

# Disease-specific gene repositioning in breast cancer

Karen J. Meaburn,<sup>1</sup> Prabhakar R. Gudla,<sup>2</sup> Sameena Khan,<sup>1</sup> Stephen J. Lockett,<sup>2</sup> and Tom Misteli<sup>1</sup>

<sup>1</sup>National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

<sup>2</sup>Optical Microscopy and Analysis Laboratory, Advanced Technology Program, Science Applications International Corporation-Frederick, Inc., and National Cancer Institute at Frederick, Frederick, MD 21702

**G**enomes are nonrandomly organized within the three-dimensional space of the cell nucleus. Here, we have identified several genes whose nuclear positions are altered in human invasive breast cancer compared with normal breast tissue. The changes in positioning are gene specific and are not a reflection of genomic instability within the cancer tissue. Reposi-

tioning events are specific to cancer and do not generally occur in noncancerous breast disease. Moreover, we show that the spatial positions of genes are highly consistent between individuals. Our data indicate that cancer cells have disease-specific gene distributions. These interphase gene positioning patterns may be used to identify cancer tissues.

## Introduction

The organization of the human genome within the cell nucleus is nonrandom (Cremer et al., 2006; Misteli, 2007). Chromosomes and individual genes occupy preferential localizations relative to each other and to nuclear landmarks such as the nuclear envelope (Misteli, 2007; Schneider and Grosschedl, 2007; Takizawa et al., 2008b). A convenient and quantifiable indicator of a gene's location is its position along the axis between the center of the nucleus and the nuclear edge, referred to as its radial position (Takizawa et al., 2008b). Although the radial position of some genes has been linked to their activity (Kosak et al., 2002; Chambeyron and Bickmore, 2004; Hewitt et al., 2004; Takizawa et al., 2008a), the functional relevance of radial positioning is not clear (Takizawa et al., 2008b).

The spatial organization of the genome changes during physiological processes such as differentiation and development (Foster and Bridger, 2005; Takizawa et al., 2008b). Importantly, large-scale alterations of spatial organization also occur in pathological states (Borden and Manuelidis, 1988; Zink et al., 2004; Meaburn et al., 2007). A major hallmark of many cancers, which is routinely exploited by pathologists, is the distinctive changes to cancer nuclei at the gross level, such as to nuclear shape and chromatin texture (Zink et al., 2004). These changes suggest there must also be major changes to the spatial genome organization in cancer nuclei (Zink et al., 2004). Indeed, sporadic evidence has suggested spatial genome reorganization in human

cancer. Human chromosome (HSA) 8 moves toward the nuclear periphery in pancreatic cancer (Wiech et al., 2005), and a significant fraction of nuclei show changes in the positioning of HSA 18 and 19 in multiple cancer types (Cremer et al., 2003; Wiech et al., 2009). In addition to entire chromosomes, the centromere of HSA 17 becomes more internally positioned in breast cancer compared with normal tissues (Wiech et al., 2005).

Little is known about changes in positioning of individual genes in cancer cells. In a 3D culture in vitro model system of early breast cancer, *AKT1*, *BCL2*, *ERBB2*, and *VEGFA* have been demonstrated to undergo repositioning (Meaburn and Misteli, 2008), but it is unclear to what degree similar changes occur in cancer tissues. The only reported gene-specific change in gene location in cancer tissues is the marginally more peripheral position of *BCL2* in a *BCL2*-positive cervical squamous carcinoma tissue (Wiech et al., 2009). In contrast, *BCL2* did not reposition in a *BCL2*-negative cervical squamous carcinoma tissue (Wiech et al., 2009), and *ERBB2* was found to not alter radial position in a breast cancer tissue (Wiech et al., 2005). However, these studies are based on only a single cancer tissue, making it difficult to assess how general repositioning events are, or if they are random events. Here, we set out to identify genes that are frequently differentially positioned in breast cancer tissues, and we explore the possibility that disease-specific spatial organization of the genome may be used to distinguish malignant from normal tissue.

Correspondence to Tom Misteli: [mistelit@mail.nih.gov](mailto:mistelit@mail.nih.gov)

Abbreviations used in this paper: BAC, bacterial artificial chromosome; EDT, Euclidean distance transform; ER, estrogen receptor; KS test, Kolmogorov-Smirnov test; PR, progesterone receptor; RRD, relative radial distribution; TMA, tissue microarray.

This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.jcb.org/misc/terms.shtml>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

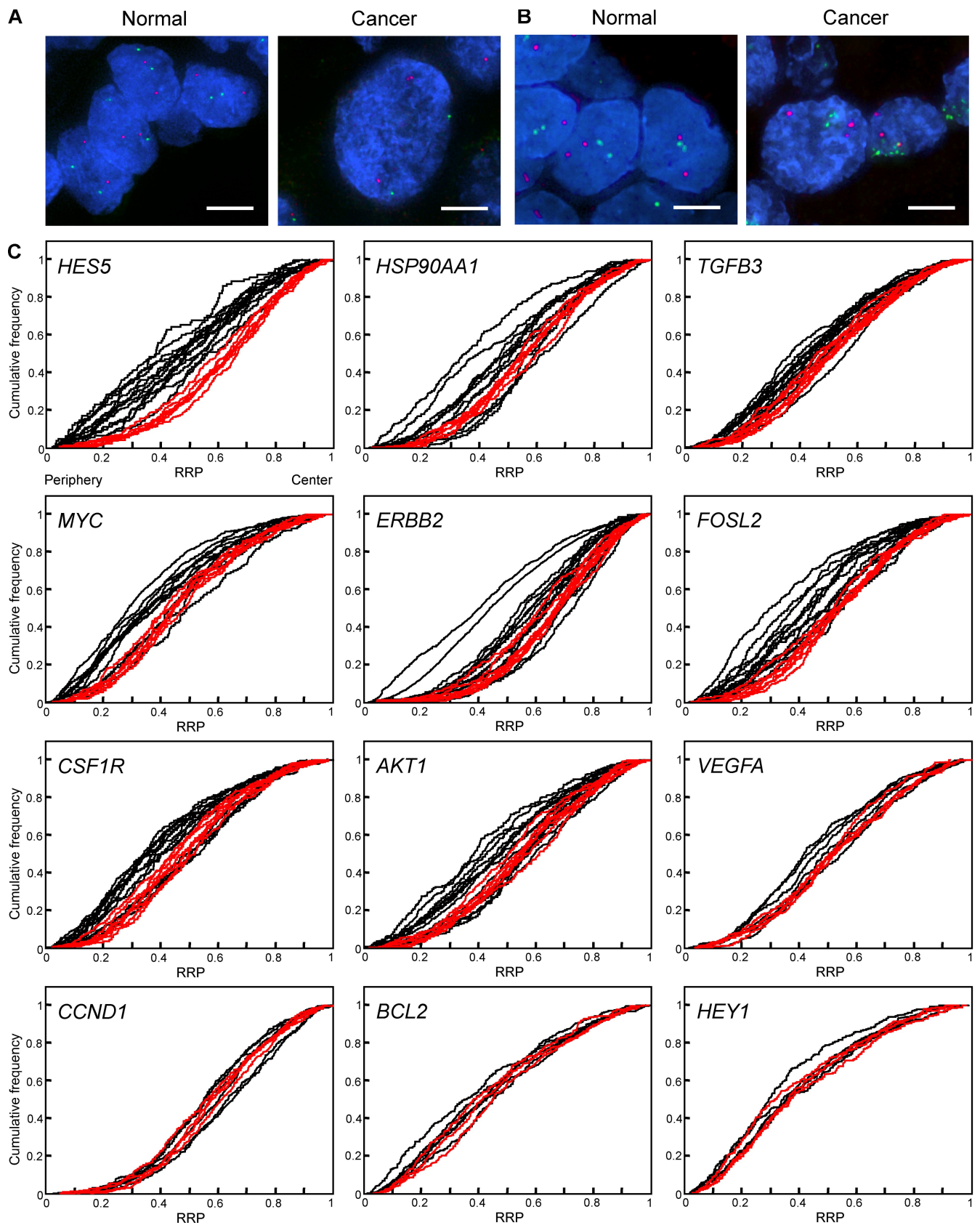


Figure 1. **Spatial reorganization of the genome in cancer.** (A) *FOLS2* (green) and *CSF1R* (red), and (B) *ERBB2* (green) and *MYC* (red) are detected by FISH in normal or cancer breast tissue. Nuclei were counterstained with DAPI (blue). Projected image stacks are shown. Bars, 5  $\mu$ m. (C) Cumulative RRDs for the indicated genes in cancerous breast tissues (black) and normal tissues (red).  $n = 88$ –220 cells.

