Article

CellPress

The Architecture and Evolution of Cancer Neochromosomes

Dale W. Garsed,[1,8,9,15] Owen J. Marshall,[5,15,17] Vincent D.A. Corbin,[2,3,13,15] Arthur Hsu,[2,15] Leon Di Stefano,[2] Jan Schröder,[2,3] Jason Li,[1] Zhi-Ping Feng,[2,3] Bo W. Kim,[5] Mark Kowarsky,[2] Ben Lansdell,[2] Ross Brookwell,[6] Ola Myklebost,[10] Leonardo Meza-Zepeda,[10] Andrew J. Holloway,[1] Florence Pedeutour,[7] K.H. Andy Choo,[5] Michael A. Damore,[11] Andrew J. Deans,[12] Anthony T. Papenfuss,[2,3,4,8,13,16,*] and David M. Thomas[1,8,14,16,*]

[1]Cancer Genomics, Peter MacCallum Cancer Centre, East Melbourne, VIC 3002, Australia
[2]Bioinformatics Division, The Walter & Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[3]Department of Medical Biology, University of Melbourne, VIC 3010, Australia
[4]Department of Mathematics and Statistics, University of Melbourne, VIC, 3010, Australia
[5]Chromosome Research, Murdoch Childrens Research Institute, and Department of Paediatrics, Royal Children's Hospital, University of Melbourne, Parkville, VIC 3052, Australia
[6]Sullivan Nicolaides Pathology, Indooroopilly, QLD 4068, Australia
[7]Laboratory of Solid Tumors Genetics, Nice University Hospital, Nice 06107, France
[8]Sir Peter MacCallum Department of Oncology, University of Melbourne, VIC 3010, Australia
[9]Department of Pathology, University of Melbourne, VIC 3010, Australia
[10]Department of Tumor Biology, Oslo University Hospital, Norwegian Radium Hospital, Oslo 0424, Norway
[11]Amgen, Thousand Oaks, CA 91320, USA
[12]St Vincent's Institute, Fitzroy, VIC 3065, Australia
[13]Bioinformatics and Cancer Genomics, Peter MacCallum Cancer Centre, East Melbourne, VIC, 3002, Australia
[14]The Kinghorn Cancer Centre, Garvan Institute of Medical Research, Sydney, NSW 2010, Australia
[15]Co-first author
[16]Co-senior author
[17]Present address: Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge CB2 1QN, UK
*Correspondence: papenfuss@wehi.edu.au (A.T.P.), d.thomas@garvan.org.au (D.M.T.)
http://dx.doi.org/10.1016/j.ccell.2014.09.010

SUMMARY

We isolated and analyzed, at single-nucleotide resolution, cancer-associated neochromosomes from well- and/or dedifferentiated liposarcomas. Neochromosomes, which can exceed 600 Mb in size, initially arise as circular structures following chromothripsis involving chromosome 12. The core of the neochromosome is amplified, rearranged, and corroded through hundreds of breakage-fusion-bridge cycles. Under selective pressure, amplified oncogenes are overexpressed, while coamplified passenger genes may be silenced epigenetically. New material may be captured during punctuated chromothriptic events. Centromeric corrosion leads to crisis, which is resolved through neocentromere formation or native centromere capture. Finally, amplification terminates, and the neochromosome core is stabilized in linear form by telomere capture. This study investigates the dynamic mutational processes underlying the life history of a special form of cancer mutation.

INTRODUCTION

Chromosomes are a fundamental organizational unit of genetic information relevant to biology and disease. Contiguous se-quences of circular or linear DNA and associated proteins from 0.1 Mb to 1 Gb in length (Paux et al., 2008), all nuclear chromosomes require centromeres and, when in linear form, telomeres (Baird and Farr, 2006). In humans, in addition to 46 normal

Significance

Different cancers reveal distinct mutation patterns, from mismatch repair defects in colorectal cancers to translocations that generate fusion genes in subtypes of leukemia. Neochromosomes are a form of mutation comprising the somatic acquisition of massive, episomal, self-replicating structures within cancer cells, specific to certain cancer types, including several sarcomas. We describe, at single-nucleotide level, the structure and content of well- and/or dedifferentiated liposarcoma neochromosomes, and we used mathematical modeling to understand their origins and oncogenic properties.
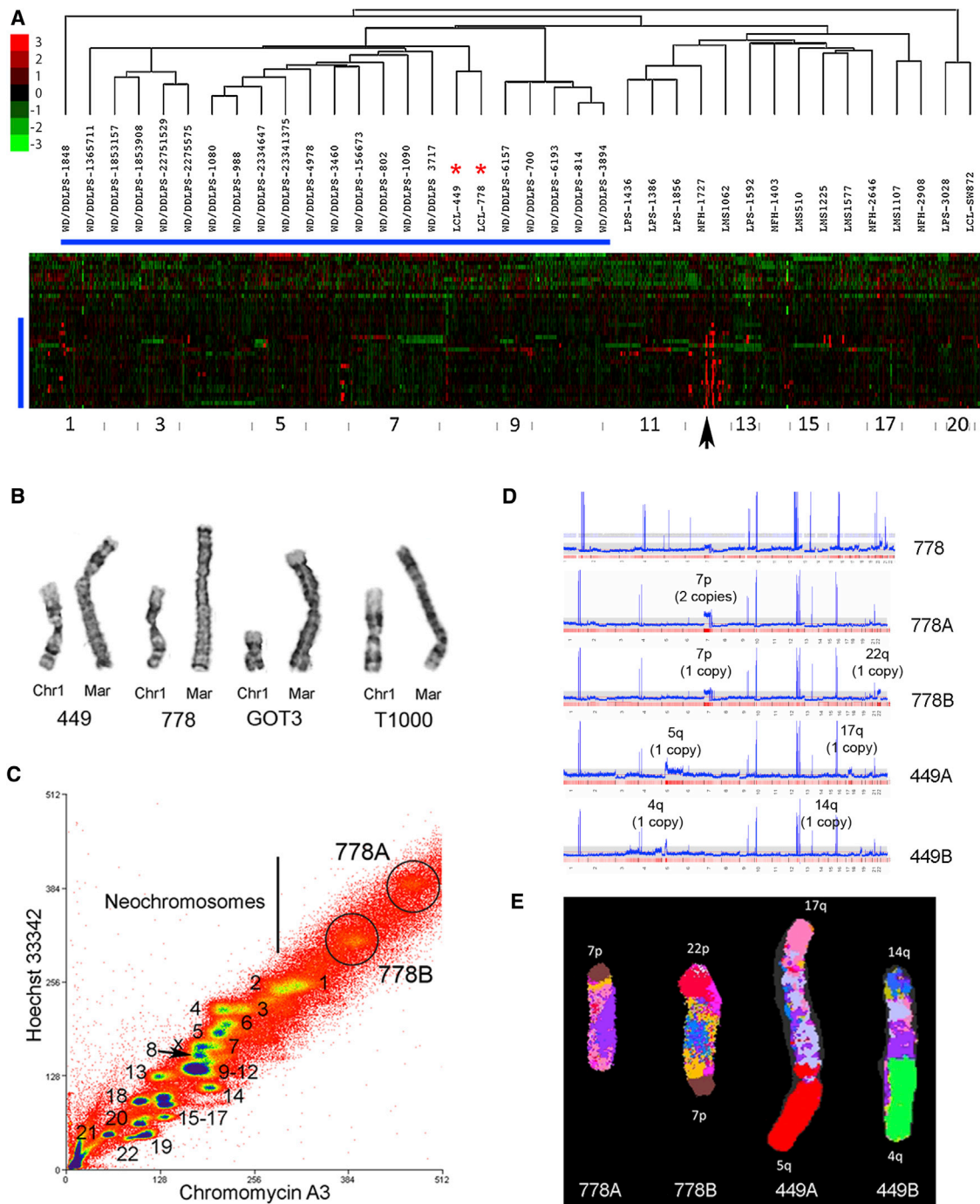
Figure 1. Characterization and Analysis of WD/DDLPS Cell Lines

(A) Hierarchical clustering of bacterial artificial chromosome array-based comparative genomic hybridization of 21 primary WD/DDLPSs, 5 pleomorphic liposarcomas (LPS), 4 undifferentiated pleomorphic sarcomas (MFH), and 5 leiomyosarcomas (LMS), as well as the 778 and 449 (indicated by stars) and SW872 cell lines (LCL). In the heatmap, rows represent samples, and columns represent probes. Beneath the dendrogram is the corresponding copy number heatmap (in landscape orientation). Heatmap color indicates the log_2 copy number ratio (normal, $log_2 = 0$); amplifications are red, and deletions are green. Blue bars demarcate WD/DDLPS samples and cell lines. The arrowhead marks the defining amplification of regions from chromosome 12q.

(B) G-banded marker neochromosomes from the 778, 449, GOT3, and T1000 cell lines, compared to chromosome 1 (Chr1).

(C) Representative flow dot plot of chromosomes from the 778 cell lines, showing the position of normal chromosomes as well as two marker neochromosome populations (circled). Chromosomes were stained for flow sorting with the fluorescent DNA-intercalating dyes as shown.

(legend continued on next page)

chromosomes, constitutive small marker supernumerary chromosome-like structures have been observed, often with developmental phenotypes (Liehr et al., 2006). Chromosomal abnormalities are also common in cancer. Double minutes, small, self-replicating extrachromosomal structures in a ring form, were originally observed in sarcomas (Buoen and Brand, 1968). Supernumerary accessory chromosomes (hereinafter referred to as neochromosomes) occur in 3% of all cancers, including 14% of mesenchymal tumors and 21% of glioblastoma (Dahlback et al., 2009; Sanborn et al., 2013). Neochromosomes are characteristic of several cancers, including well- and dedifferentiated liposarcoma (WD/DDLPS), parosteal osteosarcoma, and dermatofibrosarcoma protuberans (Garsed et al., 2009). Pathogenetically, the concomitant amplification of oncogenes such as *MDM2* and *CDK4* appears to drive carcinogenesis (Barretina et al., 2010; Italiano et al., 2008; Pedeutour et al., 1994; Pilotti et al., 1998).

Recently, new large-scale mutation patterns have been identified in cancer, including chromothripsis, kataegis, double minutes, and recurrent amplification of megabase regions of chromosome 21 (Li et al., 2014; Nik-Zainal et al., 2012; Sanborn et al., 2013; Stephens et al., 2011). Here, we anatomize cancer-associated neochromosomes, occurring in both circular form (ring) and linear form (giant markers) and possessing key functional structures in chromosomes, including centromere and telomeric domains (Lo et al., 2002a, 2002b; Marshall et al., 2008a). These features distinguish neochromosomes from double minutes, which are episomally replicating circular structures typically a few megabases in size and lacking recognizable centromeres and telomeres (Sanborn et al., 2013). Integrating structural mapping with mathematical simulations, we describe the evolution of these neochromosomes under oncogenic selection and, using transcriptional and copy number data from cell lines and primary tumors combined with functional studies, investigate a potential oncogene, *NUP107*.

