2	Q000890	1	
FMO3	Flavin-containing monooxygenase 3	Q000892 | Q018968	1	
FOXK2	Forkhead box K2	Q001283	1	
FTL	Ferritin, light polypeptide	Q000916	1	
FZR1	Fizzy/cell division cycle 20–related 1 (*Drosophila*)	Q004092	1	
GANAB	Glucosidase, alpha; neutral AB	Q003617	1	
GDF3	Growth differentiation factor 3	Q003070	1	
GDF9	Growth differentiation factor 9	Q000972	1	
GGH	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	Q002855		1
GINS2	GINS complex subunit 2 (Psf2 homolog)	Q004146 | Q022222	1	
GNG11	Guanine nucleotide binding protein (G protein), gamma 11	Q001032 | Q019108	1	
GPC4	Glypican 4	Q000858	1	
GPD2	Glycerol-3-phosphate dehydrogenase 2 (mitochondrial)	Q001044	1	
GRIA2	Glutamate receptor, ionotropic, AMPA 2	Q001061	1	
GRIK3	Glutamate receptor, ionotropic, kainate 3	Q001071	1	
GSN	Gelsolin (amyloidosis, Finnish type)	Q001086	1	
HDAC3	Histone deacetylase 3	Q002860		1

Table W1. (*continued*)

Gene Name	Description	siRNA IDs	Multinucleated	Apoptotic
HDAC4	Histone deacetylase 4	Q003117 | Q021193		1
HDAC6	Histone deacetylase 6	Q003185	1	
HERC3	Hect domain and RLD 3	Q002885 | Q020961		1
HERC6	Hect domain and RLD 6	Q004275	1	
HGFAC	HGF activator	Q001138	1	
HIP1	Huntingtin interacting protein 1	Q001141 | Q019217	1	
HNF4G	Hepatocyte nuclear factor 4, gamma	Q001159	1	
HNMT	Histamine *N*-methyltransferase	Q001160	1	
HSD11B1L	Hydroxysteroid (11-beta) dehydrogenase 1-like	Q005625	1	
HSD17B12	Hydroxysteroid (17-beta) dehydrogenase 12	Q004044		1
HSD17B4	Hydroxysteroid (17-beta) dehydrogenase 4	Q001185	1	
HTRA2	HtrA serine peptidase 2	Q003914	1	
HUS1	HUS1 checkpoint homolog (*Schizosaccharomyces pombe*)	Q001195	1	
IGF2R	Insulin-like growth factor 2 receptor	Q001234	1	
IL8	Interleukin 8	Q001262	1	
IMPG1	Interphotoreceptor matrix proteoglycan 1	Q001288	1	
INCENP	Inner centromere protein antigens 135/155 kDa	Q001289 | Q019365	1	
INDO	Indoleamine-pyrrole 2,3 dioxygenase	Q001290	1	
IQGAP3	IQ motif containing GTPase-activating protein 3	Q005193	1	
ITGA3	Integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	Q001316	1	
ITGB2	Integrin, beta 2 (complement component 3 receptors 3 and 4 subunit)	Q001328	1	
ITPR3	Inositol 1,4,5-triphosphate receptor, type 3	Q001338	1	
JAG2	Jagged 2	Q001340	1	
KCNS3	Potassium voltage-gated channel, delayed-rectifier, subfamily S, member 3	Q001398	1	
KIF12	Kinesin family member 12	Q005086 | Q023162	1	
KIF13A	Kinesin family member 13A	Q004587 | Q022663		1
KIF15	Kinesin family member 15	Q004433	1	
KIF22	Kinesin family member 22	Q001416	1	
KIF23	Kinesin family member 23	Q003039		1
KIF5C	Kinesin family member 5C	Q001404		1
KIFC1	Kinesin family member C1	Q001414	1	
KNTC1	Kinetochore-associated 1	Q003108		1
KTN1	Kinectin 1 (kinesin receptor)	Q001424	1	
LAMA5	Laminin, alpha 5	Q001429	1	
LIG1	Ligase I, DNA, ATP-dependent	Q001453 | Q019529	1	
LIG3	Ligase III, DNA, ATP-dependent	Q001454		1
LRP2	Low-density lipoprotein-related protein 2	Q001477	1	
LRRN2	Leucine-rich repeat neuronal 2	Q003324	1	
MARS	Methionyl-tRNA synthetase	Q001522	1	
MCM2	Minichromosome maintenance complex component 2	Q001532		1
MCTS1	Malignant T-cell–amplified sequence 1	Q003927	1	
MED24	Mediator complex subunit 24	Q003137 | Q021213	1	
MICB	MHC class I polypeptide-related sequence B	Q001562	1	
MLLT4	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, *Drosophila*); translocated to, 4	Q001576	1	
MMP3	Matrix metallopeptidase 3 (stromelysin 1, progelatinase)	Q001583 | Q019659	1	
MNAT1	Ménage à trois homolog 1, cyclin H assembly factor (*Xenopus laevis*)	Q001597	1	
MRAS	Muscle RAS oncogene homolog	Q003557	1	
MSH4	MutS homolog 4 (*Escherichia coli*)	Q001612	1	
MSTN	Myostatin	Q000971	1	
MTR	5-Methyltetrahydrofolate-homocysteine methyltransferase	Q001624	1	
MTUS1	Mitochondrial tumor suppressor 1	Q004489	1	
MUC1	Mucin 1, cell surface–associated	Q001626 | Q019702	1	
NEDD4	Neural precursor cell expressed, developmentally downregulated 4	Q001672	1	
NEDD4L	Neural precursor cell expressed, developmentally downregulated 4-like	Q003639	1	
NPC1L1	NPC1 (Niemann-Pick disease, type C1, gene)-like 1	Q003953		1
NPC2	Niemann-Pick disease, type C2	Q003358		1
NUDT9	Nudix (nucleoside diphosphate linked moiety X)-type motif 9	Q004173		1
P4HB	Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide	Q001757	1	
PAQR4	Progestin and adipoQ receptor family member IV	Q005170		1
PCMT1	Protein-L-isoaspartate (D-aspartate) *O*-methyltransferase	Q001786	1	
PCNA	Proliferating cell nuclear antigen	Q001787 | Q019863		1
PHF11	PHD finger protein 11	Q004041		1
PHF12	PHD finger protein 12	Q004517	1	
PMAIP1	Phorbol-12-myristate-13-acetate–induced protein 1	Q001889	1	