Table I. Characterization of breast tissues

Sample code	Pathology	Other information	Source	Source specimen ID	Tissue thickness
N1	Normal	Absence of cancer; 35 y	Biomax	HuFPT 130 a	5 μm
N2	Normal	Absence of cancer; 44 y	Biomax	HuFPT 129	5 μm
N3	Normal	Absence of cancer; 53 y	Biomax	HuFPT 128	5 μm
N4	Normal	Absence of cancer	ACSR	AA-98-3672	5 μm
N5	Normal	Absence of cancer; 56 y	BioChain	T2234086	5 μm
N6	Normal	NAT; 50 y	BioChain	TMA core A2	4 μm
N7	Normal	NAT; 51 y	Imgenex	IMH-1013	4 μm
N8	Normal	NAT (cancer C4)	ACSR	AA-98-313-Q	5 μm
N9	Normal	NAT	ACSR	AA-96-193-3	5 μm
N10	Normal	NAT	ACSR	AA-97-239-2	5 μm
N11	Normal	NAT (cancer C2)	ACSR	AA-02-161-HH	5 μm
C1	IDC with DCIS	ER+	ACSR	AA-97-108 B	5 μm
C2	IDC with DCIS	Lymph node +, ER-, PR-, HER2-	ACSR	AA-02-161 C	5 μm
C3	IDC	Lymph node +	ACSR	AA-98-31 A	5 μm
C4	IDC with DCIS	ER+, PR+, HER2+	ACSR	AA-98-313 C	5 μm
C5	IDC with DCIS	Moderate to high grade	ACSR	AA-98-051-2	5 μm
C6	IDC	Poorly differentiated	ACSR	AA-96-152-3	5 μm
C7	IDC	TMN stage T4N3M1; 50 y	BioChain	TMA core C1	4 μm
C8	IDC	TMN stage T2N1M0; 43 y	BioChain	TMA core C2	4 μm
C9	IDC	TMN stage T2N3M0; 47 y	BioChain	TMA core C3	4 μm
C10	IDC	TMN stage T2N0M0; 56 y	BioChain	TMA core C4	4 μm
C11	IDC	TMN stage T2N2M0; 34 y	BioChain	TMA core C5	4 μm
C12	ILC	TMN stage T2N0M0	BioChain	TMA core D1	4 μm
C13	ILC	TMN stage T2N0M0	BioChain	TMA core D2	4 μm
C14	ILC	Moderately differentiated; tumor size: 3 cm diameter; 38 y	Capital	CT2235086-3	5 μm
B1	Hyperplasia	41 y	BioChain	TMA core B1	4 μm
B2	Hyperplasia	23 y, male	BioChain	TMA core B2	4 μm
B3	Fibroadenoma	45 y	BioChain	TMA core B3	4 μm
B4	Fibroadenoma	25 y	BioChain	TMA core B4	4 μm
B5	Fibroadenoma	32 y	BioChain	TMA core B5	4 μm
B6	Hyperplasia	49 y	BioChain	TMA core A4	4 μm

IDC, invasive ductal carcinoma; DCIS, ductal carcinoma in situ; ILC, invasive lobular carcinoma; NAT, normal adjacent to breast cancer tissue; y, years old; +, positive; -, negative; TNM, tumor (graded 0–4, based on tumor size and extension to chest wall or skin), node (lymph node metastasis, numbered as NO [no spread to lymph nodes] or N1–3 [metastasis to lymph nodes]), metastasis (0 = no distant metastasis, 1 = distant metastasis) classification of malignant tumors; Biomax, US Biomax, Inc; ACSR, AIDS and Cancer Specimen Resource; Imgenex, Imgenex Corp.; BioChain, BioChain Institute, Inc.; TMA, TMA catalog no. Z7020010; Capital, Capital Biosciences, Inc.

## Results

We sought to identify genes that occupy distinct intranuclear positions in normal and malignant cells. To this end, we visualized a set of 20 gene loci (Table S1) by FISH in a panel of 11 normal and 14 invasive carcinoma human breast tissues (Fig. 1, A and B; and Table I). The radial position of a gene, normalized to the size of the nucleus, was determined using a modified version of a previously developed image analysis method (Meaburn and Misteli, 2008; Takizawa et al., 2008a), which takes into account the non-elliptical shape of some of the nuclei (see Materials and methods). Data from 88–220 nuclei per sample (Fig. S1), acquired from multiple randomly selected regions of the tissue sample, were analyzed and combined to determine the cumulative relative radial distribution (RRD) for each gene in a tissue (Figs. 1 C and S2 A). The RRD is a standard measure of a gene's position in a population and is defined as the statistical distribution of the radial position of all alleles in a cell population (see Materials and methods). RRDs were statistically

compared with each other using the two-sample 1D Kolmogorov-Smirnov test (KS test) as described previously (Figs. 2 and S2; see Materials and methods; Meaburn and Misteli, 2008; Takizawa et al., 2008a). The RRDs were considered distinct if  $P < 0.01$ . RRDs for a given gene were highly reproducible between experiments and were statistically indistinguishable ( $0.65 \leq P \leq 0.81$ ). The 20 genes mapped to 14 chromosomes (Table S1), and were selected randomly and irrespective of their function in order to enable an unbiased screening approach.

### Low variability of spatial gene positioning patterns among individuals

Initially, we determined to what degree spatial gene positioning patterns are reproducible between individuals by comparing the positions of a subset of genes in morphologically normal breast tissue from multiple individuals. 8 out of 15 genes were indistinguishable in all cross-comparisons between normal breast tissues ( $P > 0.01$ ; *HSP90AA1*, *TGFB3*, *MYC*, *VEGFA*, *CCND1*, *HEY1*, *MMP1*, and *ZNF217*; Figs. 1, 2, and S2; and Table II).

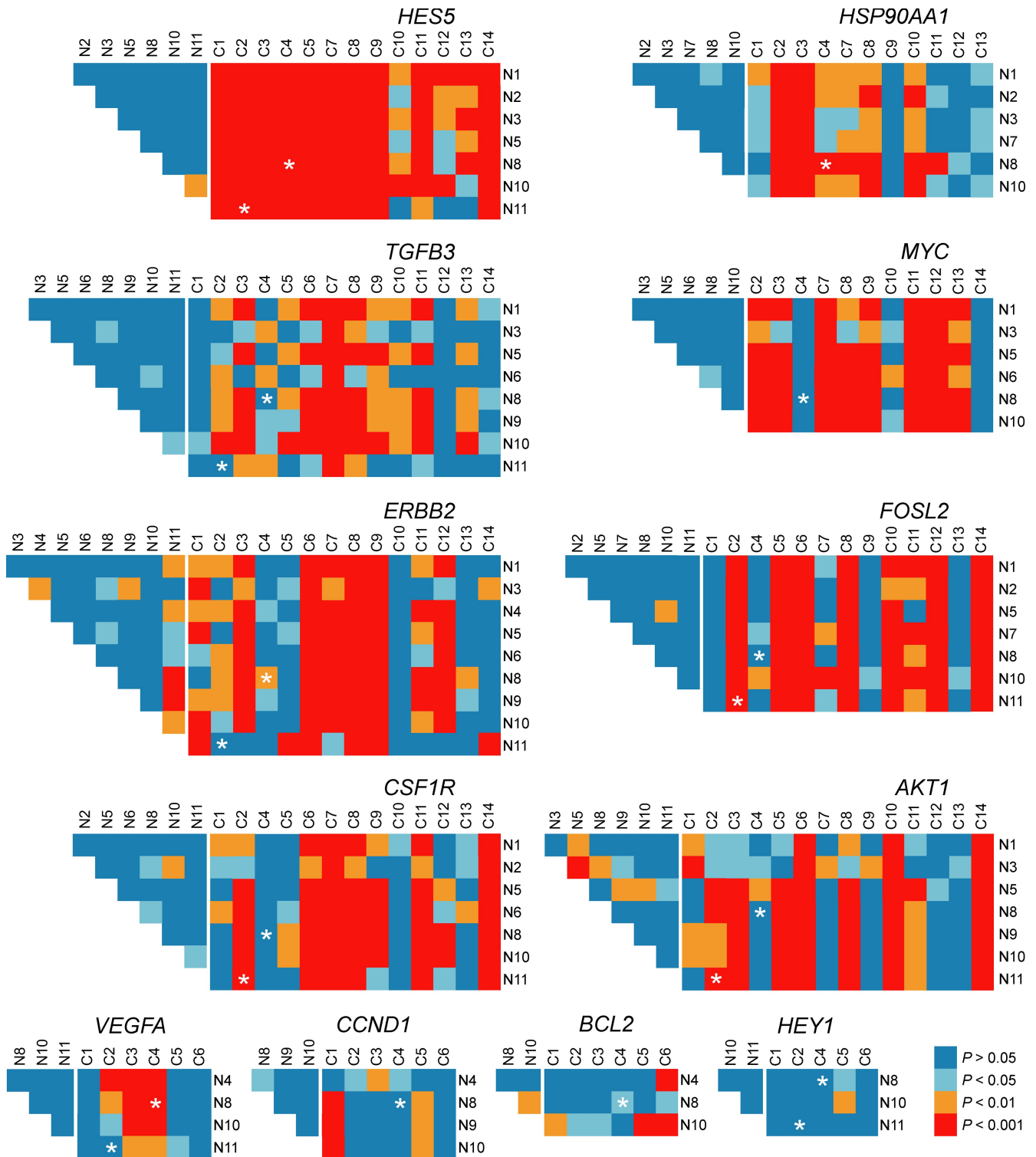


Figure 2. **Spatial reorganization of the genome in cancer.** Pairwise statistical comparisons of cumulative RRDs between individual normal (N1–11) and cancerous (C1–14) breast tissues using the two-sample 1D KS test.  $n = 88$ –220 cells. Asterisks denote a cross-comparison between a normal and cancer specimen from the same individual (see Table I).

