Original Article

TITLE

A predictive factor of microarray comparative genomic hybridization analysis quality for formalin-fixed paraffin-embedded archival tissue.

Kenjiro Nakao, MD,[1] Masahiro Oikawa, MD, PhD,[1,2] Junichi Arai, MD,[1] Zhanna Mussazhanova, MD,[3] Hisayoshi Kondo, PhD,[4] Kazuko Shichijo, PhD,[3] Masahiro Nakashima, MD, PhD,[3] Tomayoshi Hayashi, MD, PhD,[5] Koh-ichiro Yoshiura, MD, PhD,[6] Toshiko Hatachi, MD,[1] Takeshi Nagayasu, MD, PhD,[1].

[1]Department of Surgical Oncology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [2]Center for Industry, University and Government Cooperation, Nagasaki University, Nagasaki, Japan. [3]Department of Tumor and Diagnostic Pathology, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [4]Biostatistics Section, Division of Scientific Data Registry, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [5]Department of Pathology, Nagasaki University Hospital, Nagasaki, Japan. [6]Department of Human Genetics, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan.

ADDRESS FOR REPRINTS AND CORRESPONDANCE:

Masahiro Oikawa, Department of Surgical Oncology, Nagasaki University Graduate School of Biomedical Science, 1-7-1 Sakamoto, Nagasaki 852-8523, Japan

Tel.: +81-95-819-7304

Fax: +81-95-819-7306

E-mail: oimasa@iris.dti.ne.jp

DISCLOSURE OF FUNDING

This work was supported by Grant-in-Aid for Young Scientists (B) from Japan Society for the Promotion of Science KAKENHI (No. 24791382).

CONFLICTS OF INTERESTS

None declared.

ABSTRACT

Utilizing formalin-fixed paraffin-embedded (FFPE) archival tissue, the most common form of tissue preservation in routine practice, for cytogenetic analysis using microarray comparative genomic hybridization (aCGH) remains challenging. We searched for a predictive factor of the performance of FFPE DNA in aCGH analysis. DNA was extracted from 63 FFPE archival tissue samples of various tissue types (31 breast cancers, 24 lung cancers and 8 thyroid tumors), followed by aCGH analysis using high-density oligonucleotide microarrays. Tumor DNA from matched frozen samples, and from FFPE samples following whole genome amplification (WGA) were also analyzed in two and four case, respectively. The derivative log ratio spread (DLRSpread) was used to assess the overall quality of each aCGH result. The DLRSpread correlated significantly with the double-stranded DNA ratio of tumor DNA, storage time and the degree of labeling with Cy5 ($P < 0.0001$, correlation coefficients = -0.796, 0.551, -0.481, respectively). Stepwise multiple linear regression analysis revealed that the double-stranded DNA ratio of tumor DNA is the most significant predictive factor of DLRSpread (regression coefficient = -0.4798, P-value = < 0.0001). The cytogenetic profiles of FFPE and matched frozen samples showed good concordance. Although the double-stranded DNA ratios were increased after WGA, the DLRSpread was not improved. The double-stranded DNA ratio can be used to predict aCGH analysis performance for DNA from FFPE samples. Using this quality metric, valuable FFPE archival tissue samples can be utilized for aCGH analysis.

KEY WORDS

Formalin-fixed paraffin-embedded
Microarray
Comparative genomic hybridization
Cancer
Cytogenetics

INTRODUCTION

Comparative genomic hybridization (CGH) is an especially useful method for detecting chromosomal instability in tumor genomes, an important hallmark of cancer(1). Recent technological advances have enabled higher resolution CGH analysis through the use of high-density microarrays(2-10). However, utilizing formalin-fixed paraffin-embedded (FFPE) archival tissue for aCGH remains challenging because of DNA degradation, cross-linking between nucleic acid strands, formation of DNA adducts with histones or nucleic acid binding proteins, and breaking and depurination of DNA(11, 12).

Several groups have demonstrated the feasibility of performing aCGH analysis on DNA extracted from FFPE samples(13-20). FFPE is the most common form of tissue preservation in routine practice and these samples are often associated with detailed pathological data and clinical outcomes. Regarding the technical improvement in aCGH analysis with FFPE tissue, DeVries et al. reported that whole genome amplification (WGA) of DNA isolated from FFPE tissue, by random priming, is robust and reproducible(21). Lyons-Weiler et al. presented a modified protocol for single nucleotide polymorphism arrays and generated comparable results for fresh frozen tissues(22). Meanwhile, Hostetter et al. reported an improvement in aCGH analysis results for long-oligonucleotide microarrays when using DNase treatment to generate randomly fragmented DNA(23). Furthermore, Wang et al. showed that molecular inversion probe microarrays required small intact DNA and was suitable for aCGH analysis using FFPE samples(24). Recently, a one-step chemical labeling method, called the Universal Linkage System (ULS), was developed where DNA is fluorescently labeled at the N7 position of guanine without the need for enzymatic reaction, which is susceptible to the effects of DNA degradation, yielding precise, robust and high-quality aCGH data(25-27). However, a suitable predictor of the performance of DNA from FFPE samples in aCGH, which will prevent the wasting of sparse clinical samples and economic resources, is still lacking. Certain research groups have reported that a prequalifying PCR test can predict the performance of FFPE DNA on microarrays better than FFPE sample age(28, 29).

In this retrospective observational study, we search for a factor to predict performance of FFPE DNA in aCGH analysis by dissecting aCGH analysis results from 63 archival FFPE tissue samples and show that the ratio of double-strand DNA to total DNA (dsDNA ratio) is able to predict the performance of FFPE DNA in aCGH analysis better than other factors such as FFPE sample age. This quality metric simplifies aCGH quality control when using FFPE archival tissue samples, thereby improving cost effectiveness.