RESULTS

Cytogenetic and Array-Based Mapping of WD/DDLPS Genomes

We first profiled isolated neochromosomes from five WD/DDLPS cell lines (449, 778, GOT3, T1000, and LPS141) (Pedeutour et al., 2012; Persson et al., 2008; Sirvent et al., 2000; Stratford et al., 2012). It is important to note that the 449 and 778 cell lines were independently derived from the same tumor (449 was established in 1993, and 778 was established from a recurrence in 1994; Sirvent et al., 2000), a feature enabling insights into conserved and divergent aspects of neochromosome evolution. The copy number profiles of the 449 and 778 cell lines recapitulate primary WD/DDLPS tumors, characterized by high-level chromosome 12 copy number gains (Figure 1A). Initial karyotyping (Figure 1B) and multicolor fluorescence in situ hybridization

(FISH) studies of all cell lines (Figure S1 available online) indicated two to four linear neochromosomes per cell, all significantly larger than chromosome 1. The circular forms of neochromosomes commonly reported in primary tumor cultures were not seen (Pedeutour et al., 1994; Sirvent et al., 2000).

Size-based fluorescence-activated cell sorting separated and purified the neochromosomes from other chromosomes for sequencing studies (Figure 1C). The estimated enrichment of the flow-isolated neochromosomes typically varied from 6-fold for the 778 neochromosomes to 12-fold for the 449, and to 65-fold for the GOT3 neochromosomes (see Supplemental Experimental Procedures for additional details), allowing detailed analysis of these structures isolated from the remaining genome. LPS141 was an exception, with 2-fold enrichment. The 778 and 449 neochromosomes were analyzed using Affymetrix 250k SNP arrays (Figure 1D). Comparison of whole-genome copy number (778 in Figure 1D) to the isolated neochromosomes (778A, 778B, 449A, and 449B) indicated that almost all highly amplified regions are located on the neochromosomes. Striking conservation was observed of regions of high-level gain between neochromosomes from the 778 and 449 cell lines, despite their independent derivation and ex vivo culture, attesting to the primary role of the amplified core of the neochromosome in tumor development.

In contrast to amplified regions, major differences were notable between neochromosome isoforms in low-copy regions, including chromosome 7p (778A and 778B), chromosome 22 (778B), chromosome 5 (449A), chromosome 17 (449A), and chromosome 4 (449B) (Figure 1D). Multicolor FISH studies revealed that these regions are located at the telomeres in all cell lines (Figure 1E and Figure S1). Telomeres are an obligate feature of mammalian linear chromosomes. The variation between the 449 and 778 neochromosome isoforms suggests that telomeres are acquired as secondary and independent events arising after the establishment of the highly amplified core, perhaps by telomere capture (Meltzer et al., 1993). Individual metaphases showed marked variation within and between cell lines in neochromosome number and telomere patterns, consistent with chromosomal instability (Figure S1). Karyotypic analyses of the 778 cell line revealed at least two copies of the giant chromosomes per metaphase. Both the T1000 line and the GOT3 line had 45–47 chromosomes per metaphase and showed a large, variable neochromosome consisting mainly of material from chromosomes 1, 12, and X. In the LPS141 cell line, at least two populations of neochromosomes were seen, with the telomeres involving chromosomes 8 and 9, and perhaps chromosomes 2 and 3.

High-Resolution Architecture of WD/DDLPS Neochromosomes

DNA from flow-isolated neochromosomes (shown for the 778 cell line in Figure 1C) was amplified and sequenced, along with

(D) SNP array copy number analysis of the total DNA from the 778 cell line (778), as well as four flow-sorted neochromosome isoforms (449A, 449B, 778A, and 778B). Chromosomes 1 to X are represented on the x axis, while copy number is shown on the y axis. The large low-copy regions distinguishing neochromosome isoforms are labeled.

(E) Multicolor FISH staining of neochromosomes from the 778 and 449 cell lines, demonstrating homogenously staining core regions and distinctive telomeric regions captured from the labeled native chromosomes, corresponding to the SNP data in (D).
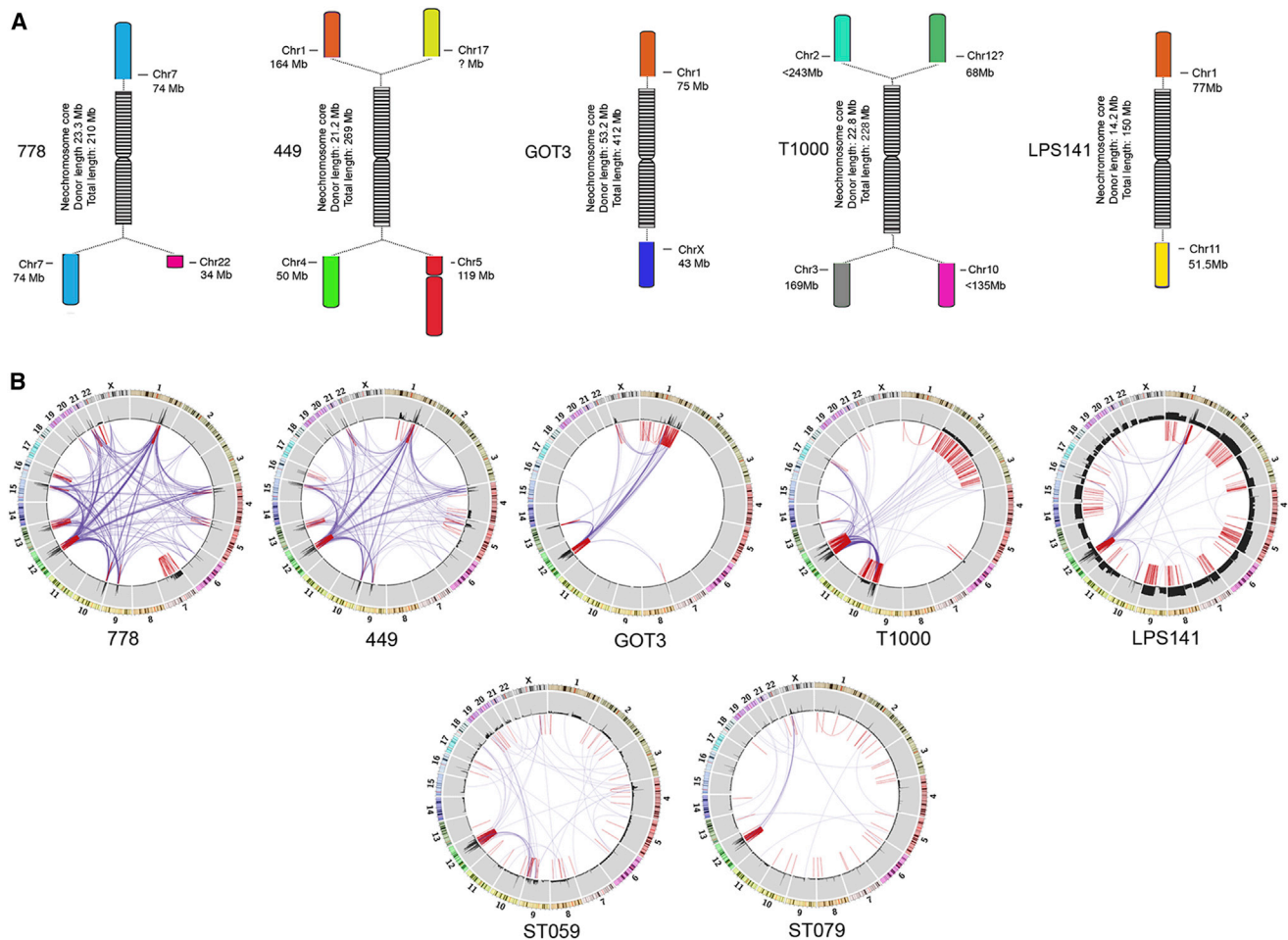
See also Figure S1.

Figure 2. Neochromosome Architecture
(A) Schematic of the structure of the 778, 449, GOT3, T1000, and LPS141 neochromosomes. Chr, chromosome.
(B) CIRCOS diagrams showing copy number profiles and intrachromosomal (red) and interchromosomal (purple) fusions for the 778, 449, GOT3, T1000, and LPS141 cell line neochromosomes, along with two primary tumors (ST059 and ST079).
See also Figure S2 and Tables S1 and S2.

the whole genomes of two primary WDLPS tumor samples, ST059 and ST079. Isolated neochromosome isoforms from each cell line were pooled. In summary, between 7 (449) and 49 Gb (T1000) of DNA sequence was generated for each cell line or primary tumor, with 32–382 Mb at greater than 20× (Table S1). Coverage of single-copy regions varied from 5× (449) to 27× (GOT3), with much greater coverage of amplified regions.

Our initial aim was to define precisely the size and structure of the neochromosomes (Figure 2A). Although lacking recurrent single nucleotide or insertion and deletion (or INDEL) variation of significance (Supplemental Experimental Procedures), extreme structural rearrangement and copy number variation was apparent (Figure 2B). Copy number was estimated by counting aligned reads in 5 kb windows, correcting for guanine and cytosine (GC) content and background and calibrating using regions of known single copy (for example, chromosome 22 for 778B; Figure S2A). Thresholding copy number (Figure S2B) and integrating fusion information identified neochromosomal material (see Supplemental Experimental Procedures for de-

tails). The reference donor size for the five cell line neochromosomes were only 14 Mb (LPS141), 21 Mb (449), 23 Mb (778), 23 Mb (T1000), and 53 Mb (GOT3), derived unequally from different chromosomes. The small size of the donor genomes was amplified to between 150 Mb (LPS141) and 412 Mb (GOT3; Figure 2A). Consistent with stability of the core, there was 91% overlap between the 449 and 778 neochromosomes. Including telomeric regions, the total size of the isolated neochromomes varied from 279 Mb (LPS141) to at least 640 Mb (T1000). The GOT3, 449/778, T1000, LPS141, ST059, and ST079 neochromosomes share in common 1.4 Mb of donor sequence from chromosome 12, including the following genes: *MDM2*, *NUP107*, *HMGA2*, *CDK4*, *CPM*, *RAB3IP*, *FRS2*, *MARCH9*, *AC025263.3*, *CCT2*, *TSPAN31*, *CYP27B1*, *CNOT2*, *SLC35E3*, *AGAP2-AS1*, *AGAP2*, *AC124890.1*, *PTPRR*, *KCNMB4*, *METTL1*, *LRRC10*, *MYRFL*, *PTPRB*, and *BEST3*.