Five genes (*HES5*, *FOSL2*, *CSF1R*, *BCL2*, and *HES1*) showed a single discrepancy among all cross-comparisons ( $0.0015 \leq P \leq 0.0077$ ). Two genes, *ERBB2* and *AKT1*, had a significantly different RRD for a single individual (N11 and N5, respectively) compared with the majority of all normal tissues (Fig. 2).

Furthermore, the positioning patterns in tissue from healthy individuals (N1–5) were indistinguishable from those in normal tissues adjacent to the tumor from breast cancer patients (N6–11; Figs. 1, 2, and S2; and Table I). Four of nine tested genes were indistinguishably located in the two types of normal

Table II. Cross-comparisons between individual tissues

Gene	Number of cross-comparisons between individual normal tissues			Number of cross-comparisons between individual normal and cancer tissues		
	SD	Total	% SD	SD	Total	% SD
<i>HES5</i>	1	21	4.8	83	91	91.2
<i>MYC</i>	0	15	0.0	47	66	71.2
<i>FOSL2</i>	1	21	4.8	58	91	63.7
<i>HSP90AA1</i>	0	15	0.0	41	66	62.1
<i>CSF1R</i>	1	21	0.0	56	91	61.5
<i>ERBB2</i>	7	36	19.4	73	126	57.9
<i>AKT1</i>	5	21	23.8	54	98	55.1
<i>TGFB3</i>	0	28	0.0	59	112	52.7
<i>HES1</i>	1	3	33.3	6	12	50.0
<i>ZNF217</i>	1	1	0.0	3	6	50.0
<i>VEGFA</i>	0	6	0.0	10	24	41.7
<i>MMP1/3/12</i>	0	1	0.0	2	6	33.3
<i>CCND1</i>	0	6	0.0	7	24	29.2
<i>PTGS2</i>	ND	ND	ND	1	4	25.0
<i>BCL2</i>	1	3	33.3	4	18	22.2
<i>HEY1</i>	0	3	0.0	1	15	6.7
<i>BRCA1</i>	ND	ND	ND	0	2	0.0
<i>PTEN</i>	ND	ND	ND	0	2	0.0
<i>TLE1</i>	ND	ND	ND	0	2	0.0
<i>TJP1</i>	ND	ND	ND	0	1	0.0
Total	18	201	9.0	505	857	58.9

SD, significantly different based on a two-sample 1D KS test,  $P < 0.01$ ; ND, not determined.

samples for all cross-comparisons (Figs. 2 and S2;  $P > 0.01$ ), with a further three genes having only a single significant difference. The remaining two genes had three cross-comparisons that were significantly different. Collectively, and consistent with earlier data (Wiech et al., 2005), these observations demonstrate a low degree of variability in spatial gene positioning between individuals.

#### Identification of repositioned genes in breast cancer

To identify genes that are repositioned in a wide range of breast cancer tissues, we deliberately analyzed the positions of genes in a diverse population of invasive breast carcinomas, which included both ductal and lobular carcinomas, HER2-positive and HER2-negative cancers, cancers that are estrogen receptor (ER)- and progesterone receptor (PR)-positive or ER/PR negative, and carcinomas that have metastasized to the lymph nodes and tumors without known metastases (Table I).

When multiple cancer tissues were individually cross compared with multiple normal tissues, six genes (*BRCA1*, *PTEN*, *TJP1*, *TLE1*, *HEY1*, and *BCL2*) showed very little change in position between normal and cancer breast tissue, and repositioned in only 0–22% of all possible cross-comparisons (Fig. 2, Fig. S2 B, and Table II). Six loci (*PTGS2*, *CCND1*, *MMP1/3/12*, *VEGFA*, *ZNF217*, and *HES1*) showed differential positioning in 25–50% of cross-comparisons. Seven genes (*TGFB3*, *AKT1*, *ERBB2*, *CSF1R*, *HSP90AA1*, *FOSL2*, and *MYC*) repositioned in 53–71% of comparisons (Fig. 2 and Table II). One gene, *HES5*, repositioned in 91% of the pairwise comparisons (83/91 tissue comparisons; Fig. 2 and Table II). We focused

further analysis on the eight genes with the highest degree of repositioning.

For these eight genes, approximately half of the pairwise comparisons in a given tumor were consistent between all normal tissues, with 55/103 comparisons showing either repositioning of a given locus in a tumor relative to all normal tissues (Fig. 2, vertical red/orange columns; 40/103) or no repositioning to all normal tissues (Fig. 2, vertical blue columns; 15/103). In two individuals for which both cancer and adjacent normal breast tissue were available (Table I; asterisks in Figs. 2 and S2), the majority of comparisons of a gene's position between the tumor relative to its corresponding adjacent normal tissue showed similar repositioning behavior to that of the tumor compared with the majority of other normal tissues, with only 4 of 21 comparisons showing a differential positioning behavior. Moreover, for 7 of the 21 comparisons, the gene had a significantly different RRD in the cancer tissue than in its corresponding adjacent normal tissue.

Only a minority of the tested genes underwent significant repositioning in a given cancer tissue (Table II), which suggests that repositioning is not a reflection of global genome reorganization but is more gene specific. Furthermore, in several cases, genes on the same chromosome behaved differently (Figs. 1, 2, and S2; and Table S1). For example, *AKT1*, *HSP90AA1*, and *TGFB3* all map to HSA 14. Although all three genes were repositioned in some cancers (C2, C3, C8, and C10) and none of them repositioned in another cancer (C12), in yet other cancers, only one (C1, C4, C9, and C13) or two (C7, and C11) of the three genes repositioned. Differential repositioning behavior of genes on the same chromosome was also observed for *BRCA1*