MATERIALS AND METHODS

Tumor samples and clinical information

Sixty-three formalin fixed paraffin embedded (FFPE) tissue samples were obtained from the Department of Pathology, Nagasaki University Hospital. The samples included 31 invasive breast cancers, 24 adenocarcinoma of the lung and 8 thyroid follicular tumors. Twenty-eight of the 31 invasive breast cancers and the 24 lung cancers formed part of the data set that we previously reported(30, 31). Resected specimens were immediately fixed in 10% neutral buffered formalin for 24 hours, processed and embedded into paraffin blocks. Routine Hematoxylin and Eosin (H&E) stains were performed on 3μm sections of tissue cut from the FFPE blocks. Pathological diagnoses were made by independent pathologists. For the breast cancer samples, Fluorescence in situ hybridization (FISH) analysis of the HER2 gene was conducted. A minimum 2-fold increase in HER2 signals over chromosome enumeration probe 17 (CEP17) signals in cancer cells was considered positive for gene amplifications. Storage times for FFPE samples ranged between 1 and 43 years, with a mean of 14.5 years. A summary of these samples is provided in Supplementary Table 1. Matched fresh frozen tissue samples were available for two of the invasive breast cancer samples. All experimental procedures for this study were approved by the Committee for Ethical Issues on the Human Genome and Gene Analysis in Nagasaki University.

DNA extraction

Tumor DNA was extracted from each FFPE sample, as previously reported(32). Briefly, samples were macrodissected using between 10 and 20 10-μm-thick sections. Tumor tissue areas containing more than 70% tumor cells, as identified by a guide slide stained with hematoxylin, were selected and manually dissected using surgical scalpels. Paraffin removal was performed in 80% xylene after which samples were washed twice with absolute ethanol. Deparaffinized tissue pieces were spun down, after which the dried up pellet was resuspended in 360μL buffer ATL (QIAmp DNA Mini Kit, Qiagen, Dusseldorf, Germany) and incubated at 95 °C for 15 min. Samples were allowed to cool down to room temperature and subsequently digested with proteinase K for 3 days at 56 °C in a rotation oven with periodic mixing and adding of fresh proteinase K every 24 h. DNA was extracted using the QIAmp DNA Mini Kit according to the manufacturer's instructions with some modifications. Briefly, 400μL buffer AL was added to the sample, which was then incubated at 70 °C for 10 min followed by the addition of 400μL

absolute ethanol and mixing by vortex. The sample solution was placed into the spin column followed by centrifugation for 1 min at 8000 ×g. The spin column was washed twice with 500μL buffer AW1 by centrifugation at 8000×g for 1 min and washed with 80% ethanol by centrifugation at 14,000 ×g for 3 min. Finally, the DNA was eluted with 55μL nuclease free water.

For frozen tissue samples, three to four pierce of tumor tissue were collected and DNA was isolated using the QIAmp DNA Mini Kit according to the manufacturer's instructions. For 4 of the breast cancer samples, 100ng DNA extracted from FFPE tissue was used in each case for whole-genome amplification using the GenomePlex Complete Whole Genome Amplification Kit (Sigma-Aldrich, St. Louis, MO) followed by purification using the GenElute PCR Clean-up Kit (Sigma-Aldrich) following the manufacturer's instructions, so that the ratio of double-strand to total DNA (dsDNA ratio) of the samples increased.

DNA was quantitated using the ratio of absorbance at 260 and 280nm (A260/A280), and 260 and 230nm (A260/A230) on a Nano Drop ND-2000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). The double-strand DNA concentration in each sample was quantitated on a Qubit fluorometer with the Qubit dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA) following the manufacturer's instructions.

aCGH analysis

The Genomic DNA ULS Labeling Kit (Agilent Technologies, Palo Alto, CA) was used to chemically label 500ng of DNA from each sample and 250ng of reference female genomic DNA (Promega, Madison, WI) with Cy5 or Cy3 dye for 30 min at 85 °C, respectively, followed by purification using Agilent-KREApure™ columns. The amount of input DNA was determined on a Qubit fluorometer using the Qubit dsDNA HS Assay Kit. The degree of Cy5 labeling (absorbance at 650nm) was calculated using a Nano Drop ND-2000 spectrophotometer. Purified, labeled samples were then combined and mixed with human Cot-1 DNA (Invitrogen, Carlsbad, CA), Agilent 10× Blocking Agent and Agilent 2× Hybridization Solution. Prior to array hybridization, hybridization mixtures were denatured at 95 °C for 3 min and incubated at 37 °C for 30 min. Agilent CGHblock was added and samples were hybridized to the SurePrint G3 Human CGH 8×60K Microarray, which contains 8 identical arrays consisting of ~63,000 in situ synthesized 60-mer oligonucleotide probes that span coding and noncoding sequences with an average spatial resolution of ~54 kb. Hybridization was carried out at 65 °C for 40 h before washing with Agilent Oligo aCGH Wash Buffer 1 at room temperature for 5 min, followed by washing with Agilent Oligo aCGH Wash Buffer 2 at 37 °C for 1 min.

Scanning and image analysis were done on an Agilent DNA Microarray Scanner. Agilent Feature Extraction Software (version 9.5) was used for data extraction from raw microarray image files. Agilent Genomic Workbench (version 5.0) was used to visualize, detect and analyze chromosomal patterns using the Aberration Detection Method 2 (ADM-2) algorithm with the default settings. The derivative log ratio spread (DLRSpread) of each sample was also calculated using Agilent Genomic Workbench (version 5.0). A copy number gain or loss was defined as a log2 ratio > 0.25 or < -0.25, respectively.

Statistical analysis

The DLRSpread, which estimates the log ratio noise by calculating the spread of log ratio differences between consecutive probes along all chromosomes, was used to assess the overall quality of each aCGH result, thus, lower is better.

To search for predictive factors of the performance of FFPE DNA in aCGH analysis, the Spearman rank-correlation was used to evaluate the relationship between DLRSpread and A260/A280 ratio, A260/A230 ratio, dsDNA ratio, storage time and degree of labeling (Cy5) for each sample. For the DLRSpread, multiple regression analysis with stepwise selection was used to examine the associations with A260/A280 ratio, A260/A230 ratio, dsDNA ratio, storage time and degree of labeling (Cy5) for each sample. These statistical analyses were performed using SAS statistical software (version 9.2, SAS Institute Inc., Cary, NC). Reported p-values are two-sided and those < 0.05 were considered to be statistically significant.