Synoptically, neochromosomes comprise hundreds of highly amplified and rearranged fragments of DNA from every chromosome (Figure 2B). High-level copy number changes are

extremely focal and show a nonlinear distribution of copy number states (Figures S2A and S2C). For example, in 778, quantized copy number clusters equivalent to 1, 2, 4, 8, 16, and 32 copies were observed across different chromosomes of origin, with single-copy states for chromosome 22 and a mixture of single and double- or triple-copy states for chromosome 7. The modal copy number state on chromosome 12 was ∼16 copies, with a maximum of 64 copies. Spatially, clusters of anomalously aligned paired end reads were identified (see Supplemental Experimental Procedures), yielding between 256 (449) and 586 (778) fusions, involving 6–18 different chromosomal partners (Table S2). The 449 and 778 neochromosomes shared 70% of fusions, with the gap likely attributable to differential read depth (data not shown). Between 23% and 86% of fusions were intrachromosomal. Forty-four of 50 (88%) randomly selected representative fusions from the 778 neochromosome data were validated by conventional PCR and sequencing. A detailed analysis of the 778 neochromosome breakpoints showed features of blunt end joining, or imperfect homology (<20 base pairs [bp] of homology) (data not shown). Notably, many nonneochromosomal translocations are also observed in our cell lines (Figure S1).

Despite the complexity, clear patterns were discernible. Neochromosomes comprise hundreds of spatially distinct, contiguous genomic regions (CGRs), illustrated by the CGR containing the oncogene *CDK4* (58,048–58,167 kb on 778 chromosome 12; Figures 3A and 3B; Table S3). CGRs are islands in the reference genome, flanked by regions effectively absent in the flow-isolated neochromosome sequence data. The absence of signal outside CGRs is not an artifact of read depth, since known single-copy regions (e.g., chromosome 22) are readily detected at 5×–27×. CGRs range from 88 bp to 1.15 Mb (median, 26 kb), and are highly rearranged and amplified. These rearrangements generate two types of breakpoints (Figure 3B). The first defines the edges of CGRs (e), with contiguous reads mapping on only one side of the breakpoint. The second type lies internal to CGRs (i) and is characterized by reads flanking both sides of the breakpoint. These breakpoints give rise to three types of fusion (Figure 3C). Fusions between the boundaries of two CGRs are termed "edge-to-edge" (e2e) fusions. e2e fusions linking CGRs are probably early events, arising at, or soon after, incorporation of new material into the neochromosome. A second fusion type links a break internal to a CGR to the edge of another CGR (termed "edge-to-internal," or e2i), and the third type ("internal-to-internal," or i2i) links two CGRs internally.

The extraordinarily complex topology of the neochromosome is shown by a graph of CGRs in the 778 neochromosome, with nodes whose radius is proportional to maximal copy number, and fusion types as edges (Figure S3A). A total of 177 e2e, 203 e2i, and 166 i2i fusions were identified within the 778 neochromosome core. The highest copy CGRs are typically linked within a tight network of i2i and e2i fusions, suggesting that these fusions are related to amplification.

Different types of fusions have distinct properties. Analysis of e2e fusions indicates that the neochromosome probably formed from chromosome 12, with or without other contributory chromosomes. For example, in the 778 neochromosome, although chromosome 12 contributes less than 30% of the amplified core of the neochromosome, 75% of chromosome 12 e2e

fusions are intrachromosomal (Figure S3B). For comparison, 23% of chromosome 1 e2e fusions and 7% of chromosome 15 e2e fusions are intrachromosomal, despite representing 19% and 12% of the neochromosome content, respectively. Moreover, 19 of the top 25 (76%) most highly amplified e2e fusions involve at least one partner on chromosome 12, and both partners come from chromosome 12 in 15 cases (60%). By contrast, 778 chromosome 12 e2i and i2i fusions involve diverse partners, including itself (26%), chromosome 1 (12%), and chromosome 15 (9%). Assuming that e2e fusions primarily represent early events, identifying "walkable" chains of e2e fusions may allow insight into the early structure of neochromosomes. Continuous walkable chains of CGRs entirely located on chromosome 12 were identified in several samples, with the longest chain of chromosome 12 CGRs connecting 22 of 50 CGRs in the 778 neochromosome (Figure 3D). In T1000, a mixture of CGRs from chromosomes 12 and 10 form a similar set of walkable chains, similar to another chain involving two distinct chromosomes (chromosomes 13 and 16) in the 778 neochromosome (Figure S3C), and suggesting the synchronous engagement of two distinct chromosomes. The average read depth for chromosome 12 e2e fusions is 165 ± 12, compared to 76 ± 10 for chromosome 1, the next most abundant chromosome. The earliest CGRs incorporated into the neochromosome likely are the most highly amplified, consistent with the early involvement of chromosome 12.

e2e fusions are characterized by equal copy numbers on each side of the fusions (average copy number ratio, 1.02 ± 0.03; range = 0.6–1.6 in 778; Figures S3D and S3E), a likely consequence of two CGRs introduced into the neochromosome simultaneously without amplification. By contrast, e2i fusions typically comprise asymmetric copy number states (for example, in 778, the ratio of "e" side to "i" side averages 0.32 ± 0.02), although the range (0.02–1.2) includes copy-number-neutral fusions. i2i fusions are characterized by a broad range of copy number ratios (1–16). In summary, e2i and i2i events appear dominated by progressive amplification and deletion events but also include some early copy-number-neutral events. The distribution of fusion types, their copy number states, and the presence of walkable chains of e2e fusions likely reflects the order of events during the life history of the neochromosome. The data suggest an order of events in which chromosome 12 fusions demonstrating copy number neutrality across the breakpoints (typically but not exclusively e2e) most likely arise during initial assembly or soon thereafter, while fusions involving marked differences in copy number states on either side of the breakpoint (again typically but not exclusively e2i and i2i) are generated during subsequent amplification and deletion events on the neochromosome.

Modeling the Evolution of WD/DDLPS Neochromosomes

We sought to understand in greater detail the order of events leading to neochromosome formation, starting with the initial formation of an episomal, self-replicating structure. Double-minute episomal structures have been observed in the context of chromothripsis (Sanborn et al., 2013; Stephens et al., 2011). Chromothripsis is a temporally and spatially limited phenomenon characterized by chromosomal shattering and rearrangement, without significant copy number change (Korbel and Campbell,
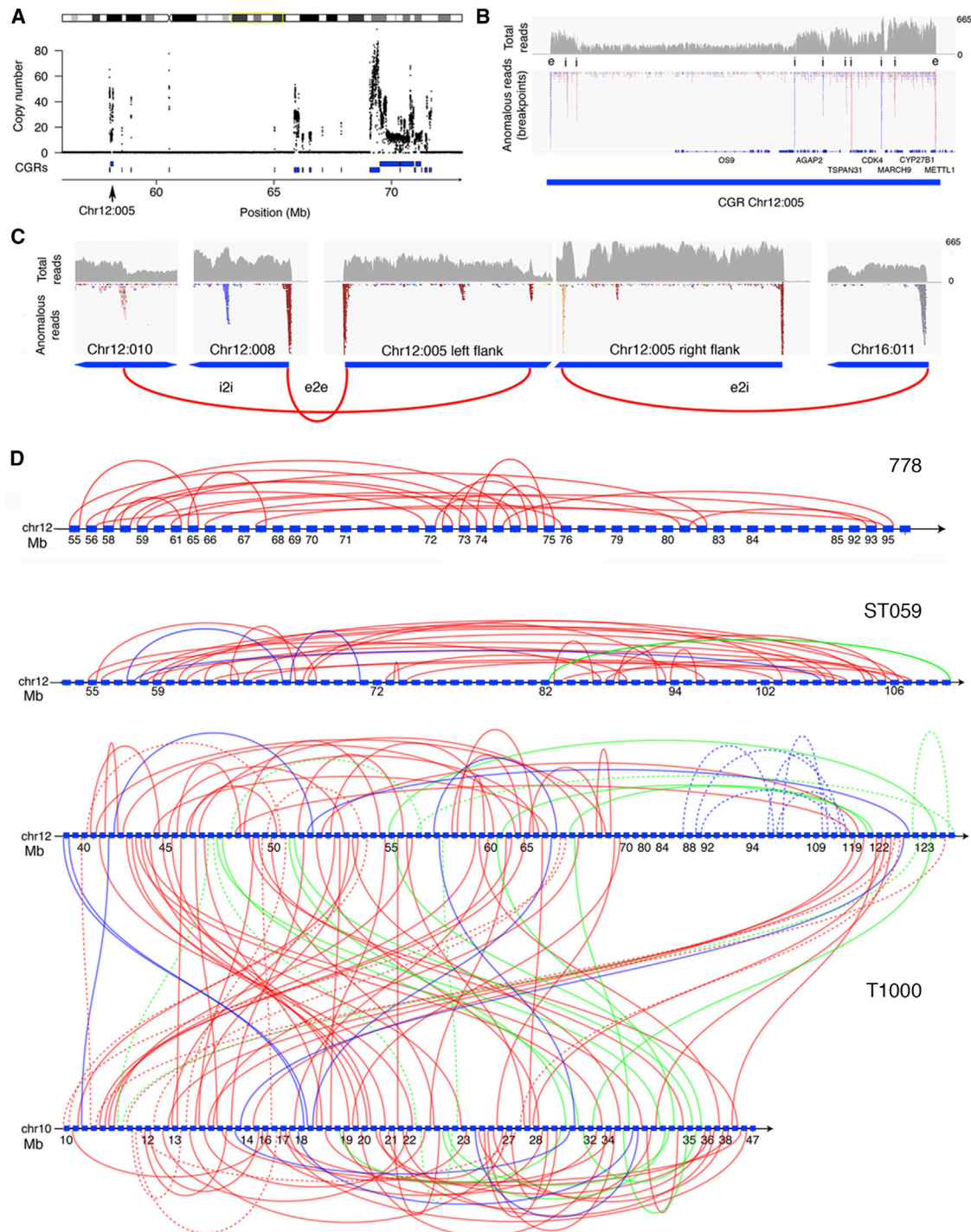
Figure 3. Classification of Breakpoints and Fusions in CGRs on the Neochromosome

(A) Copy number profile from 778 showing adjacent CGRs containing the oncogene *CDK4*. Chr, chromosome.

(B) Higher resolution view of CGR at chromosomes 12:58.070-58; 165 Mb showing total read depth (top) and anomalously aligned reads (bottom). Edge breakpoints (e) define the boundaries of the CGR, while internal breakpoints (i) correspond to rearrangements linking CGR internally to noncontiguous regions of the genome. Note that the flanking regions on either side of the CGR contain only a low level of reads consistent with the absence of representation on the neochromosome.

(C) The three classes of fusion: e2e, e2i, and i2i.

(D) The longest chains of e2e fusions between CGRs (blue rectangles) involving chromosome 12 on the 778 and ST059 neochromosomes and involving both chromosomes 12 and 10 on the T1000 neochromosome. These represent the inferred structure of the original chromothriptic rearrangements.

See also Figure S3 and Table S3.