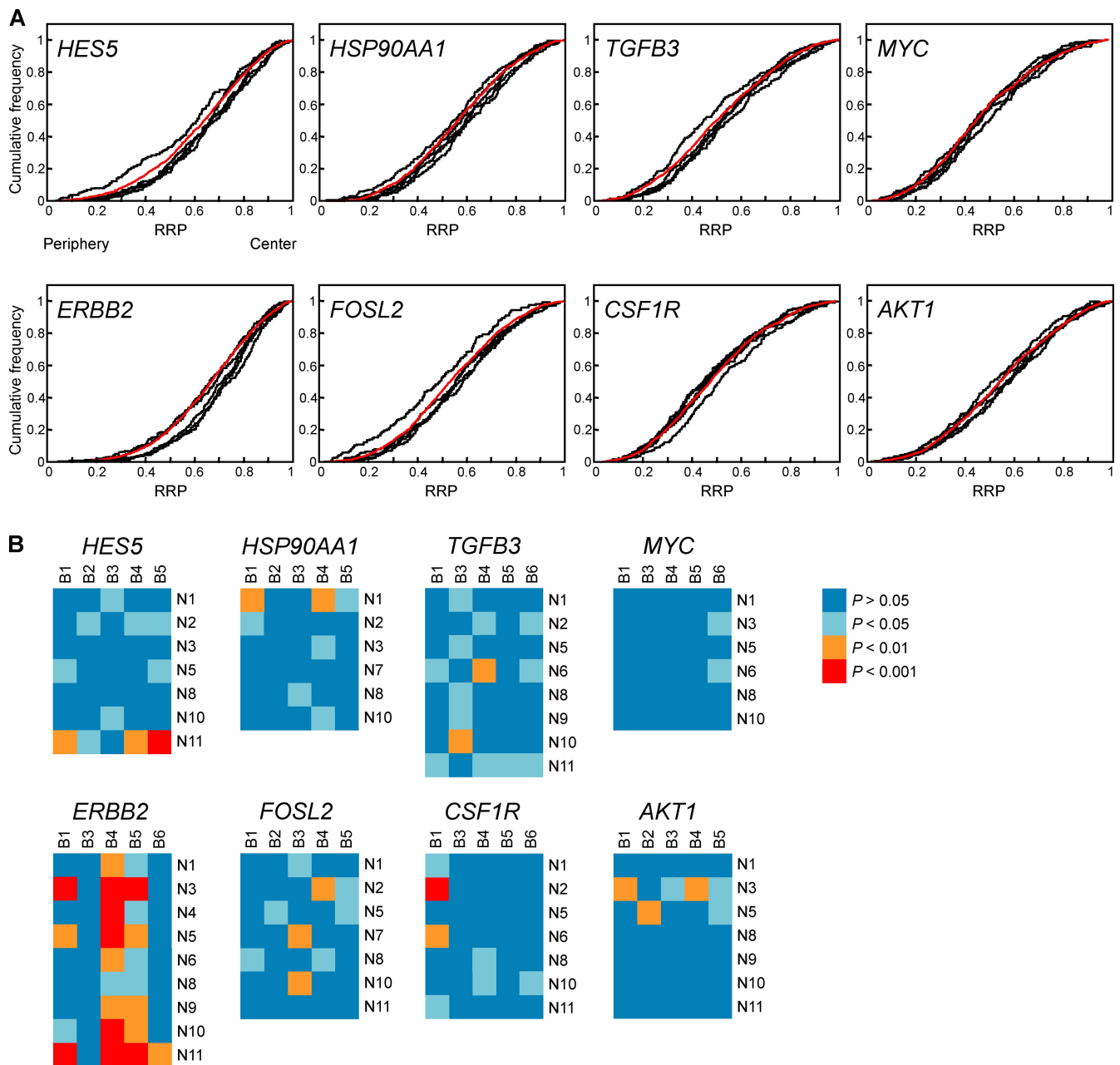


Figure 3. **Gene repositioning is specific to cancer.** (A) Cumulative RRDs for the indicated genes in noncancerous breast tissues (black) and their mean distribution in normal breast tissue (red). (B) Pairwise statistical comparisons of RRDs between individual normal (N1–11) and noncancerous disease (B1–6; B1 and B2, B6 hyperplasia, and B3–5 fibroadenoma) breast tissues using the two-sample 1D KS test.

and *ERBB2* on HSA 17 (C1 and C6), *HES5* and *PTGS2* on HSA 1 (C1 and C4), *MYC* and *HEY1* on HSA 8 (C2), and *CCND1* and *MMP1/3/12* on HSA 11 (C3; Figs. 2 and S2). Further evidence for the specificity of repositioning events is indicated by the finding that individual tumors exhibited distinct numbers and sets of reorganized genes, with the proportion of repositioned genes varying from 18% (3/16 in C1) to 100% (8/8 in C8; Figs. 1, 2, and S2). The apparent cancer-specific repositioning events are not caused by genomic instability, which is often associated with cancer, because repositioning did not correlate with alterations in gene copy number (Table S2;  $P < 0.02$ , using Yates correlated  $\chi^2$  analysis). The only gene that seems to

be an exception is *VEGFA*, which repositioned only in cancers where it was amplified. The degree of genomic instability in a cancer also did not correlate with the number of genes that repositioned (Table S2). Furthermore, the likelihood of a locus to reposition in breast cancer was not related to the gene's genomic context because no correlation between propensity to reposition and gene density in the surrounding genome region was found ( $P = 0.64$ ,  $t$  test; Table S1).

**Repositioning events are specific to cancer**  
 Because genomes can spatially reorganize in disease states other than cancer (Borden and Manuelidis, 1988; Meaburn et al., 2007),

Table III. Comparison of individual cancer and normal tissues to a standard normal distribution

Tissue	<i>HES5</i>	<i>HSP90AA1</i>	<i>TGFB3</i>	<i>MYC</i>	<i>ERBB2</i>	<i>FOSL2</i>	<i>CSF1R</i>	<i>AKT1</i>
C1	<1 × 10 <sup>-6</sup>	0.0042	0.281	ND	3 × 10 <sup>-6</sup>	0.516	0.0079	4.4 × 10 <sup>-4</sup>
C2	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	4.6 × 10 <sup>-4</sup>	<1 × 10 <sup>-6</sup>	0.0012	<1 × 10 <sup>-6</sup>	4 × 10 <sup>-6</sup>	1.7 × 10 <sup>-4</sup>
C3	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	ND	ND	1.6 × 10 <sup>-5</sup>
C4	<1 × 10 <sup>-6</sup>	7 × 10 <sup>-5</sup>	6.1 × 10 <sup>-4</sup>	0.860	0.060	0.028	0.190	0.513
C5	<1 × 10 <sup>-6</sup>	ND	0.0021	ND	0.0043	<1 × 10 <sup>-6</sup>	0.025	5 × 10 <sup>-6</sup>
C6	ND	ND	<1 × 10 <sup>-6</sup>	ND	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>
C7	<1 × 10 <sup>-6</sup>	2.2 × 10 <sup>-4</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	0.0066	<1 × 10 <sup>-6</sup>	0.282
C8	<1 × 10 <sup>-6</sup>	7 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	2.1 × 10 <sup>-5</sup>
C9	<1 × 10 <sup>-6</sup>	0.094	8.9 × 10 <sup>-4</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	0.077	4 × 10 <sup>-6</sup>	0.064
C10	0.0013	1.3 × 10 <sup>-5</sup>	0.0039	0.019	0.242	3 × 10 <sup>-6</sup>	0.095	<1 × 10 <sup>-6</sup>
C11	<1 × 10 <sup>-6</sup>	0.0071	7 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	3.4 × 10 <sup>-4</sup>	1.3 × 10 <sup>-4</sup>	<1 × 10 <sup>-6</sup>	8.8 × 10 <sup>-4</sup>
C12	0.0012	0.049	0.987	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	5.2 × 10 <sup>-4</sup>	0.384
C13	2.6 × 10 <sup>-5</sup>	0.0070	0.0073	<1 × 10 <sup>-6</sup>	0.063	0.424	0.019	0.143
C14	<1 × 10 <sup>-6</sup>	ND	0.099	0.364	0.067	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>	<1 × 10 <sup>-6</sup>
N1	0.766	0.220	0.633	0.949	0.767	0.915	0.390	0.596
N2	0.498	0.997	ND	ND	ND	0.285	0.058	ND
N3	0.954	0.914	0.151	0.276	0.022	ND	ND	0.037
N4	ND	ND	ND	ND	0.342	ND	ND	ND
N5	0.969	ND	0.701	0.998	0.555	0.112	0.917	0.0014
N6	ND	ND	0.353	0.325	0.981	ND	0.315	ND
N7	ND	0.967	ND	ND	ND	0.656	ND	ND
N8	0.763	0.081	0.543	0.178	0.129	0.579	0.158	0.144
N9	ND	ND	0.673	ND	0.510	ND	ND	0.873
N10	0.116	0.889	0.133	0.694	0.995	0.062	0.083	0.924
N11	0.133	ND	0.297	ND	0.0011	0.687	0.569	0.652

Individual p-values for the comparison of RRDs in individual cancer tissues (C1–14) or normal tissues (N1–11) to a pooled normal distribution (two-sample 1D KS test). Orange ( $P < 0.01$ ) and red ( $P < 0.001$ ): significantly different radial positions; blue ( $P > 0.05$ ) and light blue ( $0.01 < P < 0.05$ ): statistically similar distributions. ND, not determined.

it is possible that the repositioning events we detected in the cancer tissues are not specific to carcinogenesis per se, but instead represent a general response to disease. To test this, we analyzed the RRDs of the eight genes that are robustly repositioned in cancer tissues (*HES5*, *HSP90AA1*, *AKT1*, *FOSL2*, *TGFB3*, *ERBB2*, *MYC*, and *CSF1R*) in breast tissue with the noncancerous breast diseases hyperplasia or fibroadenoma (Fig. 3 and Table I). The positioning of genes in the noncancerous disease samples was similar to that in normal tissue, with only 11.2% (32/285) of cross-comparisons showing differences (two-sample 1D KS test;  $P < 0.01$ ). For seven of the eight genes, the rates of significantly different cross-comparisons were low and ranged from 0% (0/30; *MYC*) to 8.6% (3/35; *HES5*, *FOSL2*, and *AKT1*). Similarly to normal tissues, *ERBB2* had a higher rate of differences, with 37.8% (17/45) of the cross-comparisons being statistically different. We conclude that repositioning of most genes is specific to cancer and is not a general indicator of disease.