RESULTS

aCGH analysis

Regarding the quality of DNA extracted from FFPE tissue, the mean dsDNA, A260/280 and A260/A230 ratios were 0.30 (0.017 and 0.79), 1.84 (1.68 and 2.01) and 2.03 (0.92 and 2.28), respectively, with minimum and maximum values indicated in parentheses. After labeling with Cy5, the mean degree of labeling of DNA was 1.30 (0.27–2.02). Finally, the mean DLRSpread was 0.45 (0.20–1.05). Conversely, the mean dsDNA ratio, A260/A280 ratio, A260/A230 ratio, degree of labeling and DLRSpread for DNA extracted from fresh frozen tissue was 0.87 (0.77 and 0.96), 1.84 (1.83 and 1.84), 2.09 (2.02 and 2.16), 1.73 (1.72 and 1.74) and 0.18 (0.17 and 0.19), respectively (Supplementary Table 1).

For the 2 breast cancer cases for which matched fresh frozen tissue was available, the cytogenetic profiles of FFPE and matched fresh frozen samples showed good

concordance (Figure 1). For the 31 breast cancer samples, we compared the amplification status results for HER2 from aCGH and FISH to confirm the validity of using FFPE samples for aCGH. HER2 amplification was identified in 11 of 28 samples using FISH. Using aCGH, 8 of these 11 samples had log2 ratios > 0.25 for probe sets corresponding to the HER2 gene (A_14_P121276, A_14_P114826 and triplicate of A_16_P20643178), which met our criteria for a copy number gain. The sensitivity, specificity and overall accuracy for the HER2 gene were 81.8%, 94.1% and 89.3%, respectively (Supplementary Table 1).

Investigating predictive factors of DLRSpread

To determine factors predictive of DLRSpread, Spearman rank correlations between DLRSpread and the other factors were calculated. The dsDNA ratio and degree of labeling of DNA (Cy5) were significantly negatively correlated ($P < 0.0001$, $rs = -0.796$, -0.481, respectively) whereas storage time was significantly positively correlated ($P < 0.0001$, $rs = 0.551$) with DLRSpread. The A260/A280 ratio was positively correlated though not significantly so ($P = 0.0589$, $rs = 0.237$) (Table 1).

To elucidate the effect of each factor on DLRSpread, stepwise multiple linear regression analysis was conducted because each factor correlated with each other factor to some extent (Table 2). The adjusted R2 value was 0.53. Of the selected variables, the dsDNA ratio had the most significant effect on DLRSpread (regression coefficient -0.4798 ± 0.0914, $P < 0.0001$), while the degree of labeling of DNA (Cy5) had a mild yet significant effect on DLRSpread (regression coefficient -0.1796 ± 0.047, $P = 0.0003$) (Figure 2, Table 3).

The effect of WGA on aCGH analysis

For the 4 FFPE breast cancer samples for which WGA was performed, the dsDNA ratio for each sample was successfully increased. However, the degree of DNA labeling (Cy5) was slightly decreased and the DLRSpread was increased compared with the original DNA (Table 4, Figure 1). Thus, the increase in dsDNA ratio by whole-genome amplification did not improve the quality of aCGH analysis using DNA extracted from FFPE tissues.

DISCUSSION

The wealth of clinical samples and comprehensive clinical information, such as clinical outcomes, associated with FFPE archival tissue, make these samples an attractive

source for research. In fact, scarce clinical samples are often only available in FFPE as opposed to snap frozen form. However, utilizing FFPE tissues for molecular cytogenetic analysis is challenging because of DNA degradation caused by exposure to formaldehyde and acidic environments associated with the fixation and preservation process. Enzymatic fluorescence labeling is especially susceptible to the effects of DNA degradation. The ULS labeling system, which labels DNA non-enzymatically, is therefore an ideal method for aCGH analysis with degraded samples.

In this study, we conducted aCGH analysis on various types of cancer archival tissue samples using the ULS labeling system. Our aCGH results, which had a mean DLRSpread of 0.45 and an overall accuracy of 89.3% for the HER2 gene, were comparable to the results from former aCGH studies using FFPE archival tissue, though relatively lower in quality compared to that expected from DNA from fresh frozen tissue or peripheral blood lymphocytes(16, 17, 23). Remarkably, very old archival tissue, preserved for more than 20 years, could still generate meaningful data. Nevertheless, FFPE samples tend to generate poor aCGH results, which indicate the importance of identifying a predictive factor for the performance of FFPE DNA in aCGH analysis.

According to the manufacturer's instructions for ULS labeling, sample quality is only assessed by A260/A280 and A260/A230 ratios before labeling, thus DNA samples which show A260/A280 ratios of 1.8 to 2.0 and A260/A230 ratios > 2.0 are considered high-quality samples. The degree of labeling is used as a quality criterion for aCGH results with optimal ranges for Cy5 and Cy3 between 0.75% and 2.5%, and between 1.75% and 3.5% respectively, with a Cy3 minus Cy5 range between 1% and 2%. However, our investigation revealed that neither the A260/A280 nor the A260/A230 ratio could predict aCGH performance. The degree of labeling indeed predicted aCGH performance, but it cannot predict aCGH performance before the labeling step. Furthermore, stepwise multiple linear regression analysis revealed that the dsDNA ratio was the most significant predictive factor of aCGH performance. According to the manufacturer's instructions for the ULS labeling system, the DLRSpread for FFPE samples should be less than 0.4, which corresponds to a dsDNA ratio of 0.3 according to our results. Hence, extremely old samples can be used for aCGH analysis when using the ULS labeling system, as long as the dsDNA ratio is > 0.3.This finding is profoundly useful because we can predict aCGH performance before the labeling step. This enables the user to predict which archival samples will generate high-quality data before processing the samples, without wasting scarce specimen and financial resources.

From our results, we hypothesized that an increase in dsDNA ratio could improve

aCGH performance when using the ULS labeling system. We then employed the GenomePlex Whole Genome Amplification Kit, which generates a representative 500 fold amplification of genomic DNA using universal primers and optimized enzymes that decreases the background in the reaction(33), to enlarge the dsDNA ratio of FFPE samples and compared the aCGH results. Disappointingly, the increase in dsDNA ratio did not translate to an improvement in DLRSpread, indicating that the poor performance of degraded DNA on aCGH platforms could not be attributed to the dsDNA ratio, but that the dsDNA ratio was an indicator of the degree of DNA degradation in the sample.