Table W1. (*continued*)

Gene Name	Description	siRNA IDs	Multinucleated	Apoptotic
PMS1	PMS1 postmeiotic segregation increased 1 (*S. cerevisiae*)	Q001893	1	
PMS2	PMS2 postmeiotic segregation increased 2 (*S. cerevisiae*)	Q001895	1	
POLB	Polymerase (DNA-directed), beta	Q001902	1	
PPIL5	Peptidylprolyl isomerase (cyclophilin)-like 5	Q005158		1
PSCD3	Pleckstrin homology, Sec7 and coiled-coil domains 3	Q002967	1	
PSMC4	Proteasome (prosome, macropain) 26S subunit, ATPase, 4	Q001977		1
PTGES3	Prostaglandin E synthase 3	Q003397	1	
PYGO2	Pygopus homolog 2 (*Drosophila*)	Q005012	1	
RAB28	RAB28, member RAS oncogene family	Q002993	1	
RAB40B	RAB40B, member RAS oncogene family	Q003457	1	
RAB4A	RAB4A, member RAS oncogene family	Q002014		1
RAB8A	RAB8A, member RAS oncogene family	Q001544	1	
RAB9A	RAB9A, member RAS oncogene family	Q002994	1	
RAC1	Ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	Q002024 | Q020100		1
RAD50	RAD50 homolog (*S. cerevisiae*)	Q003221		1
RANBP5	RAN binding protein 5	Q001420 | Q019496	1	
RAPGEFL1	Rap guanine nucleotide exchange factor (GEF)-like 1	Q004059 | Q022135	1	
RBL2	Retinoblastoma-like 2 (p130)	Q002067		1
RERG	RAS-like, estrogen-regulated, growth inhibitor	Q004966 | Q023042	1	
RNF14	Ring finger protein 14	Q003080 | Q021156	1	
RNF7	Ring finger protein 7	Q003084	1	
RRM2B	Ribonucleotide reductase M2 B (TP53-inducible)	Q003992	1	
RSF1	Remodeling and spacing factor 1	Q004167	1	
SIRT6	Sirtuin (silent mating type information regulation 2 homolog) 6 (*S. cerevisiae*)	Q004130	1	
SLC9A3R1	Solute carrier family 9 (sodium/hydrogen exchanger), member 3 regulator 1	Q002995 | Q021071	1	
SMAD7	SMAD family member 7	Q001509	1	
SMUG1	Single-strand selective monofunctional uracil-DNA glycosylase 1	Q003700	1	
SNX6	Sorting nexin 6	Q004556	1	
SOX4	SRY (sex-determining region Y)-box 4	Q002305	1	
SPATA20	Spermatogenesis-associated 20	Q004656	1	
SRPX2	Sushi-repeat-containing protein, X-linked 2	Q003898	1	
TAP2	Transporter 2, ATP-binding cassette, subfamily B (MDR/TAP)	Q002369	1	
TFDP1	Transcription factor Dp-1	Q002410		1
TGFB1I1	Transforming growth factor beta 1-induced transcript 1	Q002420		1
TMEM30A	Transmembrane protein 30A	Q004351	1	
TNN	Tenascin N	Q004582	1	
TRPM4	Transient receptor potential cation channel, subfamily M, member 4	Q004249		1
TUBG2	Tubulin, gamma 2	Q003884 | Q021960	1	
TXN	Thioredoxin	Q002530	1	
TYMS	Thymidylate synthetase	Q002532 | Q020608		1
UNC13B	Unc-13 homolog B (*C. elegans*)	Q003337		1
WDHD1	WD repeat and HMG-box DNA binding protein 1	Q003510 | Q021586		1
WFS1	Wolfram syndrome 1 (wolframin)	Q002594	1	
VWF	Von Willebrand factor	Q002590		1
YEATS4	YEATS domain containing 4	Q002676	1	

Table W2. GO Categories Associated with RNAi Hit Genes Inducing Polyploidy.