2013; Stephens et al., 2011). Consistent with chromothripsis, the walkable chromosome 12 e2e chains noted earlier typically do not show marked copy number differences on either side of the fusions (Figure 3D; Figures S3D and S3E). Other classical features of chromothripsis affecting chromosome 12 seen in all samples include clustering of breakpoints, monoallelic amplification, and broadly equal numbers of deletion-type, tandem duplication-type, and head-to-head and tail-to-tail intrachromosomal fusions (Korbel and Campbell, 2013) (Figure 4A, right panel; Figures S4A and S4B). Chromothripsis is also suggested by, for example, oscillating copy number states in the low-copy regions of chromosomes 7 and 22 in the 778 neochromosomes (Figure 4B), the telomeric fragment of chromosome 1 in the GOT3 neochromosome (Figures 2B and S2A), and multiple regions in LPS141 (Figure S2A). Taken together, these features suggest that chromothripsis occurs early in the life cycle of the neochromosome and may later account for subsequent episodic integration of new material (see Supplemental Experimental Procedures and Table S4 for further details).

By definition, chromothripsis cannot account for high-level copy number variation. Breakage-fusion-bridge (BFB) has been linked to gene amplification in WD/DDLPS (Gisselsson et al., 2000; Lo et al., 2002a). The dominant model for BFB assumes a linear form, in which telomeric association and fusion of homologous chromosomes form a dicentric chromosome that breaks during each mitosis, leading to asymmetric segregation (Gisselsson et al., 2000). However, two features led us to postulate a circular form of BFB (Figure 4C). First, the circular neochromosomes frequently seen in primary WD/DDLPS karyotypes are subsequently replaced by linear forms in cell lines, where we observe stabilization of the core content of the neochromosome (discussed earlier). More important, the inverted duplications that are characteristic of linear BFB are not observed in any neochromosome, and the proportions of different fusion types (deletion type, tandem duplication type, and inversion type with head-to-head or tail-to-tail orientation; Korbel and Campbell, 2013), were uniform (Figures S4A and S4B).

Modeling circular BFB without chromothripsis over several thousand cell cycles, beginning with chromosome 12 (see Supplemental Experimental Procedures for details), qualitatively recapitulated the focal and discretized amplification, as well as many features of e2e, e2i, and i2i fusions (Figures S4C and S4D). However, BFB alone does not generate the inversions observed empirically (Figure 4A, right panel; Figures S4A–S2C and S4E). Moreover, even after several thousand cycles, the number of chromosome 12 CGRs consistently reached an asymptote <30, less than that empirically observed in all isolated neochromosomes (50–90 CGRs; Figure 4D, left panel). Finally, BFB alone fails to generate empirically observed CGR sizes (Figure S4F) or sustained walkable e2e chains of chromosome 12 CGRs (Figure S4G).

By contrast, incorporating chromothripsis (modeled by randomly fragmenting chromosome 12q using the fragment length distribution empirically observed in the chromothriptically rearranged region of chromosome 7 in the 778 neochromosome) into our model consistently recapitulated important observed phenomena, including the number of chromosome 12 CGRs in all lines (Figure 4D, right panel); eq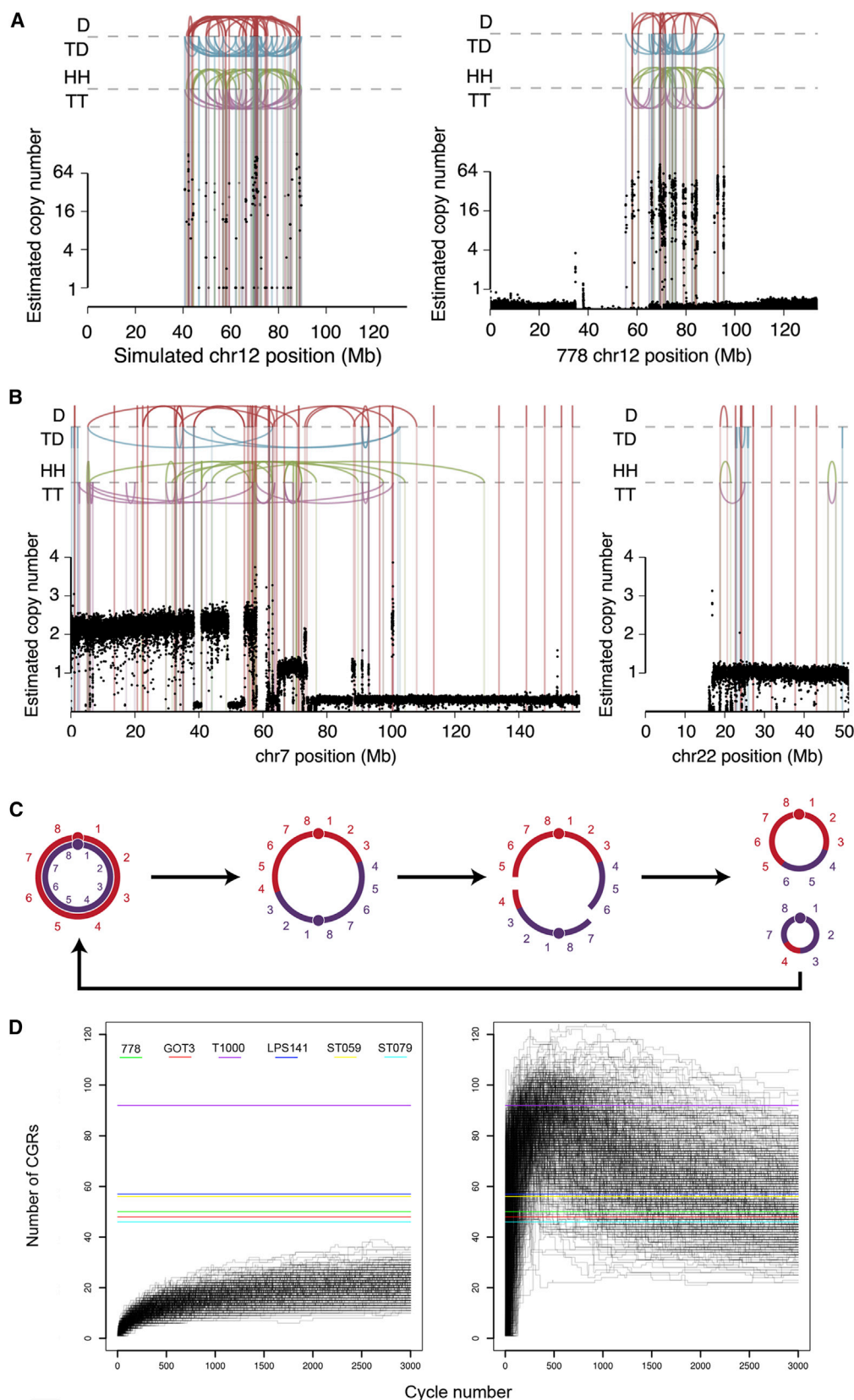ual proportions of fusion orientations (Figure 4A, left panel; Figure S4H); copy number ratios for e2e, e2i, and i2i fusion types (Figure S4I); sustained walkable chains of chromosome 12 CGRs linked by e2e fusions (Figure S4J); total number of e2e fusions (Figure S4K); and the median CGR size (Figure S4L; Movie S1). Together with empiric evidence for chromothriptic events at other sites in the genome, these data are consistent with a model in which chromothriptic rearrangement of chromosome 12 is involved early in the formation of the ring chromosome, perhaps as the initiating event.

Notably, these models depend on the application of heuristic locus-specific selective values, without which simulations invariably resulted in complete loss of the neochromosome, suggesting that any regions recurrently amplified in primary tumors depend on a selective advantage of some kind to the evolving cancer clone. A complete summary of the assumptions used in developing this model is provided in the Supplemental Experimental Procedures.

Centromeric Corrosion by BFB

Large-scale loss of material, a feature of both BFB and chromothripsis, also has implications for key structural domain of chromosomes. Normal human centromeres comprise highly repetitive α-satellite DNA. WD/DDLPS neochromosomes commonly possess neocentromeres arising on nonrepetitive regions (Marshall et al., 2008a), observed as centromeric constrictions in the absence of alphoid DNA. For example, the 778 line neochromosomes only possesses nonalphoid neocentromeres (Figure S5). To fully characterize the neocentromere on the 778 neochromosomes, centromere protein A (CENP-A) chromatin immunoprecipitation and sequencing (ChIP-seq) was performed in the 778 cell line. Compared to either the input or a nonspecific control, only two regions on the neochromosome were significantly enriched in the CENP-A pulldown, constituting a 77.5 kb region in chromosome 1q25.3 and an 8.5 kb region in chromosome 6q21 (Figure 5A). Both regions are linked by an e2i fusion, suggesting that the neocentromere assembled on a single, contiguous domain from both chromosomes 1 and 6. Confirming the ChIP-seq data, a bacterial artificial chromosome (BAC) from chromosome 1 that spans the 77.5 kb binding domain colocalized with CENP-A at the centromere of the neochromosomes via immuno-FISH (Figure 5B). Since the regions from chromosomes 1 and 6 that comprise the 778 neocentromere were not involved in the initial formation of the ring (discussed earlier), neocentromere formation appears to be a late acquisition during the evolution of the neochromosome.

We next looked for examples of neochromosomes that contained alphoid centromeres. The GOT3 cell line (Figure 5C) possesses a faint active alphoid centromere located within the oncogenic amplified material. This alphoid centromere, derived from chromosome 1, had lost 64% of the α-satellite-specific repeats, as determined by quantitative FISH (p < 0.001, n = 3; Figure 5Da). Similarly, an inactive, attenuated centromere derived from chromosome 5 was observed on the arms of two of the four neochromosomes present in the 449 cell line. This centromere had lost 78% of its satellite repeats, as determined via quantitative FISH (p < 0.001, n = 7; Figure 5Db), retaining single-copy adjacent pericentric DNA from chromosome 5p on both neochromosomes (Figures 5Ea and 5Eb). Additional single-copy nonalphoid material from the 5p pericentric region

(legend on next page)

was observed on a third neochromosome from this cell line (Figure 5Ec). Combining the FISH probe and sequencing data, the multiple breakpoints mapped on each neochromosome isoform (Figure 5F) provides evidence of progressive states of attrition and corrosion of α-satellite and surrounding pericentric DNA. Assuming that the centromere is eventually lost through BFB, it may be replaced by neocentromere formation utilizing nonalphoid regions as seen in the 778 cell line.