Recently, next generation sequencing technologies have enabled novel findings with regards to the molecular architecture of human cancers (34-39). These technologies allow detection of not only DNA sequence changes, but also structural changes and copy number changes (40, 41) – and can be used in routine clinical practice in some cases (42). However, these technologies require high grade DNA extracted from fresh or frozen samples which is difficult to obtain, especially for rare patient cases. Furthermore, it remains costly to analyze and interpret huge data sets obtained through these technologies, although the related running costs are decreasing. On the other hand, degraded DNA extracted from FFPE samples can be used for conventional Sanger sequencing and aCGH analysis, which may be sufficient for some molecular genetics analyses. Furthermore, the cost of aCGH analysis is much lower than next generation sequencing technologies, for example, the aCGH platform used in this study, with about sixty thousand probes, cost only two hundred dollars per sample. Therefore, when considering copy number analyses of FFPE samples as well as cost effectiveness thereof, aCGH analysis remains an attractive even in the era of next generation sequencing.

In conclusion, the dsDNA ratio can be used to predict the performance of DNA from FFPE samples on the aCGH analysis when using the ULS labeling system. Using this quality metric, valuable FFPE archival tissue samples can be utilized for aCGH analysis without wasting scarce specimen and financial resources.

REFERENCES

1 Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. Nat Rev Mol Cell Biol. 2004 Mar;11(3):220-8.

2 Barrett MT, Scheffer A, Ben-Dor A, et al. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. Proc Natl Acad Sci U S A. 2004 Dec 21;101(51):17765-70.

3 Naylor TL, Greshock J, Wang Y, et al. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. Breast Cancer Res. 2005;7(6):R1186-98.

4 Reis-Filho JS, Simpson PT, Gale T, Lakhani SR. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. Pathol Res Pract. 2005;201(11):713-25.

5 Fridlyand J, Snijders AM, Ylstra B, et al. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer. 2006;6:96.

6 Climent J, Garcia JL, Mao JH, Arsuaga J, Perez-Losada J. Characterization of breast cancer by array comparative genomic hybridization. Biochem Cell Biol. 2007 Aug;85(4):497-508.

7 Loo LW, Ton C, Wang YW, et al. Differential patterns of allelic loss in estrogen receptor-positive infiltrating lobular and ductal breast cancer. Genes Chromosomes Cancer. 2008 Dec;47(12):1049-66.

8 Melchor L, Honrado E, Garcia MJ, et al. Distinct genomic aberration patterns are found in familial breast cancer associated with different immunohistochemical subtypes. Oncogene. 2008 May 15;27(22):3165-75.

9 Andre F, Job B, Dessen P, et al. Molecular characterization of breast cancer with high-resolution oligonucleotide comparative genomic hybridization array. Clin Cancer Res. 2009 Jan 15;15(2):441-51.

10 Hu X, Stern HM, Ge L, et al. Genetic alterations and oncogenic pathways associated with breast cancer subtypes. Mol Cancer Res. 2009 Apr;7(4):511-22.

11 Johnson NA, Hamoudi RA, Ichimura K, et al. Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells. Laboratory investigation; a journal of technical methods and pathology. 2006 Sep;86(9):968-78.

12 Gnanapragasam VJ. Unlocking the molecular archive: the emerging use of formalin-fixed paraffin-embedded tissue for biomarker research in urological cancer. BJU international. 2010 Jan;105(2):274-8.

13 Nessling M, Richter K, Schwaenen C, et al. Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. Cancer Res. 2005 Jan 15;65(2):439-47.
14 Thompson ER, Herbert SC, Forrest SM, Campbell IG. Whole genome SNP arrays using DNA derived from formalin-fixed, paraffin-embedded ovarian tumor tissue. Hum Mutat. 2005 Oct;26(4):384-9.
15 Paik S, Kim CY, Song YK, Kim WS. Technology insight: Application of molecular techniques to formalin-fixed paraffin-embedded tissues from breast cancer. Nat Clin Pract Oncol. 2005 May;2(5):246-54.
16 Paris PL, Sridharan S, Scheffer A, Tsalenko A, Bruhn L, Collins C. High resolution oligonucleotide CGH using DNA from archived prostate tissue. Prostate. 2007 Sep 15;67(13):1447-55.
17 Tuefferd M, De Bondt A, Van Den Wyngaert I, et al. Genome-wide copy number alterations detection in fresh frozen and matched FFPE samples using SNP 6.0 arrays. Genes Chromosomes Cancer. 2008 Jul 28.
18 Oosting J, Lips EH, van Eijk R, et al. High-resolution copy number analysis of paraffin-embedded archival tissue using SNP BeadArrays. Genome Res. 2007 Mar;17(3):368-76.
19 Gunn S, Yeh IT, Lytvak I, et al. Clinical array-based karyotyping of breast cancer with equivocal HER2 status resolves gene copy number and reveals chromosome 17 complexity. BMC Cancer. 2010;10:396.
20 Krijgsman O, Israeli D, Haan JC, et al. CGH arrays compared for DNA isolated from formalin-fixed, paraffin-embedded material. Genes Chromosomes Cancer. 2012 Apr;51(4):344-52.
21 Devries S, Nyante S, Korkola J, et al. Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. J Mol Diagn. 2005 Feb;7(1):65-71.
22 Lyons-Weiler M, Hagenkord J, Sciulli C, Dhir R, Monzon FA. Optimization of the Affymetrix GeneChip Mapping 10K 2.0 Assay for routine clinical use on formalin-fixed paraffin-embedded tissues. Diagn Mol Pathol. 2008 Mar;17(1):3-13.
23 Hostetter G, Kim SY, Savage S, et al. Random DNA fragmentation allows detection of single-copy, single-exon alterations of copy number by oligonucleotide array CGH in clinical FFPE samples. Nucleic Acids Res. 2010 Jan;38(2):e9.
24 Wang Y, Carlton VE, Karlin-Neumann G, et al. High quality copy number and genotype data from FFPE samples using Molecular Inversion Probe (MIP) microarrays. BMC medical genomics. 2009;2:8.