Category	GO Term	Count	*P*	Benjamini
GOTERM_BP_ALL	Response to DNA damage stimulus	21	3.7*e*–10	1.9*e*–06
GOTERM_BP_ALL	Response to endogenous stimulus	23	4.7*e*–10	8.3*e*–07
GOTERM_BP_ALL	Cell cycle	34	4.7*e*–10	1.2*e*–06
GOTERM_BP_ALL	DNA repair	19	7.2*e*–10	9.5*e*–07
GOTERM_BP_ALL	Response to stress	36	4.1*e*–09	4.3*e*–06
GOTERM_BP_ALL	DNA metabolic process	31	1.2*e*–08	0.000011
GOTERM_BP_ALL	Cell cycle process	28	3.7*e*–08	0.000028
GOTERM_BP_ALL	DNA replication	14	3.5*e*–06	0.002
GOTERM_BP_ALL	M phase	14	.000015	0.0074
GOTERM_BP_ALL	Negative regulation of cellular process	29	.000034	0.015
GOTERM_BP_ALL	Cell cycle phase	15	.000035	0.014
GOTERM_BP_ALL	Regulation of progression through cell cycle	18	.000061	0.023
GOTERM_BP_ALL	Regulation of cell cycle	18	.000065	0.023
GOTERM_BP_ALL	Negative regulation of biological process	29	.000071	0.023
GOTERM_BP_ALL	Microtubule-based movement	8	.00022	0.066
GOTERM_BP_ALL	Localization	55	.00023	0.066
GOTERM_BP_ALL	Chromosome organization and biogenesis	14	.00038	0.099
GOTERM_BP_ALL	Biological regulation	79	.00047	0.12
GOTERM_BP_ALL	Cytoskeleton-dependent intracellular transport	8	.00057	0.13
GOTERM_BP_ALL	Regulation of cellular process	69	.00059	0.13
GOTERM_BP_ALL	Microtubule-based process	10	.00068	0.14
GOTERM_BP_ALL	Mitotic cell cycle	12	.00072	0.15
GOTERM_BP_ALL	Mitosis	10	.00086	0.16
GOTERM_BP_ALL	M phase of mitotic cell cycle	10	.00091	0.17
GOTERM_BP_ALL	Developmental process	54	.00095	0.17
GOTERM_BP_ALL	Meiotic recombination	4	.0011	0.18
GOTERM_BP_ALL	DNA-dependent DNA replication	7	.0013	0.21
GOTERM_BP_ALL	Regulation of biological process	71	.0017	0.26
GOTERM_BP_ALL	Regulation of apoptosis	15	.0023	0.31
GOTERM_BP_ALL	Regulation of programmed cell death	15	.0025	0.33
GOTERM_BP_ALL	Meiosis I	4	.0041	0.47
GOTERM_BP_ALL	Cytoskeleton organization and biogenesis	14	.005	0.53
GOTERM_BP_ALL	Establishment of localization	45	.0053	0.54
GOTERM_BP_ALL	DNA recombination	7	.0062	0.57
GOTERM_BP_ALL	Localization of cell	12	.0064	0.58
GOTERM_BP_ALL	Cell motility	12	.0064	0.58
GOTERM_BP_ALL	Establishment of cellular localization	19	.0074	0.61
GOTERM_BP_ALL	Regulation of biological quality	19	.0074	0.61
GOTERM_BP_ALL	Amino acid and derivative metabolic process	11	.0081	0.62
GOTERM_BP_ALL	Transport	43	.0083	0.62
GOTERM_BP_ALL	Anatomical structure development	36	.0088	0.63
GOTERM_BP_ALL	Death	18	.0094	0.65
GOTERM_BP_ALL	Cell death	18	.0094	0.65
GOTERM_BP_ALL	Cell migration	9	.0096	0.65
GOTERM_BP_ALL	Cellular localization	19	.0097	0.64
GOTERM_CC_ALL	Nuclear part	29	.000004	0.0017
GOTERM_CC_ALL	Microtubule	12	.000031	0.009
GOTERM_CC_ALL	Intracellular organelle part	60	.000031	0.0067
GOTERM_CC_ALL	Organelle part	60	.000034	0.0059
GOTERM_CC_ALL	Macromolecular complex	48	.000059	0.0085
GOTERM_CC_ALL	Spindle	7	.00016	0.019
GOTERM_CC_ALL	Nucleoplasm	15	.00066	0.069
GOTERM_CC_ALL	Microtubule cytoskeleton	13	.0011	0.098
GOTERM_CC_ALL	Nuclear lumen	17	.0015	0.12
GOTERM_CC_ALL	Organelle lumen	21	.0016	0.12
GOTERM_CC_ALL	Membrane-enclosed lumen	21	.0016	0.12
GOTERM_CC_ALL	Microtubule associated complex	7	.0018	0.11
GOTERM_CC_ALL	Endoplasmic reticulum part	14	.0019	0.11
GOTERM_CC_ALL	Endoplasmic reticulum	19	.0023	0.12
GOTERM_CC_ALL	Cytoplasmic membrane-bound vesicle	11	.003	0.15
GOTERM_CC_ALL	Membrane-bound vesicle	11	.0034	0.16
GOTERM_CC_ALL	Cytoplasmic vesicle	12	.0038	0.17
GOTERM_CC_ALL	Vesicle	12	.0045	0.19
GOTERM_CC_ALL	Nucleoplasm part	12	.005	0.19
GOTERM_CC_ALL	Intracellular membrane-bound organelle	92	.0056	0.2
GOTERM_CC_ALL	Endoplasmic reticulum membrane	12	.0056	0.19
GOTERM_CC_ALL	Membrane-bound organelle	92	.0057	0.19
GOTERM_CC_ALL	Nuclear envelope-endoplasmic reticulum network	12	.0065	0.2
GOTERM_CC_ALL	Intracellular organelle	102	.0098	0.28