Transcriptional Consequences of Genomic Rearrangement on the Neochromosome

In integrating data from the 449/778, GOT3, and T1000 cell lines, we found that 107 coding units were highly amplified (>10 copies) in all cell lines (Table S5). Fifty-two represented curated protein coding transcripts, as well as 7 microRNAs, 11 spliceosomal or small nucleolar RNAs, 14 pseudogenes, 8 large, intergenic noncoding RNAs, and several less well-curated transcripts. Transcriptome sequencing of the 449/778, GOT3, and T1000 cell lines was performed and correlated with copy number (Table S5). Across the 449/778, GOT3, and T1000 cell lines, the average reads per kilobase million (RPKM) for genes with a median copy number >10 was 38, compared to an average RPKM of 10 for genes with a median copy number between 2 and 10, and an average RPKM of 3 for genes with a median copy number of 2 or less. Among the most abundant and curated protein coding transcripts are *MDM2* (a median of 281 RPKM), *OS9* (243 RPKM), *CDK4* (164 RPKM), *YEATS4* (85 RPKM), and *NUP107* (77 RPKM). However, some amplified genes are transcriptionally silent. For example, *FASLG* (chromosome 1: 172,628,185–172,636,012), an important proapoptotic regulator of the extrinsic death pathway, is frequently amplified (778 cell line, 36 copies; 449 cell line, 22 copies; and GOT3 cell line, 11 copies) but expressed at fewer than 1 RPKM in any line. Detailed analysis of the 778 cell line revealed that promoter methylation within amplified regions correlated with lower expression (mean ± SE: 0.46 ± 0.05 for genes with low RPKM versus 0.18 ± 0.05 for genes with high RPKM, p < 0.001). The average promoter CpG methylation of *FASLG* in the 778 cell line was 0.8, compared to 0.12 for *CDK4* and 0.11 for *MDM2*. These data suggest that epigenetic silencing is permissive for amplification of loci containing genes whose overexpression would otherwise not be tolerated. Interestingly, where a breakpoint dissociates the native promoter from the coding unit of a gene, promoter methylation correlated poorly with gene expression (r^2 = 0.01, p = 0.61), compared to intact genes (r^2 = 0.28, p < 0.001), consistent with disruption of normal mechanisms of regulatory control.

Massive rearrangement of the neochromosome may affect transcript splicing or generate fusion transcripts. By integrating predicted breakpoints and fusion transcripts identified using RNA sequencing data (see Supplemental Experimental Procedures for methods), we identified 0 (449) to 4 (778) putative fusion genes (data not shown). No recurrent fusion genes were found. Interestingly, in 778, the head-to-head or tail-to-tail orientation of these putative fusion genes indicates that they are unlikely to constitute oncogenic drivers; however, transcripts using cryptic exons located in intronic or intergenic regions were also identified. For example, a reciprocal translocation between *YEATS4* (chromosome 12) and *TAF3* (chromosome 10) results in a highly expressed exon in the antisense orientation located within the second intron of *TAF3* (Figure S6), upstream of a tail-to-tail fusion with the sixth intron in *YEATS4*. *YEATS4* was recently identified as an amplified oncogene in WD/DDLPS (Barretina et al., 2010). The resulting fusion was verified by direct sequencing of the predicted transcript, replacing a coiled-coil domain encoded in exon 7 of *YEATS4*. Notably, alternate transcripts were identified that introduced a premature stop codon after exon 6 of *YEATS4*, also deleting the coiled-coil domain. Since the C terminus coiled-coil domain of YEATS4 elicits a p53 response, this may provide a selective growth advantage for WD/DDLPS (Park and Roeder, 2006).

Functional Analysis

The model predicts that recurrently amplified regions must confer a selective advantage on the evolving cancer cell; in some cases, by oncogene overexpression. To identify putative oncogenes, we undertook a census of amplified and over-expressed genes in 131 primary connective tissue tumors, including 15 WD/DDLPS tumors (D.W.G. and D.M.T., unpublished data). The most frequently amplified genes in WD/DDLPS (>70%) are located at 12q14.1-15 (Table S6). Of the 49 recurrently amplified genes (>30% of WD/DDLPS) located on chromosome 12, there were 16 genes whose expression levels were higher in WD/DDLPS compared to other sarcoma subtypes (>1.5-fold; p < 0.001). We further refined the list of candidate driver genes by defining the subset of genes that are also amplified (copy number > 20) and overexpressed (RPKM > 100) in three WD/DDLPS cell lines (449, GOT3, and T1000; Table S5; (Figure 6A; Table S7). To determine whether these candidates are required for viability and growth, we performed small interfering RNA (siRNA) knockdown on the top ranked four genes (*OS9*, *CDK4*, *MDM2*, and *NUP107*) in two cell lines with copy number profiles representative of primary tumors (449 and 778). The effect on cell growth was measured

---

Figure 4. Evidence Supporting Chromothriptic Origins of WD/DDLPS Neochromosomes
(A) A selected copy number and fusion orientation profile from the BFB + chromothripsis model after 1,000 cycles (left) compared to the observed profile for chromosome 12 (Chr12) from the 778 neochromosome (right).
(B) Copy number profiles and different class of intrachromosomal rearrangement (D, deletion type; TD, tandem duplication type; and inversion types: HH, head-to-head and TT, tail-to-tail) of the telomeric caps of the 778 neochromosomes—chromosomes 7 (Chr7, left) and 22 (Chr22, right).
(C) Model of circular BFB: (left to right) the circular neochromosome replicates and enters a BFB cycle if an odd number of crossover events occurs; in this event, a dicentric neochromosome is formed; at cell division, the centromeres are pulled apart and two double-strand breaks occur; daughter cells inherit unequal amounts of DNA containing duplicated and deleted regions, which we score for fitness using a heuristic model.
(D) Evolution of the number of CGRs generated by the BFB-only (left) and the BFB + chromothripsis (right) models. Each black line follows the evolution of a possible lineage. Only one cell is followed at each cycle, selected at random with probability biased toward the fittest. Color lines indicate the number of chromosome 12 CGRs observed in 778 (green), GOT3 (red), LPS141 (blue), T1000 (purple), ST059 (yellow), and ST079 (cyan).
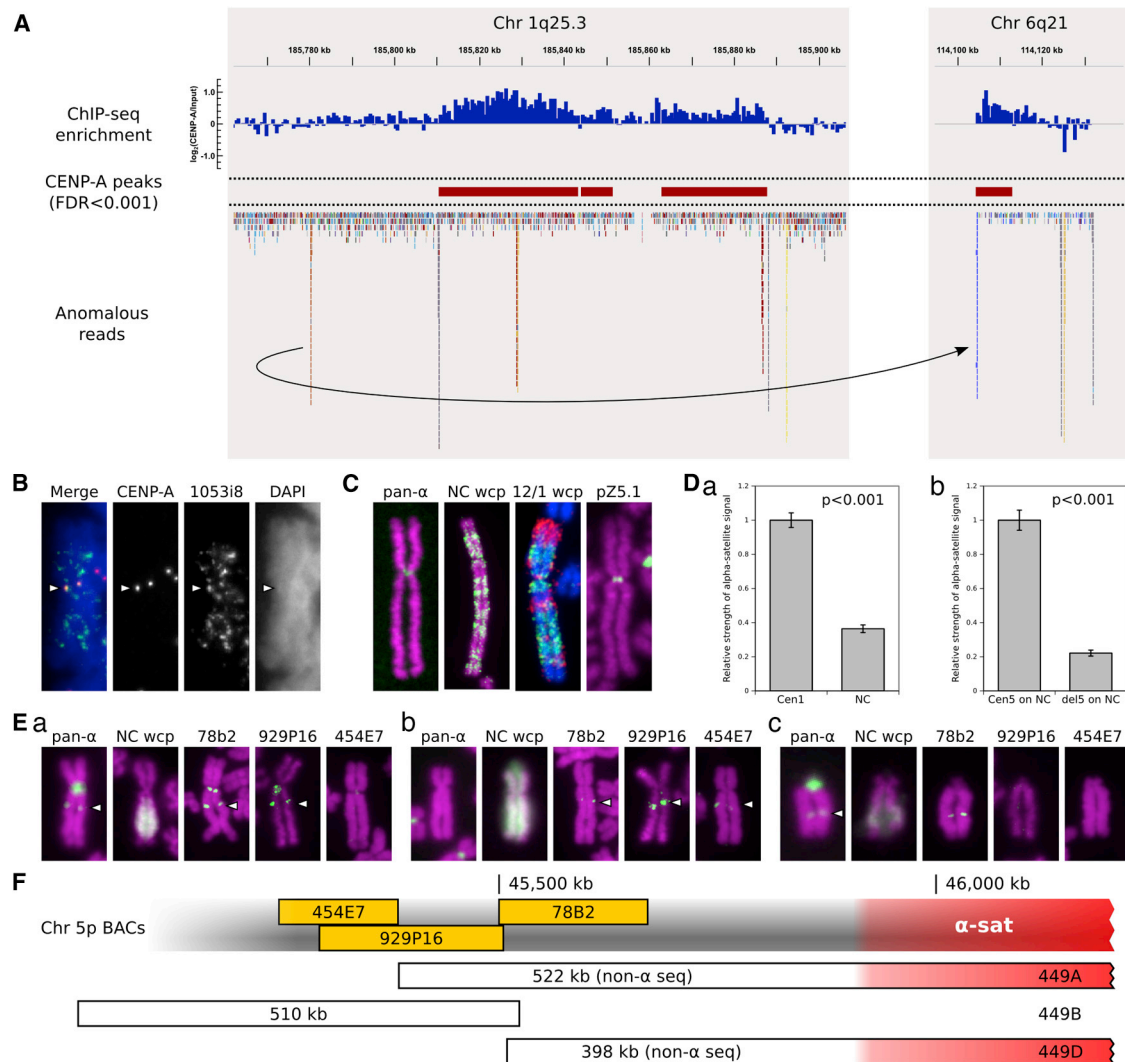See also Figure S4, Table S4, and Movie S1.

Figure 5. Characterization of WD/DDLPS Centromeres and Neocentromeres

(A) CENP-A ChIP-seq on the 778 cell line revealed a single neocentromere forming on linked regions from chromosomes 1 (Chr1) and 6 (Chr6). Log$_2$ ratio of CENP-A pulldown over input is shown in blue. CENP-A binding peaks (false discovery rate [FDR] <0.001) are demarcated in red. Anomalous reads indicating an e2i fusion between chromosomes 1 and 6 is shown. Black line with arrowhead indicates the fusion and orientation.

(B) Combined immuno-FISH using an anti-CENP-A antibody and the RP11 BAC 1053i8 as a FISH probe demonstrates that the region from chromosome 1q25.3 identified by ChIP-seq colocalizes with CENP-A at the 778 neocentromere.

(C) FISH studies of GOT3 neochromosome reveal a faint alphoid centromere (pan-α) present within amplified material (whole chromosome paint from 778 neochromosomes; NC wcp). Cohybridization with wcp for chromosomes 12 (green) and 1 (red) shows that the centromere was originally derived from chromosome 1; centromeric hybridization of the chromosome 1/5/19-specific satellite probe pZ5.1.

(D) Quantitative FISH of α-satellite sequences within the amplified core of the GOT3 (a) and 449 (b) neochromosomes demonstrates loss of satellite repeats compared to native centromeres. Error bars represent SEM.