25 Raap AK, van der Burg MJ, Knijnenburg J, et al. Array comparative genomic hybridization with cyanin cis-platinum-labeled DNAs. Biotechniques. 2004 Jul;37(1):130-4.
26 Joosse SA, van Beers EH, Nederlof PM. Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer. 2007;7:43.
27 Knijnenburg J, van der Burg M, Tanke HJ, Szuhai K. Optimized amplification and fluorescent labeling of small cell samples for genomic array-CGH. Cytometry A. 2007 Aug;71(8):585-91.
28 van Beers EH, Joosse SA, Ligtenberg MJ, et al. A multiplex PCR predictor for aCGH success of FFPE samples. British journal of cancer. 2006 Jan 30;94(2):333-7.
29 Jacobs S, Thompson ER, Nannya Y, et al. Genome-wide, high-resolution detection of copy number, loss of heterozygosity, and genotypes from formalin-fixed, paraffin-embedded tumor tissue using microarrays. Cancer Res. 2007 Mar 15;67(6):2544-51.
30 Oikawa M, Yoshiura K, Kondo H, Miura S, Nagayasu T, Nakashima M. Significance of genomic instability in breast cancer in atomic bomb survivors: analysis of microarray-comparative genomic hybridization. Radiation oncology. 2011;6:168.
31 Arai J, Tsuchiya T, Oikawa M, et al. Clinical and molecular analysis of synchronous double lung cancers. Lung cancer. 2012 Aug;77(2):281-7.
32 Oikawa M, Nagayasu T, Yano H, et al. Intracystic papillary carcinoma of breast harbors significant genomic alteration compared with intracystic papilloma: genome-wide copy number and LOH analysis using high-density single-nucleotide polymorphism microarrays. The breast journal. 2011 Jul-Aug;17(4):427-30.
33 Little SE, Vuononvirta R, Reis-Filho JS, et al. Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA. Genomics. 2006 Feb;87(2):298-306.
34 Pleasance ED, Stephens PJ, O'Meara S, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature. Jan 14;463(7278):184-90.
35 Pleasance ED, Cheetham RK, Stephens PJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. Jan 14;463(7278):191-6.
36 Shah SP, Morin RD, Khattra J, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009 Oct 8;461(7265):809-13.
37 Juul N, Szallasi Z, Eklund AC, et al. Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials. Lancet Oncol. Feb 26.

38 Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. Apr 15;464(7291):999-1005.
39 Yachida S, Jones S, Bozic I, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature. Oct 28;467(7319):1114-7.
40 Campbell PJ, Stephens PJ, Pleasance ED, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nat Genet. 2008 Jun;40(6):722-9.
41 Stephens PJ, McBride DJ, Lin ML, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature. 2009 Dec 24;462(7276):1005-10.
42 Corless CL. Medicine. Personalized cancer diagnostics. Science. 2011 Dec 2;334(6060):1217-8.

TABLES

Table 1. Spearman rank correlation coefficients (r_s) between DLRSpread and other features.

Features	r_s
A260/A280 ratio	0.237 (0.0589)
A260/A230 ratio	-0.087 (0.4953)
dsDNA ratio	-0.796 (<0.0001)
Storage time (years)	0.551 (<0.0001)
Degree of labeling (Cy5)	-0.481 (<0.0001)

P-values are shown in parentheses. A260/A280 ratio: The ratio of the absorbance at 260 and 280nm. A260/A230 ratio: The ratio of the absorbance at 260 and 230nm. dsDNA ratio: The ratio of the amount of double-strand DNA and total DNA in tumor DNA.

Table 2. Spearman rank correlation coefficients between each feature.

P-value Rho	A260/A280 ratio	A260/A230 ratio	dsDNA ratio	Storage time	Degree of Labeling
A260/A280 ratio					
A260/A230 ratio	0.384 (0.0017)				
dsDNA ratio	-0.389 (0.0015)	0.059 (0.6418)			
Storage time	-0.216 (0.0867)	-0.005 (0.9691)	-0.419 (0.0006)		
Degree of Labeling	-0.334 (0.0070)	-0.058 (0.6514)	0.492 (<0.0001)	-0.365 (0.0030)	

P-values are shown in parentheses. A260/A280 ratio: The ratio of the absorbance at 260 and 280nm. A260/A230 ratio: The ratio of the absorbance at 260 and 230nm. dsDNA ratio: The ratio of double-strand to total DNA.

Table 3. Partial regression coefficients of each feature in a step-wise multiple regression model showing the association with DLRSpread.

Features	Regression coefficient (SE)	P-value
dsDNA ratio	-0.4798 (0.0914)	<0.0001
Degree of labeling (Cy5)	-0.1796 (0.0470)	0.0003

SE: Standard error. dsDNA ratio: The ratio of double-strand to total DNA.

Table 4. The effect of WGA on aCGH analysis.

Case	Sample type	dsDNA ratio	Degree of labeling (Cy5)	DLRSpread
Case 30	FF	0.96	1.72	0.19
	FFPE	0.70	1.49	0.21
	WGA	0.75	1.30	0.99
Case 31	FF	0.77	1.74	0.17
	FFPE	0.67	1.56	0.20
	WGA	0.72	1.30	0.90
Case 11	FFPE	0.047	0.27	1.05
	WGA	0.20	0.31	1.07
Case 12	FFPE	0.035	0.28	1.04
	WGA	0.18	0.15	1.17