#### Gene positioning as a cancer marker

Differential radial repositioning of some genes in cancer tissues opens up the possibility of using spatial positioning patterns to identify cancer tissues, including for possible use in diagnostic applications. As a proof-of-principle for this

approach, we sought to test whether the genes that undergo repositioning in cancerous tissue could be used in the identification of cancer samples. To this end, we generated a standard RRD for each gene by pooling positioning data from all available normal tissues (see Materials and methods). We compared the position of genes in our known cancer samples to this standardized normal RRD (Fig. S3 and Table III). The position of *HES5* was significantly different from its standardized distribution in normal tissues in all 13 cancer tissues (Table III). Similarly, the distributions of *HSP90AA1* and *TGFB3* were distinct in 81.8% (9/11) and 78.5% (11/14) of cancers, respectively, compared with the standard distribution in normal samples (Table III). A lower, yet still significant, fraction of 64–73% of cancers showed differential distribution of the remaining five genes (8/11 for *MYC*, 10/14 for *ERBB2*, 9/13 for *FOSL2*, 9/13 for *CSF1R*, and 9/14 for *AKT1*), compared with their standard distribution. The false negative rate, defined as the percentage of cancer tissues exhibiting a gene distribution indistinguishable from the standard normal distribution, ranged from 0% (0/13 cancers) for *HES5* to 35.7% (5/14 cancers) for *AKT1* (Tables III and IV). To determine the false positive rate for the detection of cancer samples, we compared the RRDs of these genes in both normal and noncancerous breast disease tissues to the pooled

Table IV. Single gene false positive and negative rates

Gene	False negatives	Normal tissue false positives	Noncancerous breast disease tissue false positives	Total false positives
<i>HES5</i>	0/13 (0%)	0/7 (0%)	1/5 (20%)	1/12 (8.3%)
<i>HSP90AA1</i>	2/11 (18.2%)	0/6 (0%)	2/5 (40%)	2/11 (18.2%)
<i>TGFB3</i>	3/14 (21.4%)	0/8 (0%)	0/5 (0%)	0/13 (0%)
<i>MYC</i>	3/11 (27.3%)	0/6 (0%)	0/5 (0%)	0/11 (0%)
<i>ERBB2</i>	4/14 (28.6%)	1/9 (11.1%)	3/5 (60%)	4/14 (28.6%)
<i>FOSL2</i>	4/13 (30.8%)	0/7 (0%)	0/5 (0%)	0/12 (0%)
<i>CSF1R</i>	4/13 (30.8%)	0/7 (0%)	0/5 (0%)	0/12 (0%)
<i>AKT1</i>	5/14 (35.7%)	1/7 (14.3%)	0/5 (0%)	1/12 (8.3%)
Total	25/103 (24.3%)	2/57 (3.5%)	6/40 (15.0%)	8/97 (8.2%)

The number (and percentages) of tissues that give either a false negative or false positive result. For a false negative, a gene has a similar RRD in a cancer tissue to that of the pooled normal distribution (two-sample 1D KS test,  $P > 0.01$ ). A false positive is scored when a gene has a statistically different RRD than that of the pooled normal in noncancerous breast tissues ( $P < 0.01$ ).

normal distribution (Tables III–V). In 8.2% (8/97) of comparisons, there were positioning differences, with the frequency of repositioning higher in noncancerous breast disease tissues (15%; 6/40) than in normal tissues (3.5%; 2/57). The incidence of repositioning in noncancerous tissue for individual genes ranged from 0% for four of the genes (*TGFB3*, *MYC*, *FOSL2*, and *CSF1R*) to 28.6% for *ERBB2* (4/14 tissues;  $P \leq 0.0046$ ; Fig. S3 and Tables III–V).

Typical cancer specimens used for diagnosis may contain a mixture of normal and cancer cells, wherein the normal cells would dilute any repositioning events detected and could result in false negatives. To determine the minimal fraction of cancer cells required in a sample, we generated datasets containing varying proportions (10–70%) of cancer nuclei. To this end, images of a total of 160 normal or cancer nuclei were randomly selected from multiple tissues and combined into a single dataset; then the RRD of *HES5* was determined using our standard procedure (see Materials and methods; Fig. S1). Differential positioning of *HES5* could be detected in datasets containing up to 40% normal nuclei ( $P \leq 0.001$ ), demonstrating that tissue heterogeneity does not preclude accurate detection of gene position and identification of cancer tissues.

### Repositioning of gene combinations

Although a single gene, *HES5*, was repositioned in all tested cancer samples, we explored whether combinatorial use of the other seven genes could be used for the detection of cancer samples. Among 21 combinations of two genes, four showed repositioning of at least one of the genes in all tumors (Table VI).

A further 14 combinations showed repositioning of at least one of the genes in  $\geq 85\%$  of the tumors, a rate higher than seven of the eight genes when used individually (Table III). Among the possible 35 combinations of three genes, 22 showed repositioning of at least one gene in all tumors (Table S3). The overall false negative rates of all multiplexed genes was 10.2% (29/284 combinations; range of 0–21.4%) for two genes and 4.9% (19/385; range of 0–14.3%) for three genes (Tables VI and S3). The overall false positive rate for using two genes combined was 18.4% (37/201), and the rate for the three gene combinations was 27.4% (81/296; Tables III and V). As with the single gene analysis, the false positive rate was higher in the noncancerous breast disease tissue than in normal tissue. For the two gene combinations, the false positive rate in normal tissue was 8.1% (9/111), whereas in noncancerous breast disease tissues, it was 31.1% (28/90). For the three gene combinations, the false positive rate in normal tissue was 10.6% (16/151) and 44.8% (65/145) in noncancerous breast disease tissue. Collectively, these results indicate that the spatial positioning of individual or combined genes is a robust method to classify an individual tissue sample as normal or cancerous.

### Discussion

In this study, we have identified several genes that are differentially positioned in invasive breast cancers compared with normal tissue, and we show that determination of their positioning

Table V. Comparison of individual noncancerous breast disease tissues to standard normal distributions

Tissue	<i>HES5</i>	<i>HSP90AA1</i>	<i>TGFB3</i>	<i>MYC</i>	<i>ERBB2</i>	<i>FOSL2</i>	<i>CSF1R</i>	<i>AKT1</i>
B1	0.031	0.0096	0.087	0.320	0.0046	0.142	0.022	0.113
B2	0.049	0.969	ND	ND	ND	0.226	ND	0.468
B3	0.050	0.370	0.035	0.556	0.945	0.022	0.721	0.258
B4	0.104	0.0037	0.017	0.392	$4 \times 10^{-6}$	0.015	0.304	0.112
B5	0.0044	0.156	0.281	0.972	$8 \times 10^{-5}$	0.134	0.914	0.887
B6	ND	ND	0.059	0.054	0.556	ND	0.554	ND

Individual p-values for each comparison between an individual tissue and the pooled normal distribution (two-sample 1D KS test) for a given gene. Orange ( $P < 0.01$ ) and red ( $P < 0.001$ ): significantly different radial positions; blue ( $P > 0.05$ ) and light blue ( $0.01 < P < 0.05$ ): similar radial positions. ND, not determined.



Table VI. Gene multiplexing

Gene	<i>HES5</i>	<i>HSP90AA1</i>	<i>TGFB3</i>	<i>MYC</i>	<i>ERBB2</i>	<i>FOSL2</i>	<i>CSF1R</i>	<i>AKT1</i>
<i>HES5</i>	x	13/13 (100%)	14/14 (100%)	13/13 (100%)	14/14 (100%)	14/14 (100%)	14/14 (100%)	14/14 (100%)
<i>HSP90AA1</i>		x	12/13 (92%)	11/11 (100%)	13/13 (100%)	13/14 (93%)	13/13 (100%)	12/14 (86%)
<i>TGFB3</i>			x	12/13 (92%)	13/14 (93%)	13/14 (93%)	14/14 (100%)	13/14 (93%)
<i>MYC</i>				x	11/14 (79%)	12/13 (92%)	11/13 (85%)	13/14 (93%)
<i>ERBB2</i>					x	12/14 (86%)	11/14 (79%)	12/14 (86%)
<i>FOSL2</i>						x	11/13 (85%)	11/14 (79%)
<i>CSF1R</i>							x	12/14 (86%)
<i>AKT1</i>								x

The number (and percentage) of cancers where at least one of the indicated pair of genes repositioned, compared to the pooled normal distribution. Red boxes indicate a 100% detection rate; pink, >90%; yellow, >80%; and green, >70%. x, not applicable.

pattern reliably detects cancerous tissues. We suggest that interphase spatial genome positioning may be useful for diagnostic applications.