(E) Three 449 neochromosomes (a-c) show an inactive centromere derived from Chr5 (pan-alphoid probe), the amplified core (NC wcp), and adjacent pericentric material from Chr5p. Labels refer to the FISH probe (green), counterstained with DAPI (magenta). Where two FISH probes colored green/red are used, DAPI counterstain is in blue. BAC numbers refer to the RP11 library.

(F) Pericentric breakpoints identified by BAC hybridization and sequencing for each 449 neochromosome.

See also Figure S5.

by MTS assay after 3 to 5 days (Supplemental Experimental Procedures; Table S7). Consistent with the principle that genes recurrently amplified in WD/DDLPS confer a selective advantage, the median decrease in cell viability following knockdown was 31.78% (range = 0.23%–84.8%), compared to control siRNA.

This study confirms the prevalence of two known oncogenes (*MDM2* and *CDK4*) (Italiano et al., 2008; Pedeutour et al., 2012; Pilotti et al., 1998), and identified a previously noted amplified gene (*OS9*; Su et al., 1996). *NUP107*, a nucleoporin overexpressed in cancer (Agudo et al., 2004; Banerjee et al., 2010), was also identified. Located 65,499 bp 5′ of *MDM2*, *NUP107* is
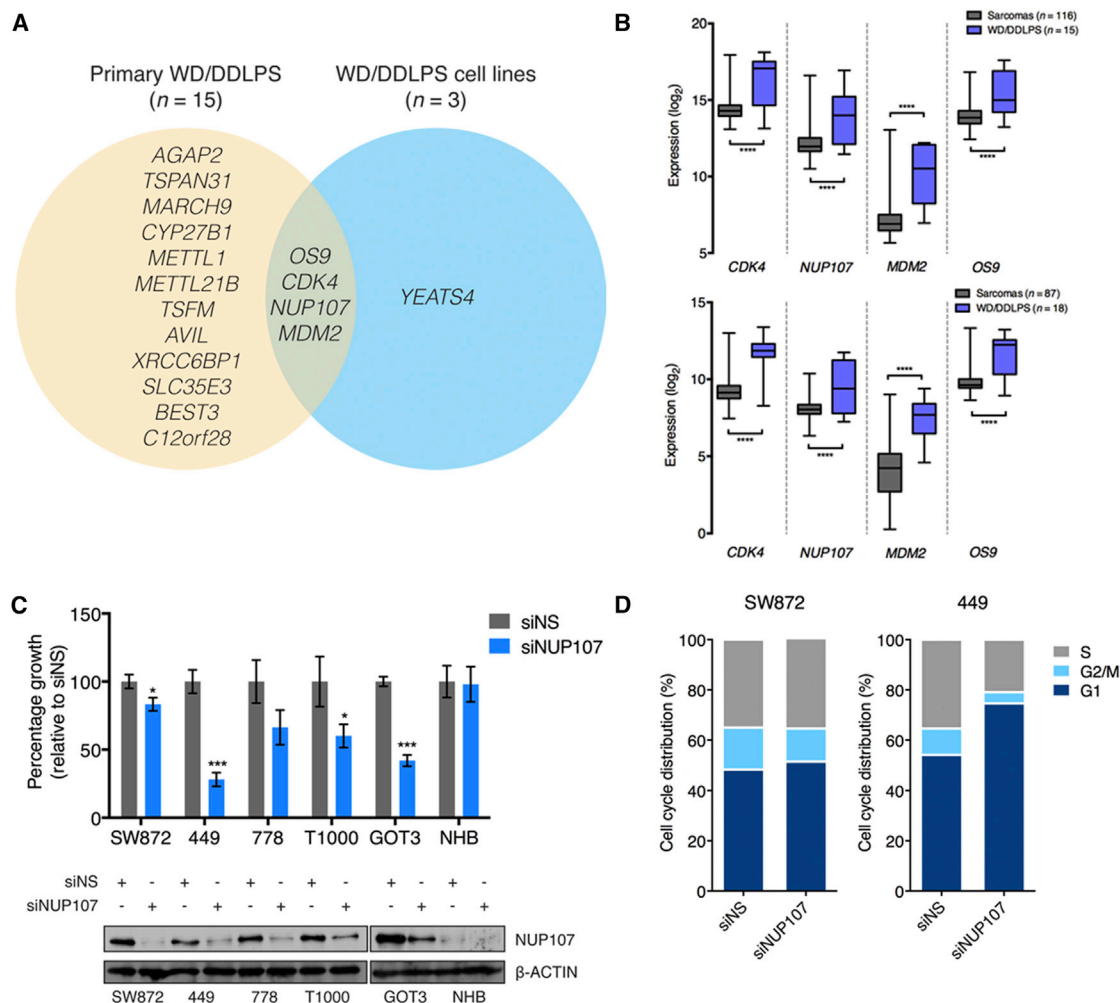
Figure 6. Screen for Candidate Driver Genes in WD/DDLPS

(A) Venn diagram showing chromosome 12 genes that are recurrently amplified and overexpressed in WD/DDLPS primary tumors and cell lines. Genes selected were frequently amplified in primary WD/DDLPS (>30%), and are expressed more than 1.5-fold higher in WD/DDLPS (n = 15) compared to a panel of other sarcoma subtypes (n = 116; corrected p value < 0.001). Candidates were further filtered to identify chromosome 12 genes that are highly amplified (median copy number >20) and expressed (median RPKM >100) in three WD/DDLPS cell lines (449, GOT3, and T1000).

(B) Top: relative expression of *CDK4*, *MDM2*, *NUP107*, and *OS9* in 15 WD/DDLPS samples compared to a pool of 116 other sarcomas. Bottom: the same analysis using an independent set of 105 sarcomas, including 18 WD/DDLPS samples (National Center for Biotechnology Information Data Set Browser record GDS2736; Nakayama et al., 2007). Box-and-whisker plots show $\log_2$ gene expression distribution (**** $p < 0.0001$, two-tailed t test).

(C) Effect of NUP107 depletion using siRNA in *NUP107* amplified cell lines (449, 778, T1000, and GOT3) and unamplified cells (SW872 and NHB). Non-silencing siRNAs (siNS) or pooled siRNAs were used as indicated, and viability was assessed by MTS assay 3 to 5 days after transfection (mean ± SE of three independent experiments; two-tailed t test: *$p < 0.05$, ***$p < 0.001$).

(D) Cell-cycle distributions of cells treated with siRNAs targeting NUP107 and assessed as described in Supplemental Experimental Procedures. Data represent the mean of two independent experiments.

See also Figure S6 and Tables S5, S6, and S7.

frequently amplified (9/17) and is overexpressed in WD/DDLPS tumors compared to other sarcomas in two independent expression data sets (Figure 6B, top and bottom panels). Validation siRNA experiments confirmed the effect of NUP107 depletion in WD/DDLPS cells (Figure 6C). Cells with amplified *NUP107* (449, 778, T1000, and GOT3) were more sensitive to NUP107 knockdown (33.7% to 71.8% decrease in cell viability) than cells lacking *NUP107* amplification (normal human-bone-derived [NHB] and SW872; 1.98% to 16.7%). NUP107 depletion induced cell cycle arrest in 449 cells, whereas SW872 cells were relatively

unaffected (Figure 6D). Taken together, these data suggest that *NUP107* amplification and overexpression provide a growth advantage in WD/DDLPS tumors.

DISCUSSION

In this study, we used deep molecular analyses to characterize the architecture, evolution, and functional properties of cancer neochromosomes. An exhaustive structural and cytogenetic analysis from several WD/DDLPS cell lines enabled us to model
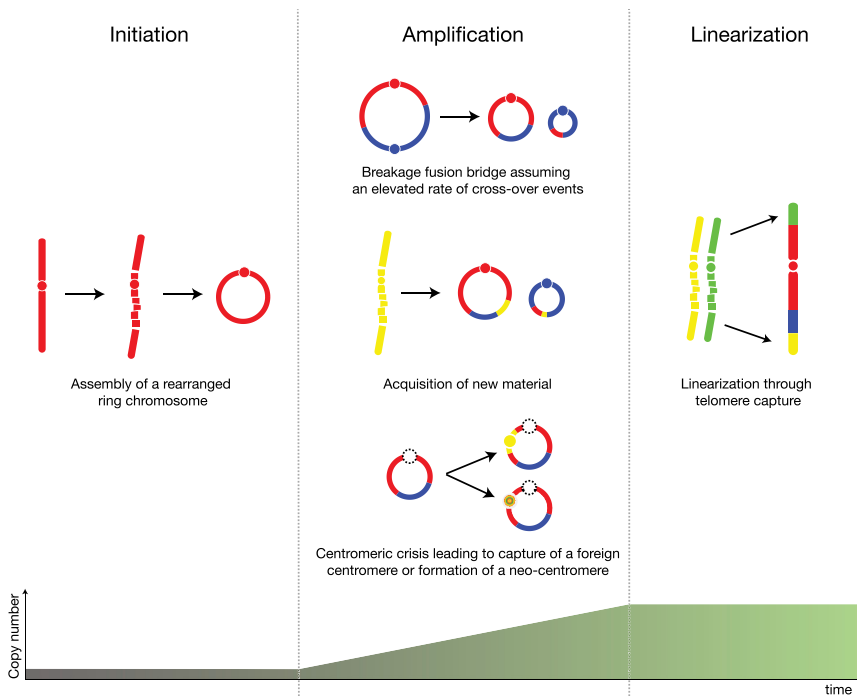
Initiation | Amplification | Linearization

Breakage fusion bridge assuming
an elevated rate of cross-over events

Assembly of a rearranged
ring chromosome

Acquisition of new material

Linearization through
telomere capture

Centromeric crisis leading to capture of a foreign
centromere or formation of a neo-centromere

Copy number

time

Figure 7. Model of Neochromosome Evolution
The life history consists of three phases, which we label Initiation, Amplification, and Linearization. During, or soon after initiation, chromothripsis of chromosome 12 contributes to the assembly of a ring chromosome consisting of rearranged chromosome 12 DNA, including oncogenic sequences and possibly containing a native alphoid centromere (red circle). The amplification phase is characterized by multiple rounds of circular BFB that shuffles and amplifies loci under selection, as well as loss of other regions. DNA from other chromosomes may be captured, possibly by a process involving chromothripsis (yellow segment). Centromeric crises (dashed circle) created by progressive corrosion are resolved by neocentromere formation (bullseye) or centromere capture (yellow circle). The amplification phase appears to be extinct in all lines studied. The final phase is defined by linearization, chromothriptic telomere capture, and continued genomic instability.

the process of neochromosome formation, from its earliest origins to the final forms visible in the cell lines and tumors studied here (summarized in Figure 7). Although the exact ordering of initial events remains a matter of speculation, chromothriptic fragmentation of chromosome 12q appears an early and important event in the formation of a circular structure (Figure 7, left panel). This interpretation is also consistent with the double minutes observed in chromothriptic tumors (Sanborn et al., 2013; Stephens et al., 2011) but does not formally exclude the possibility that chromothripsis occurs soon after the formation of the initial episomal structure; for example, by fusion of double-strand breaks on each chromosome arm or telomere fusion.