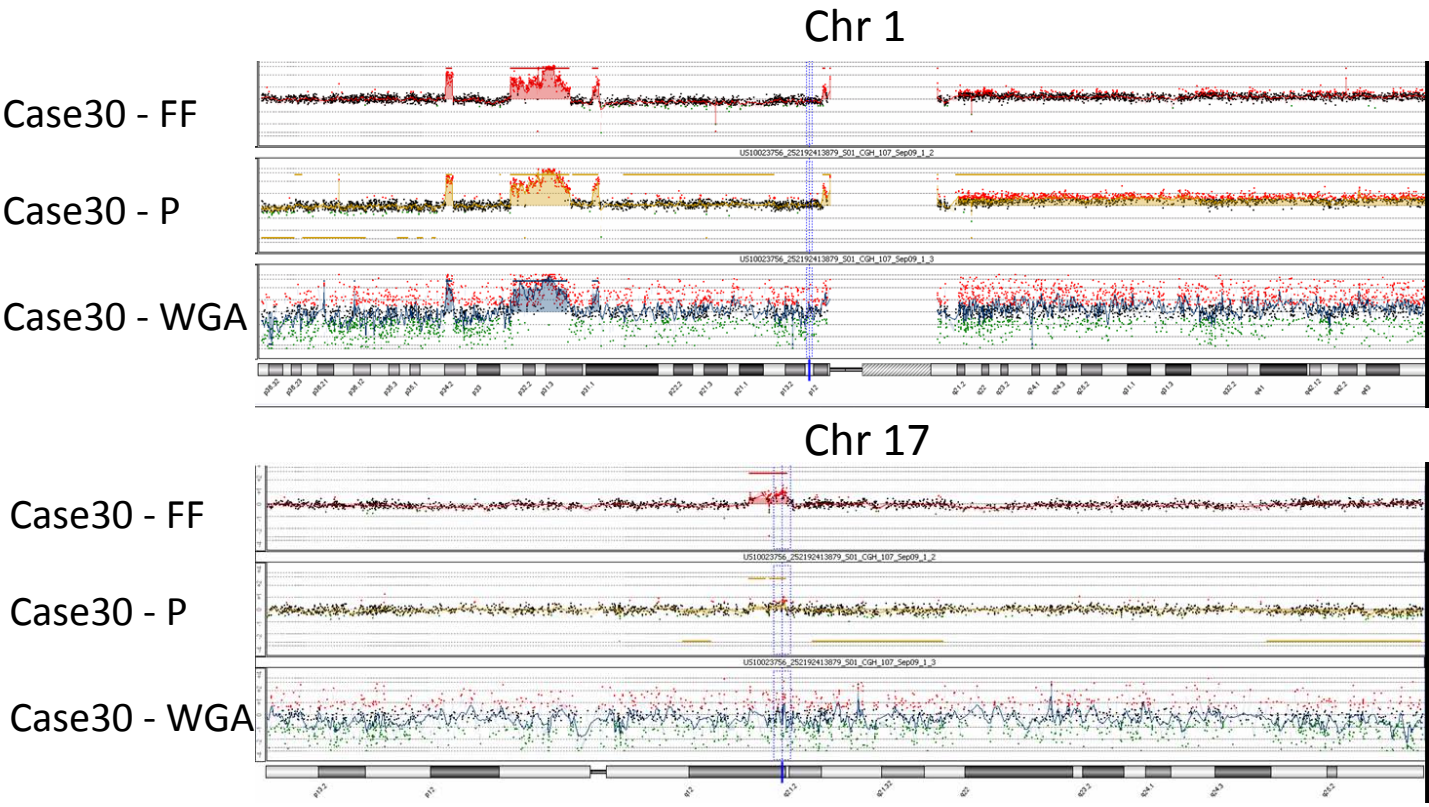
Supplementary Table 1. Summary of Clinicopathological Factors and aCGH Analysis.

ID	origine	NanoDrop (μg/mL)	Total DNA amount (μg)	260/280	260/230	Qubit (μg/mL)	dsDNA Ratio	Storage Time (years)	Dgree of Labeling of Cy5	DLR Spread	HER2 amplification (FISH)	HER2 amplification (aCGH)
1	breast	176.6	9.713	1.85	2.17	53.6	0.30	15	1.59	0.44	Present	Present
2	breast	56.4	3.102	1.68	2.14	16.4	0.29	19	1.67	0.49	Present	Absent
3	breast	42.8	2.354	1.7	2.01	13	0.30	34	1.87	0.68	Present	Present
4	breast	170.9	9.3995	1.83	2.21	61.1	0.36	12	1.42	0.51	Absent	Absent
5	breast	199	10.945	1.83	2.26	104	0.52	22	1.61	0.32	Absent	Absent
6	breast	51.5	2.8325	1.77	2.01	6.33	0.12	31	1.75	0.76	Present	Present
7	breast	690.1	37.9555	1.89	2.05	100	0.14	31	0.74	0.64	Absent	Absent
8	breast	2114.8	116.314	1.96	2.17	176	0.08	21	0.45	0.76	Present	Present
9	breast	471.7	25.9435	1.85	2.21	125	0.26	19	1.26	0.36	Absent	Absent
10	breast	606.3	33.3465	1.86	2.21	107	0.18	16	0.85	0.56	Absent	Absent
11	breast	691.6	38.038	1.96	2.14	32.3	0.05	33	0.27	1.05	Absent	Absent
12	breast	1511.5	83.1325	1.98	2.23	50.9	0.03	31	0.28	1.04	Absent	Absent
13	breast	184.1	10.1255	1.9	2.02	27	0.15	30	0.93	0.58	Absent	Absent
14	breast	176.2	9.691	1.91	2.22	38.7	0.22	30	0.88	0.50	Absent	Absent
15	breast	82.5	4.5375	1.79	2.14	39.9	0.48	15	1.57	0.30	Present	Present
16	breast	1927.2	105.996	1.88	2.19	33	0.02	15	0.6	0.52	Absent	Absent
17	breast	592.1	32.5655	1.94	2.28	163	0.28	14	1.24	0.38	Absent	Absent
18	breast	246.9	13.5795	1.86	2.12	25.8	0.10	19	0.68	0.80	Present	Present
19	breast	279.7	15.3835	1.95	2.11	44.4	0.16	27	0.83	0.58	Absent	Absent
20	breast	237.7	13.0735	1.94	2.16	46.5	0.20	17	0.99	0.52	Present	Absent
21	breast	670.2	36.861	2.01	2.08	103	0.15	15	0.78	0.75	Absent	Absent
22	breast	114.7	6.3085	1.7	2.2	30.6	0.27	23	1.24	0.40	NA	Present
23	breast	83.9	4.6145	1.68	1.77	9.35	0.11	28	1.01	0.48	Present	Present
24	breast	67.8	3.729	1.79	2.13	31	0.46	29	1.42	0.26	NA	Absent
25	breast	298.8	16.434	1.78	2.21	104	0.35	28	1.54	0.39	NA	Present
26	breast	174.3	9.5865	1.76	2.23	53.8	0.31	28	1.05	0.36	Absent	Absent
27	breast	94.8	5.214	1.7	2.18	24.5	0.26	43	1.34	0.52	Absent	Present
28	breast	164.1	9.0255	1.8	2.17	61.4	0.37	28	1.31	0.32	Absent	Absent
29	breast	119.1	6.5505	1.75	1.86	12.4	0.10	37	1.07	0.39	Absent	Absent
30	breast	117	6.435	1.83	2.16	81.4	0.70	2	1.49	0.21	Present	Present
31	breast	119.2	6.556	1.85	2.25	79.3	0.67	2	1.56	0.20	Present	Present
32	lung	253	13.915	1.82	2.16	107	0.42	10	1.76	0.29	NA	NA
33	lung	94.2	5.181	1.83	2.24	31.5	0.33	10	1.67	0.30	NA	NA
34	lung	155.7	8.5635	1.89	2.26	37.3	0.24	7	1.61	0.44	NA	NA
35	lung	53.8	2.959	1.82	2.15	17.9	0.33	7	1.82	0.40	NA	NA
36	lung	100.9	5.5495	1.9	2.19	28.2	0.28	2	1.93	0.32	NA	NA
37	lung	221.9	12.2045	1.92	2.23	77.5	0.35	2	1.64	0.27	NA	NA
38	lung	71.7	3.9435	1.96	2.15	3.74	0.05	7	1.98	0.94	NA	NA
39	lung	47.3	2.6015	1.89	1.95	8.01	0.17	7	2.02	0.47	NA	NA
40	lung	123.9	6.8145	1.79	1.71	38.9	0.31	9	1.23	0.39	NA	NA