Spatial reorganization of the genome has previously been linked to cancer and genomic instability (Cremer et al., 2003; Taslerová et al., 2003, 2006; Murmann et al., 2005; Wiech et al., 2005, 2009; Petrova et al., 2007; Sengupta et al., 2007; Meaburn and Misteli, 2008). With the exceptions of a few chromosomal translocations, however, the previously described cancer-associated spatial genome repositioning events are relatively minor and often involve large genome regions (Cremer et al., 2003; Wiech et al., 2005; Meaburn and Misteli, 2008; Wiech et al., 2009). Here, we have systematically identified cancer-specific repositioning events of several gene loci. Interestingly, the repositioned genes differ from four previously identified genes that reposition in a 3D culture in vitro model system of early breast cancer (Meaburn and Misteli, 2008). Although *AKT1* and *ERBB2* repositioned in both the 3D cell culture model and in tissue specimens, two other genes, *BCL2* and *VEGFA*, repositioned in the cell culture model but not in a significant proportion of cancer tissues. In contrast, *TGFB3* did not reposition in the cell culture model of cancer, but was repositioned in a large majority of cancer tissues. Notably, we find the degree of repositioning to be generally larger in tissues compared with the cell culture system. In agreement with previous findings in the 3D culture in vitro model system (Meaburn and Misteli, 2008) and with observation of a general conservation of spatial positioning of the genome in cancer cells (Parada et al., 2002; Cremer et al., 2003), we find that the repositioning of genes in breast cancer tissues is gene specific, independent of numerical abnormalities and unrelated to gene density in the proximity of the repositioned gene. Although not explicitly tested here, we have previously found in a 3D breast cancer model system that there is no correlation between gene activity and likelihood of repositioning of an individual locus (Meaburn and Misteli, 2008). We find that the repositioning of many genes is specific to cancer and does not occur in noncancerous breast disease nor within the normal tissue adjacent to a tumor.

Identification of genes that are differentially localized in normal and cancer cells allows for the possibility of using spatial gene positioning as a diagnostic tool. As required for such an application, we find low variability of gene positioning

between individuals. Furthermore, we demonstrate that cancer tissues can accurately be identified by comparison to a standardized normal gene distribution. This is critical for clinical applications because normal tissue is not necessarily available from a proband. Moreover, the fact that the positioning of gene loci in tissue from healthy individuals is identical to that in morphologically normal tissue adjacent to a tumor indicates that normal tissue from a proband can equally serve as a reference.

Cancer detection using spatial genome positioning promises to be a robust method. We discovered a single gene, *HES5*, that detected invasive breast cancer tissue with near 100% accuracy. *HES5* is both a primary target and an effector of the Notch signaling pathway, and has been implicated in cancer (Iso et al., 2003; Hallahan et al., 2004; Liu et al., 2007). The overall rates of false positives and false negatives for single markers were low (8.2% and 24.3%, respectively). In addition, we identified several combinations of two or three multiplexed markers that detect cancerous tissues with no false negative signals and a low rate of false positive outcomes. The false detection rate of this approach compares favorably to the current, standard breast cancer diagnostic tests, such as fine-needle aspiration cytology and core-needle biopsy cytology, for which false positive rates of 0–44% and 0–13.4%, respectively, and false negative rates of 1.3–39% and 4.5–17%, respectively, have been described (Arisio et al., 1998; Young et al., 2002; Chaiwun and Thorner, 2007; Ciatto et al., 2007; Bukhari and Akhtar, 2009); however, these rates, in contrast to our analysis, also include specimen sampling errors. Importantly, however, the broad range in false detection rates in conventional cytological assays is also largely related to the experience of the examining cytopathologist. The use of spatial genome positioning for detection of tumors should reduce human error in making a diagnosis because the method is not based on subjective criteria but gives a quantifiable readout, making it independent of the expertise of the individual performing the analysis. A distinct advantage of this approach is the very small quantity of material required. Differences in spatial positioning were routine detected by analysis of 100–200 cells and despite inherent heterogeneity in samples, although a requirement appears to be that at least 60% of the nuclei analyzed are cancerous. This approach is suitable for adaptation in a routine laboratory setting, as all individual steps of the procedure rely on standard methods including paraffin

embedding of biopsy material, FISH detection, and image analysis methods. We obtained the required 100–200 cells for analysis from a single 5- $\mu$ m-thick slice of a biopsy section or 2.5-mm tissue microarray (TMA) core, and we typically used 12–30 image fields containing a total of  $\sim$ 200–1,000 cells to obtain the required 100–200 analyzable cells. Imaging of a typical sample takes no longer than 60 min, and can be automated.

If validated in a larger number of samples, we envision that this approach may be a useful first molecular indicator of cancer after an abnormal mammogram using tissue from a core needle biopsy, and would be used in combination with standard pathological indicators such as gene amplification. Interestingly, *ERBB2* and *MYC*, which are currently both screened for amplification status by FISH in the diagnosis of breast cancer, were both repositioned in a large proportion of cancer tissues. The proof of concept we describe here lays the foundation for future studies to examine whether gene positioning analysis will reveal differences not apparent by gross morphological changes detected by pathologists and whether it will be useful in identifying early stage cancers. Moreover, the observed variability in gene repositioning patterns between individual cancers in our analysis hints at the promise of spatial positioning patterns to go beyond simply discriminating cancerous from normal tissues and implies the possibility of a prognostic value, such as distinguishing between cancer subtypes or survival outcomes. Finally, our method of cancer diagnosis is not limited to breast cancer and may be applied to any cancer type in which repositioned genes can be identified.

## Materials and methods

### FISH

4–5- $\mu$ m-thick, formalin-fixed, paraffin-embedded human breast tissue sections containing morphologically normal tissue, invasive carcinoma, or noncancerous breast disease (hyperplasia and fibroadenoma) were obtained from the AIDS and Cancer Specimen Resource or purchased from US Biomax, Inc., Imgenex Corp., Capital BioSciences, Inc., and BioChain Institute, Inc. (Table I). The tissues were in the form of standard single tissue slides, with the exceptions of tissues N6, C7–13, and B1–6, which formed part of a TMA (catalog no. Z7020010, core size of 2.5 mm; BioChain Institute, Inc.). The FISH procedure was identical for all specimens, except where stated otherwise.

To generate FISH probes, the bacterial artificial chromosome (BAC) clones detailed in Table S1 (BACPAC Resources Center) were labeled by nick translation with dUTPs conjugated with either biotin (Roche) or digoxigenin (Roche) as described in detail in Parada et al. (2002b). Dual-probe FISH experiments were routinely performed using 600–800 ng each of digoxigenin- and biotin-labeled probe DNA, 10  $\mu$ g of human COT1 (Roche), and 40  $\mu$ g tRNA (Sigma-Aldrich), then resuspended in 10  $\mu$ l of hybridization mix (10% [wt/vol] dextran sulfate [Sigma-Aldrich], 50% [vol/vol] formamide [Sigma-Aldrich], 2 $\times$  SSC, and 1% [wt/vol] Tween 20 [Sigma-Aldrich]).

Slides were dewaxed by two 30-min incubations in Xylene (Mallinckrodt Baker, Inc.). Immediately before this treatment, TMA slides were baked at 60°C for 1 h. The tissues were rehydrated with sequential 5-min incubations in an ethanol series (100%, 90%, and 70%) followed by 10 min in PBS. Slides were then boiled in the microwave (1,000 W for 10 min in 700 ml of solution) in 0.01 M sodium citrate, pH 6, and left at RT until cooled to  $\sim$ 37°C. Subsequently, the tissue sections were incubated at 37°C in 10  $\mu$ g/ml RNase A (Sigma-Aldrich)/2 $\times$  SSC for 15 min and washed for 5 min in PBS. After this, the tissues were subjected to incubation with 0.25 mg/ml proteinase K (Sigma-Aldrich) at 37°C for 10 min 30 s for TMA slides, 11 min for individual morphologically normal breast tissues, and 10 min for the single cancerous tissues. After a brief rinse in PBS, the slides were taken through a dehydrating ethanol series (70%, 90%, and 100%; 5 min each) and left to air dry.

The nuclei were co-denatured with the FISH probes at 85°C for 10 min and left to hybridize overnight at 37°C in a humid container. The next day, three 5-min washes in 50% formamide/2 $\times$  SSC at 45°C followed by three 5-min washes in 1 $\times$  SSC at 60°C were performed, and the slides were placed in 0.1% Tween 20/4 $\times$  SSC at RT to cool. To block, the slides were incubated for 15–20 min in 3% bovine serum albumin (Sigma-Aldrich)/0.1% Tween 20/4 $\times$  SSC. Detection antibody anti-digoxigenin-rhodamine (Roche) and fluorescein-avidin DN (Vector Laboratories) were diluted 1:200 in blocking solution and incubated with the tissues for 2 h at 37°C. Slides were mounted in DAPI-containing Vectashield mounting medium (Vector Laboratories) after three 5-min washes in 0.1% Tween 20/4 $\times$  SSC at RT.