A core group of genes located on chromosome 12 is almost invariably amplified in WD/DDLPS, including *MDM2* and *CDK4*. *Mdm2* was originally cloned from circular, episomal double-minute structures in murine 3T3 cells (Cahilly-Snyder et al., 1987). Recently, acentric double minutes in glioblastoma were studied by whole-genome sequencing (Sanborn et al., 2013). These studies, performed at 30× coverage on whole-tumor DNA, demonstrated a similar pattern of structural contiguity within structures typically around a megabase in size and that, in many cases, also contained *MDM2*. Almost-perfect homozygosity of amplified material in our isolated neochromosomes suggests that DNA is incorporated in a single event from one donor chromatid. The typically chromothriptic patterns at the junction of captured telomeric regions suggests that this process is also associated with subsequent integration of new DNA into established neochromosomes. Such new integrations may involve the fusion of isoforms of neochromosomes themselves, where (as frequently observed) more than one neochromosome exists per cell.

After assembly of the initial single-copy, circular, extrachromosomal structure, hundreds of BFB cycles subsequently mediate amplification of loci under selection, along with deletion of other regions (Figure 7, center panel). BFB was first identified in maize (McClintock, 1941) and has been proposed as a mechanism for gene amplification in cancer (Bignell et al., 2007; Gisselsson et al., 2000; Lo et al., 2002a). Since BFB of linear dicentric chromosomes results in inverted duplications (Bignell et al., 2007) not observed in our data, and since the proportions of different classes of fusions (deletion type, tandem duplication type, head to head, and tail to tail) remain uniform despite amplification, we propose an alternative form of BFB, in which a doubled, fused dicentric ring forms following an odd number of sister chromatid exchanges. Ring forms are common in primary WD/DDLPS (Nord et al., 2013) and were originally observed in tumor metaphases of the 449 cell line (Sirvent et al., 2000). Their replacement by the observable exclusively linear forms seen in the cell lines may explain the ex vivo stabilization of the neochromosome core in vitro. All cell lines now also demonstrate a near-tetraploid state, perhaps explaining the presence of multiple neochromosomes per cell, despite early metaphases of the original primary tumors being frequently diploid. The progressive replacement of ring by linear episomal structures has been observed previously and may represent a general phenomenon (Levan et al., 1978). Indeed, episodic linearization during tumor development may account for incorporation of additional regions of DNA into the neochromosome by telomere capture, originally described in melanoma (Meltzer et al., 1993). Progressive genomic instability, a feature of tumor chaotropy, may favor the linear form of the neochromosome and contribute to dedifferentiation in vivo. Such a transition depends on the lack of selection for BFB-mediated amplification of oncogenes, a state that inevitably occurs when ongoing amplification of genes such as *MDM2* and *CDK4* provides no additional selective benefit.

BFB is also associated with loss of DNA from the neochromosome, potentially accounting for the prevalence of neocentromeres in WD/DDLPS neochromosomes. The neocentromere

on the 778 neochromosomes spanned conjoined regions from chromosomes 1 and 6. Given that these regions are nonaboriginal, this neocentromere is a descendant of the original centromere used by the 778 neochromosomes. Not all WD/DDLPS neochromosomes contain neocentromeres. Two WD/DDLPS cell lines (GOT3 and 449) showed clear evidence of either active or inactive alphoid centromeres on neochromosomes, with attenuated levels of α-satellite DNA likely due to progressive BFB-mediated centromere corrosion. Reduction of centromeric α-satellite DNA has previously been noted in human chromosomes containing repositioned human neocentromeres (Amor et al., 2004; Tyler-Smith et al., 1999), supporting the idea that centromere deactivation is a driver for subsequent neocentromere formation (Marshall et al., 2008a). In WD/DDLPS, the original alphoid centromere in the newly formed neochromosome may be replaced either by neocentromere formation or by the capture of another native alphoid centromere.

Active selection is required for recurrent representation of donor regions within the WD/DDLPS neochromosome, including those acquired later in the life cycle. One kind of selection involves the amplification and overexpression of oncogenes. Many genes that are amplified in more than half of all primary WD/DDLPS are functionally important, including *MDM2*, *CDK4*, *YEATS4*, *OS9*, and *HMGA2*, all genes located in a 12 Mb region of chromosome 12q (58–70 Mb) pivotal to the initial formation of the episomal ring structure (Nord et al., 2013). *NUP107*, also identified recently in an integrative array-based study of soft tissue tumors (Nord et al., 2013), may have oncogenic properties. Located on chromosome 12q, NUP107 belongs to the nucleoporin family of holoenzyme subunits constituting the nuclear pore complex. The yeast homolog of NUP107, Nup84, anchors telomeric regions of chromosomes and is required for efficient repair of subtelomeric DNA double-strand breaks (Nagai et al., 2008; Palancade et al., 2007; Therizols et al., 2006). NUP107 is overexpressed in breast cancer, where expression correlates with poor prognosis (Agudo et al., 2004), and appears to have antiapoptotic functions in astrocytoma (Banerjee et al., 2010). However, not all amplified genes are overexpressed. A layer of epigenetic transcriptional regulation silences amplified genes whose expression may be deleterious, such as the proapoptotic gene *FASLG.* Regions may also be selected because of locus-specific noncoding functions, exemplified by neocentromere formation on regions from chromosomes 1 and 6. Both chromosomes 1 and 6 are amplified recurrently in WD/DDLPS (Forus et al., 2001; Nilsson et al., 2004; Nord et al., 2013; Tap et al., 2011).

Cancers often have signature mutagenic patterns, ranging from single-nucleotide changes to gross genomic instability. The present study adds to other recently described and unique chromosome-scale mutations, including chromothripsis itself (Stephens et al., 2011), kataegis (Nik-Zainal et al., 2012), the low-level amplification of megabase-scale regions of chromosome 21 in subtypes of childhood leukemia (Li et al., 2014), and double minutes (Sanborn et al., 2013). Confirming recent data (Li et al., 2014), we provide further evidence for the combination of chromothripsis and BFB cycles in tumor evolution. Our data indicate that the generation of neochromosomes in WD/DDLPS requires hundreds of circular BFB cycles to generate the recurrent high-level, focal amplification observed in these massive structures. These corrosive forces are sculpted by

intense selection. The region of chromosome 12 that nucleates formation of the neochromosome contains two potent oncogenes critical to cellular immortalization, targeting the p53 (*MDM2*) and retinoblastoma (*CDK4*) tumor suppressor pathways (Nord et al., 2013). Taken together, these studies reveal the complex architecture and life cycle of a cancer-specific mutation on a chromosome scale.

EXPERIMENTAL PROCEDURES

Cell Lines and Tumor Samples

A total of 449 cells were established from a primary retroperitoneal WDLPS from a 68-year-old woman in 1993, and the 778 cells were derived from the same patient's retroperitoneal relapse in 1994 (Sirvent et al., 2000). GOT3 cells were established from the first local recurrence of a WDLPS located close to the left kidney of a 76-year-old woman (Persson et al., 2008). T1000 cells were generated from a mixed WD/DDLPS arising from the thigh in a 43-year-old male (Pedeutour et al., 2012). The LPS141 cell line was generated from a dedifferentiated liposarcoma in a patient with recurrent disease (Snyder et al., 2009). SW872 cells were from the American Type Culture Collection (HTB-92). Cells were maintained in RPMI 1640 or Dulbecco's modified Eagle's medium/F-12 supplemented with 10% FBS and 200 mM GlutaMAX (Invitrogen). NHB cells were grown as described elsewhere (Kansara et al., 2009). Fresh frozen samples (152) were collected from patients diagnosed with a variety of bone and soft tissue tumors, including WD/DDLPS, myxoid liposarcoma, Ewing's sarcoma, synovial sarcoma, leiomyosarcoma, undifferentiated pleomorphic sarcoma (also known as malignant fibrous histiocytoma), osteosarcoma, malignant peripheral nerve sheath tumor, chondrosarcoma, lipoma, schwannoma, and desmoid tumor. Samples were obtained with appropriate informed consent under the auspices of the Peter MacCallum Cancer Centre Institutional Human Research Ethics Committee (Research Protocol 03/49). None of the cell lines or primary tumors were exposed to clastogenic therapy (chemotherapy or ionizing radiation), with the exception that this information was not available for the LPS141 cell line.

Chromosome Preparation, Staining, and Isolation

Adherent cells were grown to approximately 70% confluence and treated with 100 ng/ml KaryoMAX Colcemid Solution (GIBCO) for 16 hr to arrest cells in metaphase. Chromosomes were extracted and stained, and neochromosomes were isolated by flow cytometry, as previously described (Marshall et al., 2008b), on a flow cytometer (MoFlo, DAKO). Individual and pooled populations of neochromosomes were isolated based on the flow karyotype.

Next-Generation Sequencing

Pooled isolated neochromosomes from the 778, 449, GOT3, T1000, and LPS141 cell lines; whole-genome DNA from two patient samples, ST059 and ST079; and the ChIP-seq libraries were sequenced on the Illumina GAIIx and HiSeq 2000 platforms (see Table S1 for details).

The complete materials and methods are provided in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, seven tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.ccell.2014.09.010.

AUTHOR CONTRIBUTIONS

D.W.G., A.J.H., A.T.P., and D.M.T. conceived the ideas and design of this study. D.W.G. and O.J.M. carried out chromosome sorting. D.W.G.,

REFERENCES

Agudo, D., Gómez-Esquer, F., Martínez-Arribas, F., Núñez-Villar, M.J., Pollán, M., and Schneider, J. (2004). Nup88 mRNA overexpression is associated with high aggressiveness of breast cancer. Int. J. Cancer 109, 717–720.

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., and Choo, K.H. (2004). Human centromere repositioning "in progress". Proc. Natl. Acad. Sci. USA 101, 6542–6547.

Baird, D.M., and Farr, C.J. (2006). The organization and function of chromosomes. EMBO Rep. 7, 372–376.

Banerjee, H.N., Gibbs, J., Jordan, T., and Blackshear, M. (2010). Depletion of a single nucleoporin, Nup107, induces apoptosis in eukaryotic cells. Mol. Cell. Biochem. 343, 21–25.

Barretina, J., Taylor, B.S., Banerji, S., Ramos, A.H., Lagos-Quintana, M., Decarolis, P.L., Shah, K., Socci, N.D., Weir, B.A., Ho, A., et al. (2010). Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. Nat. Genet. 42, 715–721.

Bignell, G.R., Santarius, T., Pole, J.C., Butler, A.P., Perry, J., Pleasance, E., Greenman, C., Menzies, A., Taylor, S., Edkins, S., et al. (2007). Architectures of somatic genomic rearrangement in human cancer amplicons at sequence-level resolution. Genome Res. 17, 1296–1303.