41	lung	76.6	4.213	1.78	1.63	21.7	0.28	9	1.42	0.49	NA	NA
42	lung	360.6	19.833	1.87	1.96	132	0.37	7	1.5	0.23	NA	NA
43	lung	330.9	18.1995	1.9	1.69	75.1	0.23	7	1.35	0.31	NA	NA
44	lung	272.9	15.0095	1.9	2.23	83.6	0.31	7	1.55	0.33	NA	NA
45	lung	114.4	6.292	1.72	1.09	30.9	0.27	7	1.75	0.22	NA	NA
46	lung	375.7	20.6635	1.95	2.15	103	0.27	2	1.54	0.38	NA	NA
47	lung	80.5	4.4275	1.79	1.37	26.7	0.33	2	1.78	0.35	NA	NA
48	lung	305.1	16.7805	1.85	1.8	17.9	0.06	8	0.58	1.01	NA	NA
49	lung	207.2	11.396	1.79	1.96	39.7	0.19	8	1.11	0.51	NA	NA
50	lung	1654.9	91.0195	1.92	2.27	389	0.24	7	0.78	0.34	NA	NA
51	lung	230.3	12.6665	1.91	2.05	32.1	0.14	7	0.84	0.51	NA	NA
52	lung	46.9	2.5795	1.76	0.92	15.9	0.34	7	1	0.47	NA	NA
53	lung	57.2	3.146	1.75	1.23	22.2	0.39	7	1.53	0.39	NA	NA
54	lung	630.5	34.6775	1.76	2.17	181	0.29	11	1.59	0.31	NA	NA
55	lung	101	5.555	1.76	2	32	0.32	11	1.67	0.34	NA	NA
56	thyroid	290.8	15.994	2	2.27	72.4	0.25	1	0.95	0.46	NA	NA
57	thyroid	196.7	10.8185	1.86	1.36	116	0.59	3	1.56	0.21	NA	NA
58	thyroid	279.8	15.389	1.84	2.24	196	0.70	4	1.47	0.24	NA	NA
59	thyroid	164.4	9.042	1.7	1.8	129	0.78	14	1.24	0.35	NA	NA
60	thyroid	110.1	6.0555	1.75	1.97	50.5	0.46	6	1.59	0.24	NA	NA
61	thyroid	167.3	9.2015	1.79	1.86	78.6	0.47	7	1.53	0.21	NA	NA
62	thyroid	286.8	15.774	1.99	2.22	83.2	0.29	1	1.37	0.33	NA	NA
63	thyroid	158.5	8.7175	1.86	2.08	125	0.79	4	1.44	0.31	NA	NA

Figure 1.

[A]


Chr 1
Case30 - FF
Case30 - P
Case30 - WGA
Chr 17
Case30 - FF
Case30 - P
Case30 - WGA

[B]