### Image acquisition

For analysis of the FISH signals, tissue sections were imaged using a microscope (IX70; Olympus) controlled by a Deltavision System (Applied Precision) with SoftWoRx 3.5.1 software (Applied Precision) and fitted with a charge-coupled device camera (CoolSnap; Photometrics). We used a 60 $\times$ , 1.4 NA oil objective lens and an auxiliary magnification of 1.5. Z stacks were acquired to cover the thickness of the tissue section with a step size of 0.2  $\mu$ m or 0.5  $\mu$ m in the z direction. The increase of step size to 0.5  $\mu$ m gave identical results to 0.2  $\mu$ m. Images were acquired at a 1,024  $\times$  1,024 pixel resolution, with a pixel size equivalent to 0.07427  $\mu$ m in x and y. Deconvolution was performed on the image stacks using the following settings of SoftWoRx 3.5.1: enhance ratio (aggressive) method, 10 cycles, medium noise filtering, border rolloff (voxel) set to 16, size for z transformation set to 128, Wiener filter enhancement set at 0.9, Wiener filter smoothing set at 0.8, and intensity scale factor set at 1. Maximum intensity projections of the deconvolved stacks were generated and analyzed for the radial distribution of the FISH signals, as described in the next section.

### Quantitative analysis of FISH signal distributions

Initially, images were contrast-enhanced based on visual inspection, and individual cell nuclei from the blue color channel were manually delineated using Photoshop 7.0 (Adobe). Manual segmentation was required to avoid nuclei that were overlapping and because of the close proximity of nuclei in many tissues. The automatic detection of FISH signals was performed as described previously (Meaburn and Misteli, 2008; Takizawa et al., 2008a); in brief, a three-stage process was used involving: (1) noise reduction, (2) segmentation, and (3) post-processing. (1) Background noise was removed in each channel by applying an adaptive nonlinear noise reduction technique (“SUSAN”; Smith and Brady, 1997). (2) A fuzzy-C-means clustering algorithm was applied on the noise-reduced images to probabilistically assign each image picture element (pixel) into two classes. The two classes corresponded to background and objects in the image. The images from this process were segmented into binary images whereby each pixel with >50% probability of being in the object class was classified as corresponding to actual objects in the specimen, whereas the remaining pixels were classified as background. The integrated intensities of each group of contiguous object pixels were calculated, and those groups that exceeded a threshold value, which was calculated automatically by the isodata threshold method in DIPimage toolbox (Technical University of Delft), were considered to correspond to individual FISH signals.

For quantification of the spatial distributions of FISH signals, the following procedure was used: for each segmented nucleus, the binary Euclidean distance transform (EDT) was computed (Danielsson, 1980). The EDT is a morphological operation that assigns each pixel in the nucleus a Euclidean distance to the closest boundary point; i.e., the EDT value assigned to each pixel in a segmented nucleus equals the shortest distance to the edge of the nucleus. To account for variations in nuclear size, this distance-transformed image was normalized with the maximum EDT value for the given nucleus, such that the normalized EDT values varied between 0 (nuclear periphery) and 1 (nuclear center). The position of each FISH signal was defined by its geometric gravity center, and the normalized EDT value corresponding to that position was used to determine the relative radial position of each signal. Using this method, no assumption regarding nuclear shape is made when determining the radial position of a gene. In addition, the number of red and green FISH signals in each nucleus was recorded. All analysis tools were implemented using custom software written in MATLAB (Mathworks, Inc.) with DIPimage toolbox.

### Statistical analysis

All alleles in a nucleus were included for analysis, and nuclei were included regardless of the number of alleles present for a gene, unless the

nucleus contained no FISH signals for either gene. The relative radial positions of FISH signals across multiple nuclei of the same specimen were combined to generate RRDs. The RRDs from different samples were then compared using a two-sample 1D KS test. To account for the multiple comparisons, differences were considered significant if there was <1% probability that the two distributions arose from the same parent distribution ( $P < 0.01$ ). 100 or 200 randomly chosen nuclei from a dataset gave highly similar results to 526 or 876 nuclei from the same dataset ( $0.88 \leq P \leq 0.96$ ; Fig. S1). Although 25 or 50 nuclei did not reach significance when compared with the full dataset of 526 or 876 nuclei ( $0.22 \leq P \leq 0.68$ ), a greater variability was seen; thus, 110–220 nuclei per gene per tissue were analyzed (200–220 nuclei for *AKT* and *TGFB3* in N8–10, *BCL2* and *VEGFA* in N4 and N10, *CCND1* and *ERBB2* in N4, N8–10, and for these six genes in C2, C5 and C6; *ERBB2* was analyzed in 161 nuclei in C13; 110–143 nuclei were analyzed for the remaining datasets), with the exception of *HSP90AA1* in C11, where 88 nuclei were analyzed. The number of nuclei for the pooled normal datasets was as follows: *HES5*,  $n = 921$  nuclei from a total of seven individuals; *HSP90AA1*,  $n = 797$  nuclei from six individuals; *AKT1*,  $n = 1,149$  nuclei from seven individuals; *FOSL2*,  $n = 919$  nuclei from seven individuals; *TGFB3*,  $n = 1,269$  nuclei from eight individuals; *ERBB2*,  $n = 1,406$  nuclei from nine individuals; *CSFR1*,  $n = 885$  nuclei from seven individuals; and *MYC*,  $n = 756$  nuclei from six individuals. Pilot experiments demonstrated high reproducibility between repeat experiments (typically  $0.65 \leq P \leq 0.81$ ).

Datasets with varying ratios of normal and cancer nuclei, for the gene *HES5*, were blindly generated and analyzed. Each dataset contained a total of 160 individual nuclei with varying ratios (10–70%) of cancer/normal cells. A master dataset containing 200 randomly selected nuclei each from known normal or cancer samples, where *HES5* had been detected, was first generated. To ensure there was no biasing toward a particular tissue, the randomly selected nuclei were taken from four different normal tissues and 12 different cancer tissues, with approximately equal numbers of nuclei used from each tissue of the same type (~50 nuclei per tissue for normal and ~17 nuclei per tissue for cancer tissues). From these pools of nuclei datasets with varying ratios of cancer, nuclei were generated. The person mixing the nuclei had no prior knowledge on which of the master datasets contained normal and which contained the cancer nuclei. Furthermore, the individual performing the RRD analysis was blinded as to the ratio of nuclei from each of the master datasets. Cumulative RRDs from these datasets were generated and compared with the pooled normal distribution of *HES5* using the two-sample 1D KS test.  $P < 0.01$  was considered significant.

Statistical significance between the incidence of repositioning and changes to copy number were determined by Yates correlated  $\chi^2$  analysis, where a  $P$  value  $< 0.05$  was considered significant. The null hypothesis being tested was that a gene repositioned in cancer due only to genomic instability. Data were used from all 20 genes. For this analysis, a gene was considered to reposition in a cancer tissue if the majority of cross-comparisons between the individual cancer tissue and the individual normal tissues were significantly different ( $P < 0.01$ ). A gain in copy number for a gene was assumed if >20% of nuclei in a tissue had three or more signals; a loss of copy number for a gene was recorded if  $\geq 40\%$  of nuclei in the tissue had only one allele detected. Using these cut-off values, normal tissues were never classified as having a copy number variation.

To probe for an effect of gene density on repositioning probability, the number of genes 1 MB on either side a gene were calculated using data from [http://www.ncbi.nlm.nih.gov/mapview/map\\_search.cgi?taxid=9606](http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=9606). The statistical significance between the local gene densities for genes that have a propensity to reposition in cancer tissues compared with normal (*HES5*, *HSP90AA1*, *TGFB3*, *MYC*, *ERBB2*, *FOSL2*, *CSFR1*, and *AKT1*) and those loci that do not (*VEGFA*, *CCND1*, *HES1*, *PTGS2*, *MMP1/3/12*, *ZNF217*, *BCL2*, *HEY1*, *BRAC1*, *PTEN*, *TLE1*, and *TPI1*) was assessed using a two-tailed Student's  $t$  test in Excel (Microsoft).

#### Online supplemental material

Fig. S1 shows the determination of appropriate sample size and of the minimal fraction of cancer cells required for analysis. Fig. S2 contains additional data showing loci-specific reorganization of the genome in breast cancer. Fig. S3 shows standardized pooled normal distributions, with individual normal and cancer specimens for comparison. Table S1 provides additional information about the candidate genes. Table S2 demonstrates that repositioning events are not due to numerical changes. Table S3 provides false negative rates based on combinations of three genes. Online supplemental material is available at <http://www.jcb.org/cgi/content/full/jcb.200909127/DC1>.

We are grateful to the AIDS and Cancer Specimen Resource (ACSR) for providing tissue sections and to Delft University (Netherlands) for providing the DIPimage. We thank Tatiana Karpova and Ty Voss for help with microscopy. Fluorescence imaging was performed at the National Cancer Institute Fluorescence Imaging Facility.

This work was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research, and with federal funds from the National Cancer Institute, National Institutes of Health, under Contract HHSN261200800001E. K.J. Meaburn was supported by an Idea Award from the Department of Defense (BC073689).