Table W3. Genes and siRNA IDs of Candidate Targets Included in the Secondary CSMA Analysis.

Gene Name	ID siRNA_A	ID siRNA_B
BARD1	SI00299383	SI02664354
CDC23	SI02653112	SI02653413
CRY2	SI00354473	SI00354480
DCLRE1C	SI00133945	SI02645286
DDIT3	SI00059535	SI00059528
DKC1	SI00364581	SI00364602
DMC1	SI00094374	SI00094367
FANCD2	SI00145306	SI00145292
GINS2	SI02653056	SI00131404
HDAC3	SI00057323	SI02653581
HDAC4	SI00083958	SI00057316
HDAC6	SI02757769	SI00083951
HUS1	SI02664837	SI02663808
LIG1	SI02757482	SI02633449
LIG3	SI02663549	SI02663542
MCM3	SI02664872	SI02757489
MNAT1	SI00037716	SI02664879
MSH4	SI00037877	SI00037723
PCNA	SI02653357	SI02628598
PMS1	SI00011347	SI02653287
PMS2	SI03115049	SI00011333
POLB	SI00041398	SI00011375
PSMC4	SI00301469	SI02663605
RAD50	SI02653826	SI02665208
RNF14	SI00062258	SI02653742
RRM2B	SI02664977	SI00062244
RSF1	SI00117005	SI02664984
SIRT6	SI02777698	SI00116991
SMUG1	SI00104706	SI00116592
TYMS	SI00021609	SI00104699