Buoen, L.C., and Brand, K.G. (1968). Double-minute chromosomes in plastic film-induced sarcomas in mice. Naturwissenschaften 55, 135–136.

Cahilly-Snyder, L., Yang-Feng, T., Francke, U., and George, D.L. (1987). Molecular analysis and chromosomal mapping of amplified genes isolated from a transformed mouse 3T3 cell line. Somat. Cell Mol. Genet. 13, 235–244.

Dahlback, H.S., Brandal, P., Meling, T.R., Gorunova, L., Scheie, D., and Heim, S. (2009). Genomic aberrations in 80 cases of primary glioblastoma multiforme: Pathogenetic heterogeneity and putative cytogenetic pathways. Genes Chromosomes Cancer 48, 908–924.

Forus, A., Larramendy, M.L., Meza-Zepeda, L.A., Bjerkehagen, B., Godager, L.H., Dahlberg, A.B., Saeter, G., Knuutila, S., and Myklebost, O. (2001). Dedifferentiation of a well-differentiated liposarcoma to a highly malignant metastatic osteosarcoma: amplification of 12q14 at all stages and gain of 1q22-q24 associated with metastases. Cancer Genet. Cytogenet. 125, 100–111.

Garsed, D.W., Holloway, A.J., and Thomas, D.M. (2009). Cancer-associated neochromosomes: a novel mechanism of oncogenesis. BioEssays 31, 1191–1200.

Gisselsson, D., Pettersson, L., Höglund, M., Heidenblad, M., Gorunova, L., Wiegant, J., Mertens, F., Dal Cin, P., Mitelman, F., and Mandahl, N. (2000). Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. Proc. Natl. Acad. Sci. USA 97, 5357–5362.

Italiano, A., Bianchini, L., Keslair, F., Bonnafous, S., Cardot-Leccia, N., Coindre, J.M., Dumollard, J.M., Hofman, P., Leroux, A., Mainguené, C., et al. (2008). HMGA2 is the partner of MDM2 in well-differentiated and dedifferentiated liposarcomas whereas CDK4 belongs to a distinct inconsistent amplicon. Int. J. Cancer 122, 2233–2241.

Kansara, M., Tsang, M., Kodjabachian, L., Sims, N.A., Trivett, M.K., Ehrich, M., Dobrovic, A., Slavin, J., Choong, P.F., Simmons, P.J., et al. (2009). Wnt inhibitory factor 1 is epigenetically silenced in human osteosarcoma, and targeted disruption accelerates osteosarcomagenesis in mice. J. Clin. Invest. 119, 837–851.

Korbel, J.O., and Campbell, P.J. (2013). Criteria for inference of chromothripsis in cancer genomes. Cell 152, 1226–1236.

Levan, A., Levan, G., and Mandahl, N. (1978). A new chromosome type replacing the double minutes in a mouse tumor. Cytogenet. Cell Genet. 20, 12–23.

Li, Y., Schwab, C., Ryan, S.L., Papaemmanuil, E., Robinson, H.M., Jacobs, P., Moorman, A.V., Dyer, S., Borrow, J., Griffiths, M., et al. (2014). Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. Nature 508, 98–102.

Liehr, T., Mrasek, K., Weise, A., Dufke, A., Rodríguez, L., Martínez Guardia, N., Sanchís, A., Vermeesch, J.R., Ramel, C., Polityko, A., et al. (2006). Small supernumerary marker chromosomes—progress towards a genotype-phenotype correlation. Cytogenet. Genome Res. 112, 23–34.

Lo, A.W., Sabatier, L., Fouladi, B., Pottier, G., Ricoul, M., and Murnane, J.P. (2002a). DNA amplification by breakage/fusion/bridge cycles initiated by spontaneous telomere loss in a human cancer cell line. Neoplasia 4, 531–538.

Lo, A.W., Sprung, C.N., Fouladi, B., Pedram, M., Sabatier, L., Ricoul, M., Reynolds, G.E., and Murnane, J.P. (2002b). Chromosome instability as a result of double-strand breaks near telomeres in mouse embryonic stem cells. Mol. Cell. Biol. 22, 4836–4850.

Marshall, O.J., Chueh, A.C., Wong, L.H., and Choo, K.H. (2008a). Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. Am. J. Hum. Genet. 82, 261–282.

Marshall, O.J., Marshall, A.T., and Choo, K.H. (2008b). Three-dimensional localization of CENP-A suggests a complex higher order structure of centromeric chromatin. J. Cell Biol. 183, 1193–1202.

McClintock, B. (1941). The stability of broken ends of chromosomes in Zea Mays. Genetics 26, 234–282.

Meltzer, P.S., Guan, X.Y., and Trent, J.M. (1993). Telomere capture stabilizes chromosome breakage. Nat. Genet. 4, 252–255.

Nagai, S., Dubrana, K., Tsai-Pflugfelder, M., Davidson, M.B., Roberts, T.M., Brown, G.W., Varela, E., Hediger, F., Gasser, S.M., and Krogan, N.J. (2008). Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science 322, 597–602.

Nakayama, R., Nemoto, T., Takahashi, H., Ohta, T., Kawai, A., Seki, K., Yoshida, T., Toyama, Y., Ichikawa, H., and Hasegawa, T. (2007). Gene expression analysis of soft tissue sarcomas: characterization and reclassification of malignant fibrous histiocytoma. Mod. Pathol. 20, 749–759.

Nik-Zainal, S., Alexandrov, L.B., Wedge, D.C., Van Loo, P., Greenman, C.D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, L.A., et al.; Breast Cancer Working Group of the International Cancer Genome Consortium (2012). Mutational processes molding the genomes of 21 breast cancers. Cell 149, 979–993.

Nilsson, M., Meza-Zepeda, L.A., Mertens, F., Forus, A., Myklebost, O., and Mandahl, N. (2004). Amplification of chromosome 1 sequences in lipomatous tumors and other sarcomas. Int. J. Cancer 109, 363–369.

Nord, K.H., Macchia, G., Tayebwa, J., Nilsson, J., Vult von Steyern, F., Brosjo, O., Mandahl, N., and Mertens, F. (2013). Integrative genome and transcriptome analyses reveal two distinct types of ring chromosome in soft tissue sarcomas. Hum. Mol. Genet. 23, 878–888.

Palancade, B., Liu, X., Garcia-Rubio, M., Aguilera, A., Zhao, X., and Doye, V. (2007). Nucleoporins prevent DNA damage accumulation by modulating Ulp1-dependent sumoylation processes. Mol. Biol. Cell 18, 2912–2923.

Park, J.H., and Roeder, R.G. (2006). GAS41 is required for repression of the p53 tumor suppressor pathway during normal cellular proliferation. Mol. Cell. Biol. 26, 4006–4016.

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., et al. (2008). A physical map of the 1-gigabase bread wheat chromosome 3B. Science 322, 101–104.

Pedeutour, F., Suijkerbuijk, R.F., Forus, A., Van Gaal, J., Van de Klundert, W., Coindre, J.M., Nicolo, G., Collin, F., Van Haelst, U., Huffermann, K., et al. (1994). Complex composition and co-amplification of SAS and MDM2 in ring and giant rod marker chromosomes in well-differentiated liposarcoma. Genes Chromosomes Cancer 10, 85–94.

Pedeutour, F., Maire, G., Pierron, A., Thomas, D.M., Garsed, D.W., Bianchini, L., Duranton-Tanneur, V., Cortes-Maurel, A., Italiano, A., Squire, J.A., et al. (2012). A newly characterized human well-differentiated liposarcoma cell line contains amplifications of the 12q12-21 and 10p11-14 regions. Virchows Arch. 461, 67–78.

Persson, F., Olofsson, A., Sjögren, H., Chebbo, N., Nilsson, B., Stenman, G., and Aman, P. (2008). Characterization of the 12q amplicons by high-resolution, oligonucleotide array CGH and expression analyses of a novel liposarcoma cell line. Cancer Lett. 260, 37–47.

Pilotti, S., Della Torre, G., Lavarino, C., Sozzi, G., Minoletti, F., Vergani, B., Azzarelli, A., Rilke, F., and Pierotti, M.A. (1998). Molecular abnormalities in liposarcoma: role of MDM2 and CDK4-containing amplicons at 12q13-22. J. Pathol. 185, 188–190.

Sanborn, J.Z., Salama, S.R., Grifford, M., Brennan, C.W., Mikkelsen, T., Jhanwar, S., Katzman, S., Chin, L., and Haussler, D. (2013). Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. Cancer Res. 73, 6036–6045.

Sirvent, N., Forus, A., Lescaut, W., Burel, F., Benzaken, S., Chazal, M., Bourgeon, A., Vermeesch, J.R., Myklebost, O., Turc-Carel, C., et al. (2000). Characterization of centromere alterations in liposarcomas. Genes Chromosomes Cancer 29, 117–129.

Snyder, E.L., Sandstrom, D.J., Law, K., Fiore, C., Sicinska, E., Brito, J., Bailey, D., Fletcher, J.A., Loda, M., Rodig, S.J., et al. (2009). c-Jun amplification and overexpression are oncogenic in liposarcoma but not always sufficient to inhibit the adipocytic differentiation programme. J. Pathol. 218, 292–300.

Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., et al. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144, 27–40.

Stratford, E.W., Castro, R., Daffinrud, J., Skårn, M., Lauvrak, S., Munthe, E., and Myklebost, O. (2012). Characterization of liposarcoma cell lines for pre-clinical and biological studies. Sarcoma 2012, 148614.

Su, Y.A., Hutter, C.M., Trent, J.M., and Meltzer, P.S. (1996). Complete sequence analysis of a gene (OS-9) ubiquitously expressed in human tissues and amplified in sarcomas. Mol. Carcinog. 15, 270–275.

Tap, W.D., Eilber, F.C., Ginther, C., Dry, S.M., Reese, N., Barzan-Smith, K., Chen, H.W., Wu, H., Eilber, F.R., Slamon, D.J., and Anderson, L. (2011). Evaluation of well-differentiated/de-differentiated liposarcomas by high-resolution oligonucleotide array-based comparative genomic hybridization. Genes Chromosomes Cancer 50, 95–112.

Therizols, P., Fairhead, C., Cabal, G.G., Genovesio, A., Olivo-Marin, J.C., Dujon, B., and Fabre, E. (2006). Telomere tethering at the nuclear periphery is essential for efficient DNA double strand break repair in subtelomeric region. J. Cell Biol. 172, 189–199.

Tyler-Smith, C., Gimelli, G., Giglio, S., Floridia, G., Pandya, A., Terzoli, G., Warburton, P.E., Earnshaw, W.C., and Zuffardi, O. (1999). Transmission of a fully functional human neocentromere through three generations. Am. J. Hum. Genet. 64, 1440–1444.