Submitted: 21 September 2009

Accepted: 6 November 2009

## References

- Arisio, R., C. Cuccorese, G. Accinelli, M.P. Mano, R. Bordon, and L. Fessia. 1998. Role of fine-needle aspiration biopsy in breast lesions: analysis of a series of 4,110 cases. *Diagn. Cytopathol.* 18:462–467. doi:10.1002/(SICI)1097-0339(199806)18:6<462::AID-DC16>3.0.CO;2-F
- Borden, J., and L. Manuelidis. 1988. Movement of the X chromosome in epilepsy. *Science.* 242:1687–1691. doi:10.1126/science.3201257
- Bukhari, M.H., and Z.M. Akhtar. 2009. Comparison of accuracy of diagnostic modalities for evaluation of breast cancer with review of literature. *Diagn. Cytopathol.* 37:416–424. doi:10.1002/dc.21000
- Chaiwun, B., and P. Thorner. 2007. Fine needle aspiration for evaluation of breast masses. *Curr. Opin. Obstet. Gynecol.* 19:48–55. doi:10.1097/GCO.0b013e328011f9ae
- Chambeyron, S., and W.A. Bickmore. 2004. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* 18:1119–1130. doi:10.1101/gad.292104
- Ciatto, S., N. Houssami, D. Ambrogetti, S. Bianchi, R. Bonardi, B. Brancato, S. Catarzi, and G.G. Risso. 2007. Accuracy and underestimation of malignancy of breast core needle biopsy: the Florence experience of over 4000 consecutive biopsies. *Breast Cancer Res. Treat.* 101:291–297. doi:10.1007/s10549-006-9289-6
- Cremer, M., K. Küpper, B. Wagler, L. Wizelman, J. von Hase, Y. Weiland, L. Kreja, J. Diebold, M.R. Speicher, and T. Cremer. 2003. Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *J. Cell Biol.* 162:809–820. doi:10.1083/jcb.200304096
- Cremer, T., M. Cremer, S. Dietzel, S. Müller, I. Solovei, and S. Fakan. 2006. Chromosome territories—a functional nuclear landscape. *Curr. Opin. Cell Biol.* 18:307–316. doi:10.1016/j.ccb.2006.04.007
- Danielsson, P.-E. 1980. Euclidean distance mapping. *Computer Graphics and Image Processing.* 14:227–248. doi:10.1016/0146-664X(80)90054-4
- Foster, H.A., and J.M. Bridger. 2005. The genome and the nucleus: a marriage made by evolution. Genome organisation and nuclear architecture. *Chromosoma.* 114:212–229. doi:10.1007/s00412-005-0016-6
- Hallahan, A.R., J.I. Pritchard, S. Hansen, M. Benson, J. Stoeck, B.A. Hatton, T.L. Russell, R.G. Ellenbogen, I.D. Bernstein, P.A. Beachy, and J.M. Olson. 2004. The SmoA1 mouse model reveals that notch signaling is critical for the growth and survival of sonic hedgehog-induced medulloblastomas. *Cancer Res.* 64:7794–7800. doi:10.1158/0008-5472.CAN-04-1813
- Hewitt, S.L., F.A. High, S.L. Reiner, A.G. Fisher, and M. Merkenschlager. 2004. Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *Eur. J. Immunol.* 34:3604–3613. doi:10.1002/eji.200425469
- Iso, T., L. Kedes, and Y. Hamamori. 2003. HES and HERP families: multiple effectors of the Notch signaling pathway. *J. Cell. Physiol.* 194:237–255. doi:10.1002/jcp.10208
- Kosak, S.T., J.A. Skok, K.L. Medina, R. Riblet, M.M. Le Beau, A.G. Fisher, and H. Singh. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science.* 296:158–162. doi:10.1126/science.1068768
- Liu, J., F. Ye, H. Chen, W. Lü, C. Zhou, and X. Xie. 2007. Expression of differentiation associated protein Hes1 and Hes5 in cervical squamous carcinoma and its precursors. *Int. J. Gynecol. Cancer.* 17:1293–1299. doi:10.1111/j.1525-1438.2007.00930.x
- Meaburn, K.J., and T. Misteli. 2008. Locus-specific and activity-independent gene repositioning during early tumorigenesis. *J. Cell Biol.* 180:39–50. doi:10.1083/jcb.200708204
- Meaburn, K.J., E. Cabuy, G. Bonne, N. Levy, G.E. Morris, G. Novelli, I.R. Kill, and J.M. Bridger. 2007. Primary laminopathy fibroblasts display altered genome organization and apoptosis. *Aging Cell.* 6:139–153. doi:10.1111/j.1474-9726.2007.00270.x

- Misteli, T. 2007. Beyond the sequence: cellular organization of genome function. *Cell*. 128:787–800. doi:10.1016/j.cell.2007.01.028
- Murmann, A.E., J. Gao, M. Encinosa, M. Gautier, M.E. Peter, R. Eils, P. Lichter, and J.D. Rowley. 2005. Local gene density predicts the spatial position of genetic loci in the interphase nucleus. *Exp. Cell Res.* 311:14–26. doi:10.1016/j.yexcr.2005.07.020
- Parada, L.A., P.G. McQueen, P.J. Munson, and T. Misteli. 2002. Conservation of relative chromosome positioning in normal and cancer cells. *Curr. Biol.* 12:1692–1697. doi:10.1016/S0960-9822(02)01166-1
- Parada, L.A., C. Elbi, M. Dundr, and T. Misteli. 2002b. Gene Expression. In *Essential Cell Biology*. Vol. 2. J. Davey, and M. Lord editors. Oxford University Press, Oxford. 47–77.
- Petrova, N.V., I.I. Yakutenko, A.V. Alexeevski, V.A. Verbovoy, S.V. Razin, and O.V. Iarovaia. 2007. Changes in chromosome positioning may contribute to the development of diseases related to X-chromosome aneuploidy. *J. Cell. Physiol.* 213:278–283. doi:10.1002/jcp.21118
- Schneider, R., and R. Grosschedl. 2007. Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev.* 21:3027–3043. doi:10.1101/gad.1604607
- Sengupta, K., M.B. Upender, L. Barenboim-Stapleton, Q.T. Nguyen, S.M. Wincovitch Sr., S.H. Garfield, M.J. Difilippantonio, and T. Ried. 2007. Artificially introduced aneuploid chromosomes assume a conserved position in colon cancer cells. *PLoS One*. 2:e199. doi:10.1371/journal.pone.0000199
- Smith, S.M., and J.M. Brady. 1997. SUSAN - a new approach to low level image processing. *Int. J. Comput. Vis.* 23:45–78. doi:10.1023/A:1007963824710
- Takizawa, T., P.R. Gudla, L. Guo, S. Lockett, and T. Misteli. 2008a. Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. *Genes Dev.* 22:489–498. doi:10.1101/gad.1634608
- Takizawa, T., K.J. Meaburn, and T. Misteli. 2008b. The meaning of gene positioning. *Cell*. 135:9–13. doi:10.1016/j.cell.2008.09.026
- Taslerová, R., S. Kozubek, E. Lukášová, P. Jirsová, E. Bártoová, and M. Kozubek. 2003. Arrangement of chromosome 11 and 22 territories, EWSR1 and FLI1 genes, and other genetic elements of these chromosomes in human lymphocytes and Ewing sarcoma cells. *Hum. Genet.* 112:143–155.
- Taslerová, R., S. Kozubek, E. Bártoová, P. Gajdusková, R. Kodet, and M. Kozubek. 2006. Localization of genetic elements of intact and derivative chromosome 11 and 22 territories in nuclei of Ewing sarcoma cells. *J. Struct. Biol.* 155:493–504. doi:10.1016/j.jsb.2006.05.005
- Wiech, T., S. Timme, F. Riede, S. Stein, M. Schuricke, C. Cremer, M. Werner, M. Hausmann, and A. Walch. 2005. Human archival tissues provide a valuable source for the analysis of spatial genome organization. *Histochem. Cell Biol.* 123:229–238. doi:10.1007/s00418-005-0768-3
- Wiech, T., S. Stein, V. Lachenmaier, E. Schmitt, J. Schwarz-Finsterle, E. Wiech, G. Hildenbrand, M. Werner, and M. Hausmann. 2009. Spatial allelic imbalance of BCL2 genes and chromosome 18 territories in nonneoplastic and neoplastic cervical squamous epithelium. *Eur. Biophys. J.* 38:793–806. doi:10.1007/s00249-009-0474-5
- Young, N.A., D.R. Mody, and D.D. Davey. 2002. Diagnosis and subclassification of breast carcinoma by fine-needle aspiration biopsy: results of the interlaboratory comparison program in non-gynecologic cytopathology. *Arch. Pathol. Lab. Med.* 126:1453–1457.
- Zink, D., A.H. Fischer, and J.A. Nickerson. 2004. Nuclear structure in cancer cells. *Nat. Rev. Cancer.* 4:677–687. doi:10.1038/nrc1430