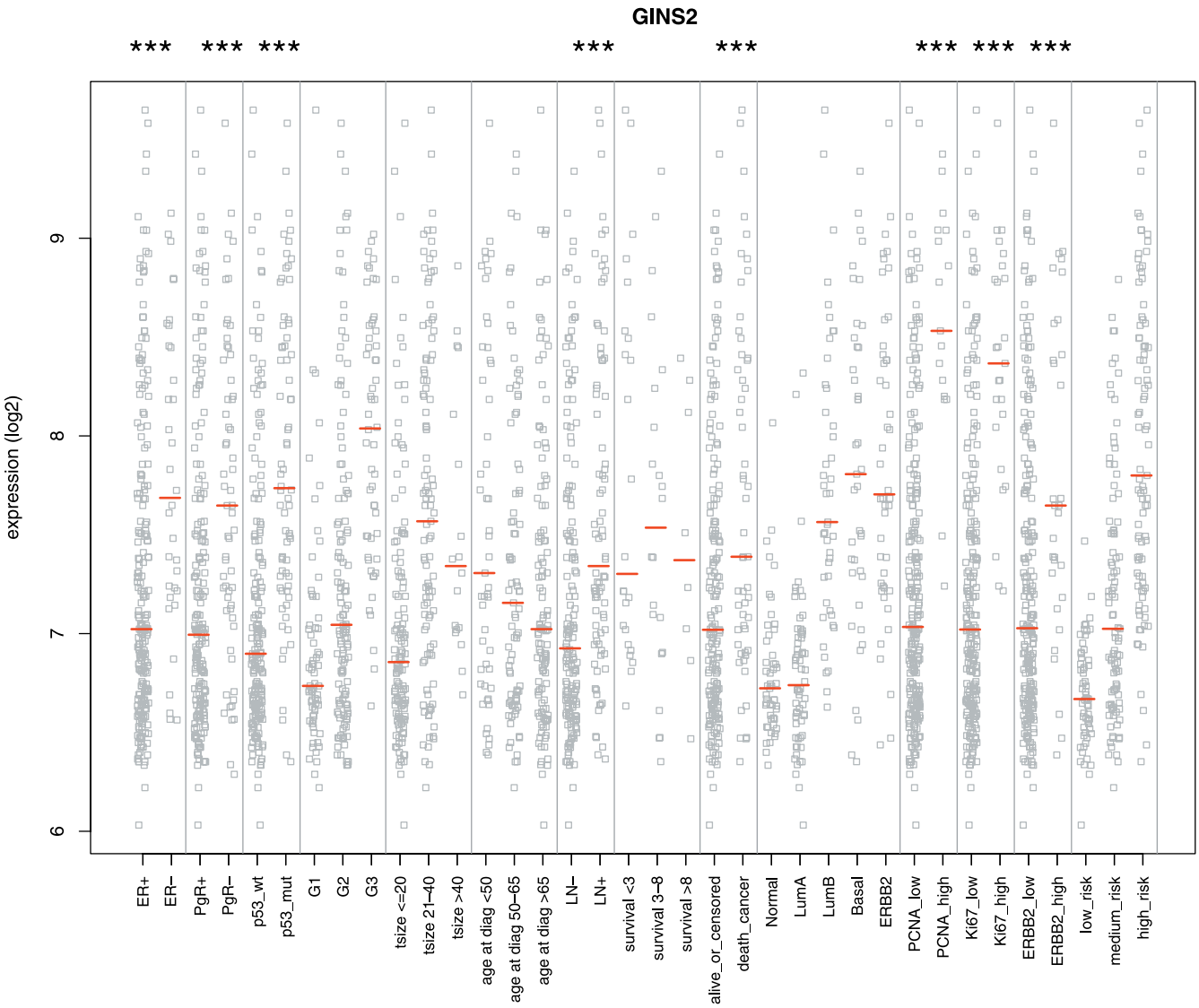


Figure W1. Clinicopathologic significance of polyploidy-inducing RNAi hit genes. The median expression level distribution of the candidate genes was assessed for correlation with 13 different clinicopathologic parameters. ***$P < .001$.