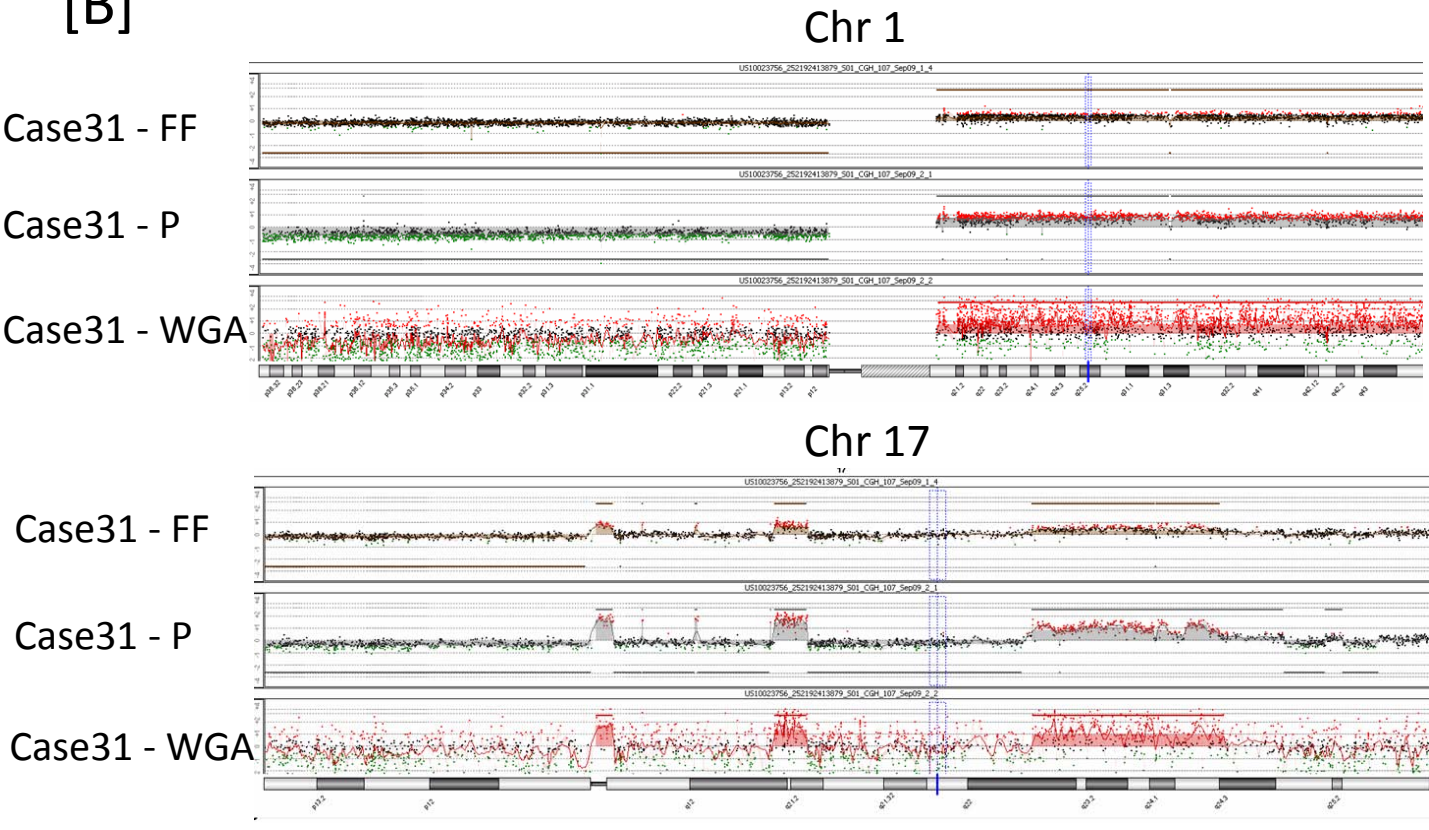
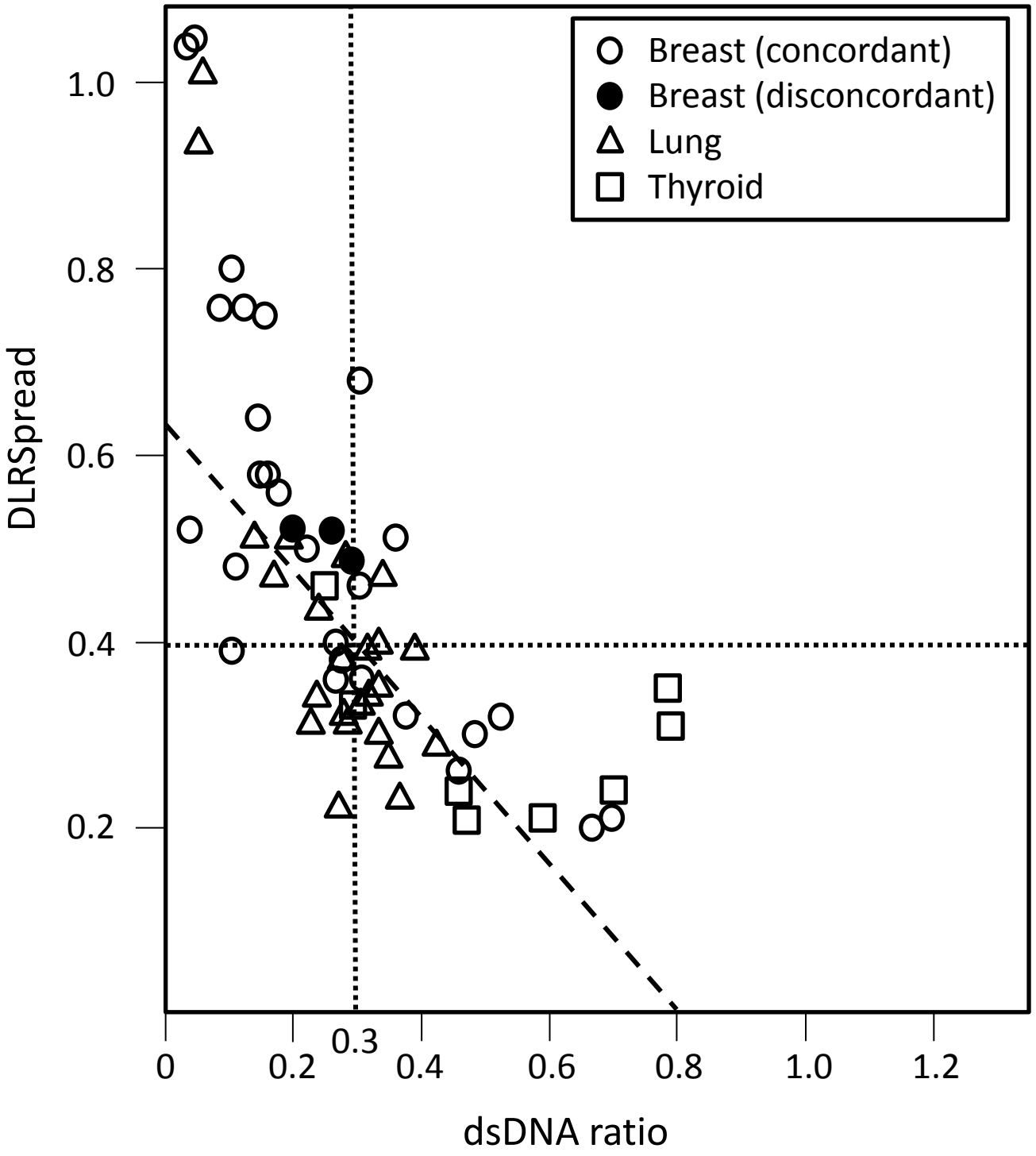

Chr 1
Case31 - FF
Case31 - P
Case31 - WGA
Chr 17
Case31 - FF
Case31 - P
Case31 - WGA

Figure 2.


Breast (concordant)
Breast (disconcordant)
Lung
Thyroid
DLRSpread
1.0
0.8
0.6
0.4
0.2
0
0.2
0.3
0.4
0.6
0.8
1.0
1.2
dsDNA ratio