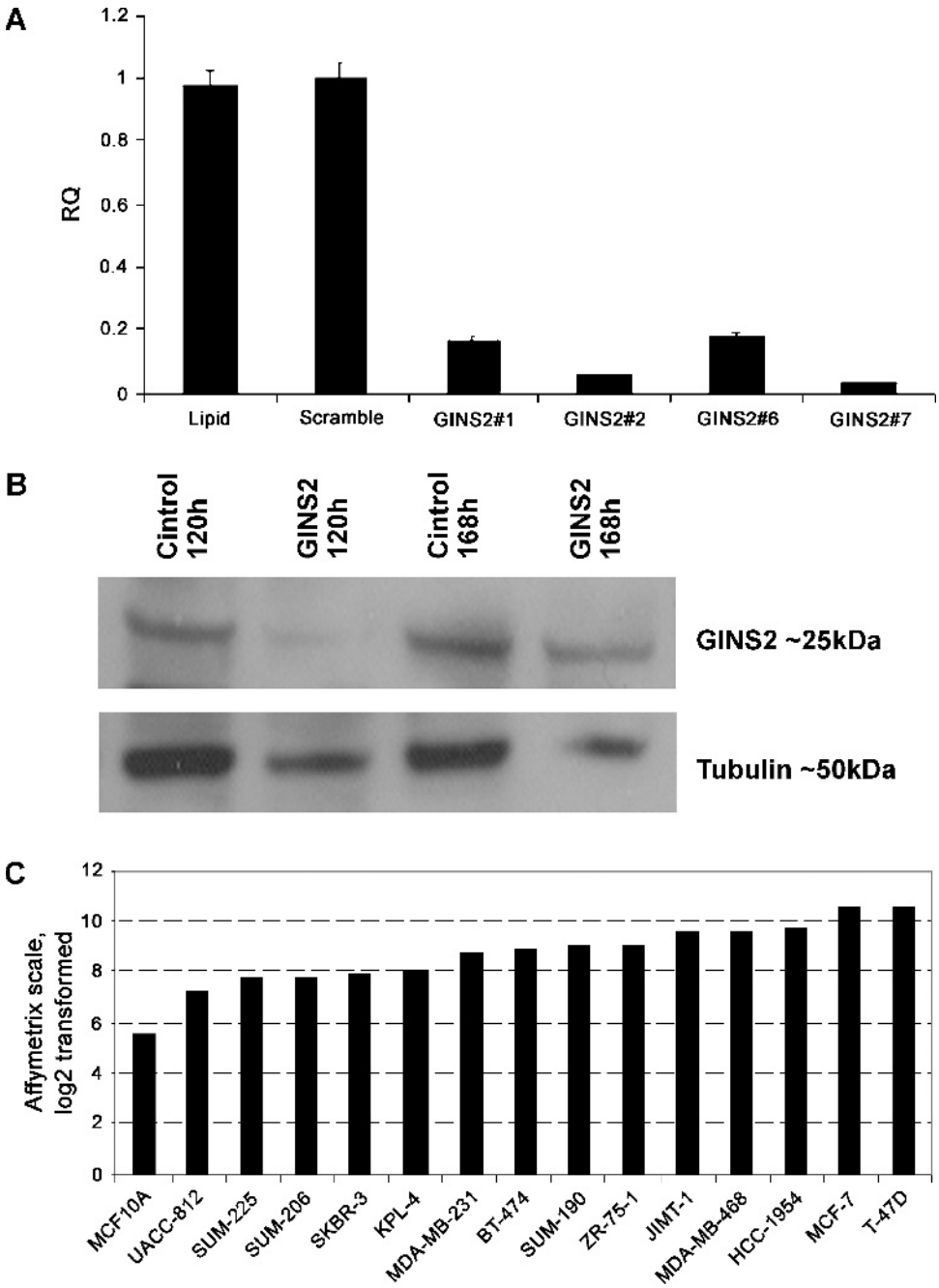


Figure W2. Target verification of GINS2. (A) qRT-PCR analysis of GINS2 transcript levels after knockdown with four distinct siRNAs in MDA-MB-231 cells. Error bars indicate the SE for three qPCR measurements. (B) Immunoblot analysis of GINS2 and α-tubulin protein levels after 5 and 7 days of siRNA knockdown on CSMA in MDA-MB-231 cells. (C) The expression profile of GINS2 mRNA in 14 different breast cancer cell lines and nonmalignant breast epithelial cells. Log_2 Affymetric scale. (D) Array-based CGH data showing amplification of GINS2 loci (chromosomal region at 16q24.1) in T-47D breast cancer cells.

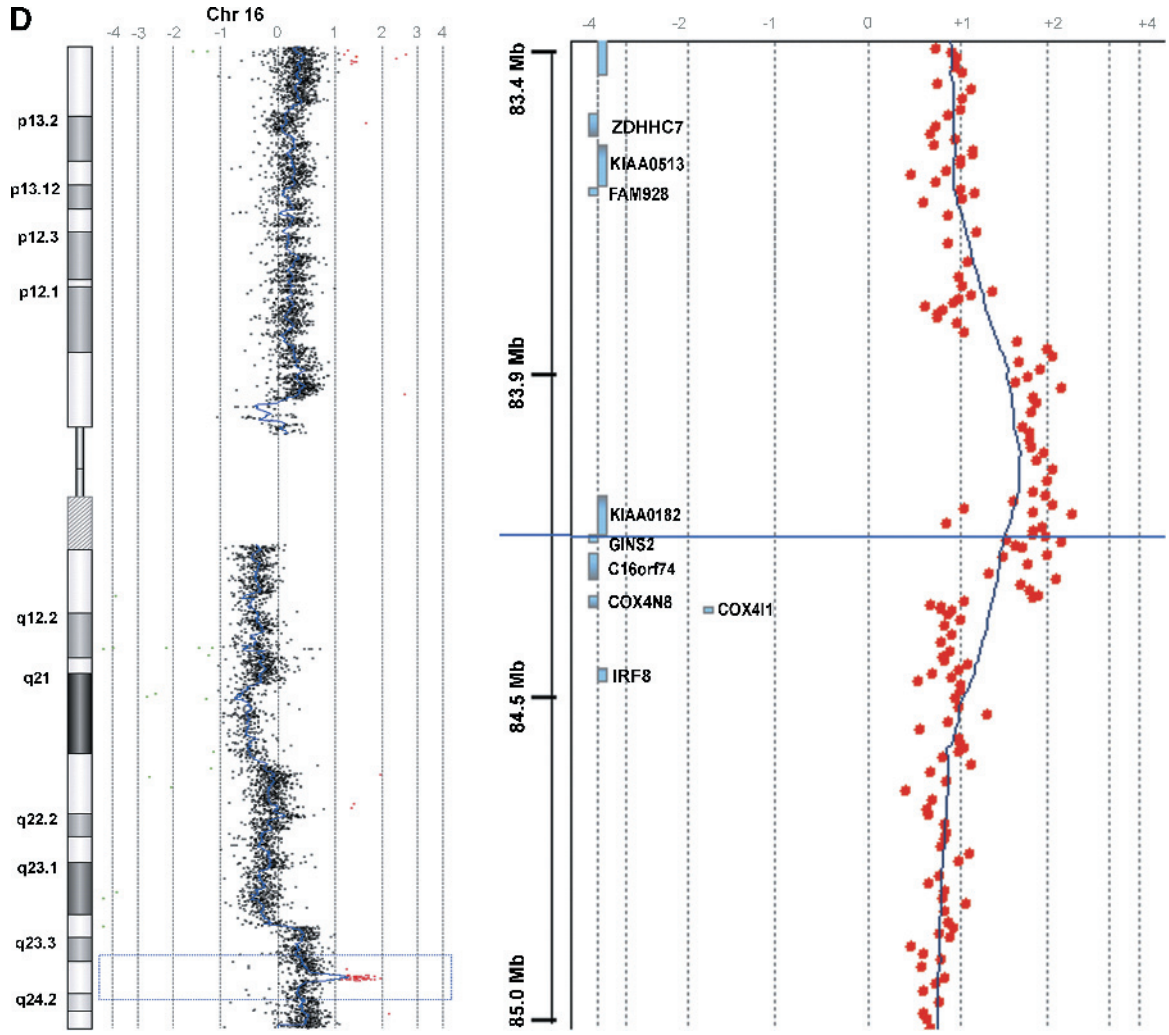


Figure W2. (*continued*).

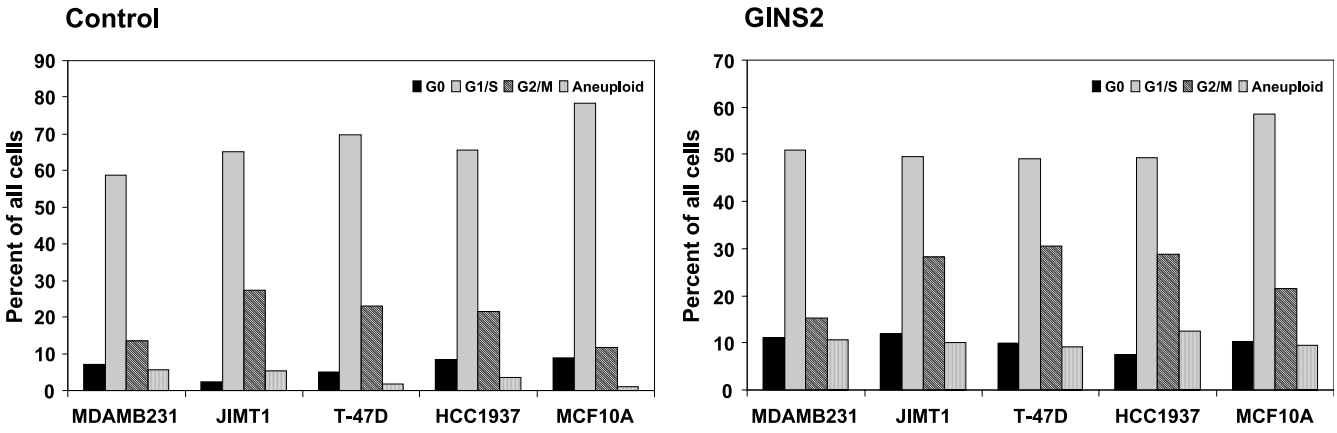


Figure W3. Image based quantification of cell cycle distribution of the analyzed cell lines after GINS2 inhibition.

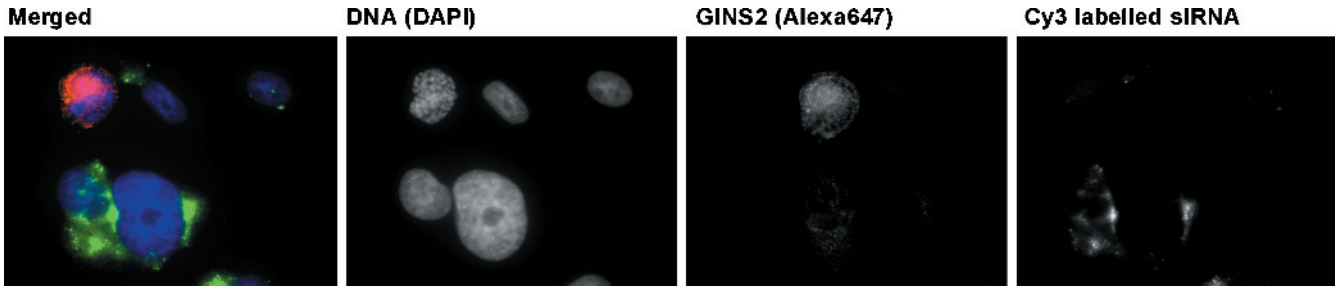


Figure W4. Comparison of GINS2 antibody specificity. Specificity of the GINS2 antibody was evaluated with cells transfected with a Cy3-labeled control siRNA and a validated siRNA for GINS2. A significant association of siRNA uptake and GINS2 protein level inhibition was verified.

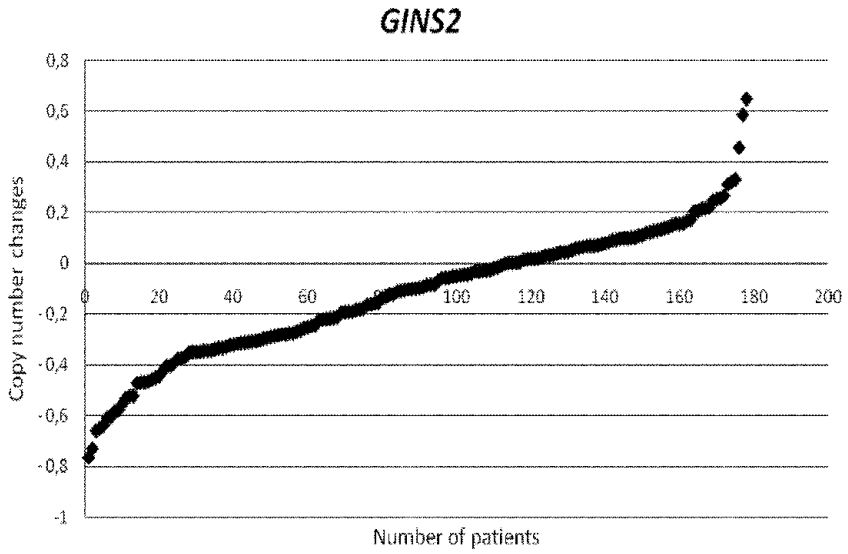


Figure W5. Copy number changes from *GINS2* on 178 breast cancer tumors. Of 178 patients, 47 (26.4%) show loss (<-0.3) of *GINS2*, whereas 6 (3.4%) show gain (>0.3).

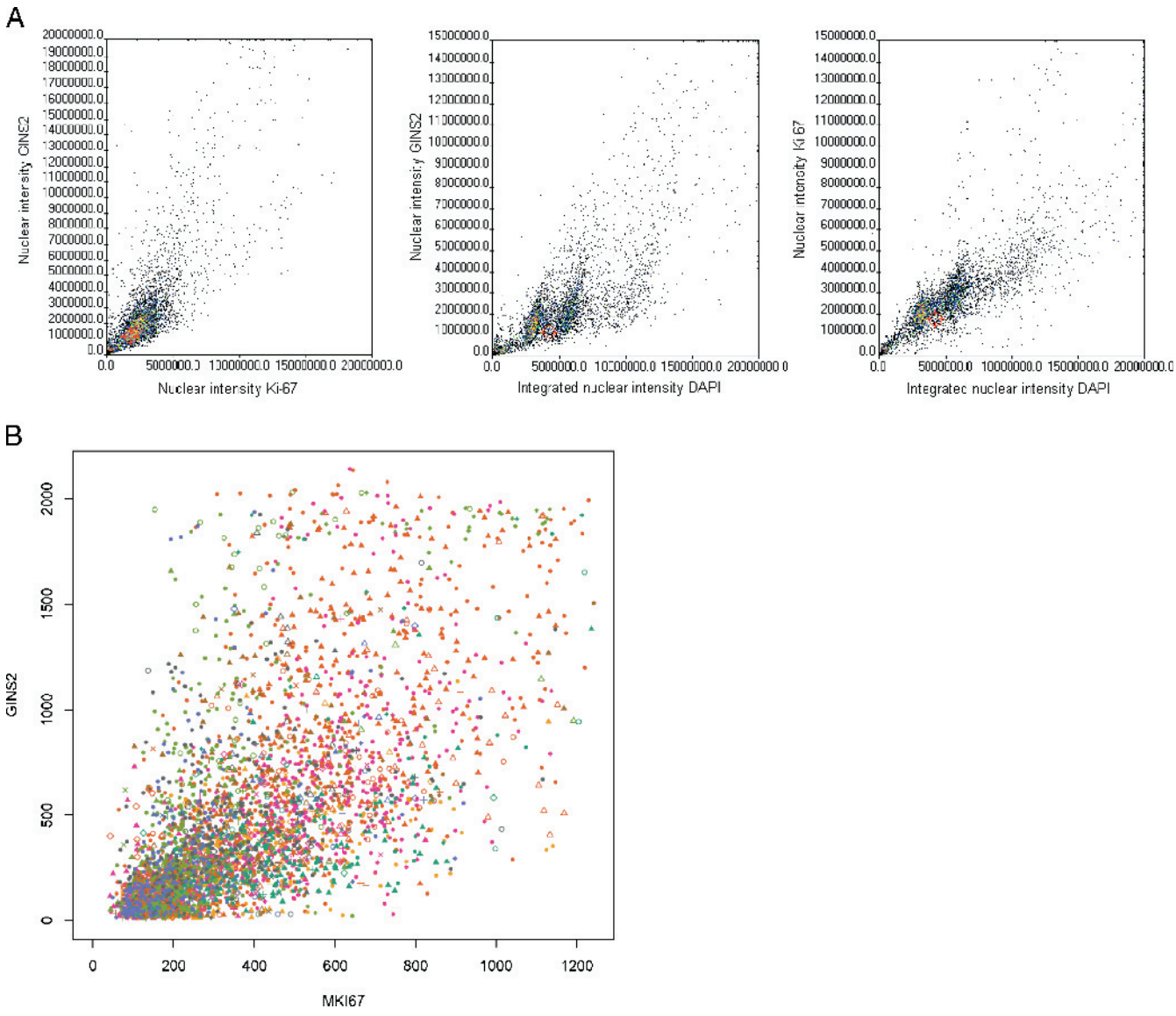


Figure W6. GINS2 protein and mRNA expression pattern. (A) Correlation of nuclear GINS2 and Ki-67 protein staining of MDA-MB-231 cells. (B) mRNA coexpression pattern of *GINS2* and MKI67 in 40 different cancer types analyzed from GeneSapiens gene expression database.