

# Genomic Instability: A Stronger Prognostic Marker Than Proliferation for Early Stage Luminal Breast Carcinomas

Anne Vincent-Salomon<sup>1,2\*</sup>, Vanessa Benhamo<sup>2,3</sup>, Eléonore Gravier<sup>3,4,5,6</sup>, Guillem Rigai<sup>3,7</sup>, Nadège Gruel<sup>2,3</sup>, Stéphane Robin<sup>7</sup>, Yann de Rycke<sup>4</sup>, Odette Mariani<sup>1</sup>, Gaëlle Pierron<sup>1</sup>, David Gentien<sup>3</sup>, Fabien Rey<sup>8</sup>, Paul Cottu<sup>9</sup>, Alain Fourquet<sup>10</sup>, Roman Rouzier<sup>8</sup>, Xavier Sastre-Garau<sup>1</sup>, Olivier Delattre<sup>2</sup>

1 Department of Tumor Biology, Institut Curie, Paris, France, 2 INSERM U830, Institut Curie, Paris, France, 3 Department of Translational Research, Institut Curie, Paris, France, 4 Department of Biostatistics, Institut Curie, Paris, France, 5 INSERMU900, Institut Curie, Paris, France, 6 Ecole des Mines de Paris, Paris, France, 7 AgroParisTech, UMR518, Paris, France, 8 Department of Surgery, Institut Curie, Paris, France, 9 Department of Medical Oncology, Institut Curie, Paris, France, 10 Department of Radiation Oncology, Institut Curie, Paris, France

## Abstract

**Background:** The accurate prognosis definition to tailor treatment for early luminal invasive breast carcinoma patients remains challenging.

**Materials and Methods:** Two hundred fourteen early luminal breast carcinomas were genotyped with single nucleotide polymorphisms (SNPs) array to determine the number of chromosomal breakpoints as a marker of genomic instability. Proliferation was assessed by KI67 (immunohistochemistry) and genomic grade index (transcriptomic analysis). IHC3 (IHC4 score for HER2 negative tumors) was also determined.

**Results:** In the training set (109 cases), the optimal cut-off was 34 breakpoints with a specificity of 0.94 and a sensitivity of 0.57 (Area under the curve (AUC): 0.81[0.71; 0.91]). In the validation set (105 cases), the outcome of patients with > 34 breakpoints (11 events / 22 patients) was poorer (logrank test  $p < 0.001$ ; Relative Risk (RR): 3.7 [1.73; 7.92]), than that of patients with < 34 breakpoints (19 events / 83 patients). Whereas genomic grade and KI67 had a significant prognostic value in univariate analysis in contrast to IHC3 that failed to have a statistical significant prognostic value in this series, the number of breakpoints remained the only significant parameter predictive of outcome (RR: 3.47, Confidence Interval (CI) [1.29; 9.31],  $p = 0.014$ ) in multivariate analysis.

**Conclusion:** Genomic instability, defined herein as a high number of chromosomal breakpoints, in early stage luminal breast carcinoma is a stronger prognostic marker than proliferation.

**Citation:** Vincent-Salomon A, Benhamo V, Gravier E, Rigai G, Gruel N, et al. (2013) Genomic Instability: A Stronger Prognostic Marker Than Proliferation for Early Stage Luminal Breast Carcinomas. PLoS ONE 8(10): e76496. doi:10.1371/journal.pone.0076496

**Editor:** William B. Coleman, University of North Carolina School of Medicine, United States of America

**Received:** February 25, 2013; **Accepted:** August 27, 2013; **Published:** October 15, 2013

**Copyright:** © 2013 Vincent-Salomon et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors have no support or funding to report.

**Competing interests:** The authors declare that no competing interests exist.

\* E-mail: anne.salomon@curie.net

## Introduction

A major challenge of breast cancer treatment is to accurately identify patients in whom adjuvant therapy can be safely avoided [1]. Over the past decade, five molecular groups of invasive breast carcinoma have been identified, each associated with a different outcome [2]. Triple-negative carcinomas are recognized as the group associated with the poorest prognosis and efficiently treated with adjuvant chemotherapy [3,4]. Identification of high-risk patients in this group by means of the classical clinicopathological parameters therefore remains difficult. As a result of mammography screening, the majority of breast cancers are now diagnosed at an early luminal stage with no axillary lymph node metastasis,

and HER2-negative. The benefit of medical adjuvant treatment in these patients also remains controversial [1]. While luminal carcinoma patients display various outcomes according to their tumor proliferation rate and clinical stage, they harbor a constellation of driver gene mutations and chromosomal alterations [5–9].

Proliferation assessed either by mitotic index according to the Ellis and Elston recommendations [10] or KI67 immunostaining improves the prognostic definition of luminal breast carcinoma patients [11]. However, molecular methods such as genomic grade index (Mapquant®)[12] have been proposed as a more accurate tool to evaluate proliferation. This technique has been developed to classify grade 2 tumors into good and poor outcome groups. IHC4 score, an algorithm

inferred from ER, PR, KI67 and HER2 immunostainings has been recently proposed as an interesting alternative tool to refine breast cancers prognostic. Its prognostic value was found to be equivalent to that of the OncotypeDX® for a lower cost [13].

Pangenomic analyses have provided substantial data describing breast carcinoma patterns of genomic alteration [9,14,15]. These analyses, consolidated by recent exome-sequencing results, have also shown that genomic alterations are different according to the molecular or histological types of breast carcinomas [5–7].

The prognostic value of genomic instability for breast cancer has been evaluated either by transcriptomic signature enriched with genes involved in chromosomal integrity maintenance [16–19] or inferred evaluation of ploidy based on fluorescent in situ hybridization [20].

We and others have proposed that the level of genomic complexity within breast carcinomas could contribute to define patient outcome [16,20–23]. In our previous work, we determined that the number of breakpoints had a prognostic power to identify high-risk T1-T2 N0 invasive ductal carcinomas of the breast. We used an in-house array-comparative genomic hybridization (CGH) tool to define a CGH classifier. This classifier was based both on the number of chromosomal breakpoints together with three regions (2p22.2, 3p23 and 8q21-24) to discriminate tumors associated with good and poor outcome. The sensitivity of this classifier was 66%, with a specificity of 84% and an accuracy of 78% [21].

The objective of the present study was to validate, in an independent series, the prognostic value of our previously established genomic signature. In this study, we used a different tool, the SNPs array (SNP 6.0 array, Affymetrix®), as this array has a higher resolution than the array CGH used in our previous study [21] and compared it to proliferation prognostic value. We focused our analysis on early luminal HER2-negative node-negative breast cancers, i.e. in the population of patients in whom prognosis assessment remains challenging, in whom treatment must be adapted to tumor biology and decreased as far as possible to minimize the side effects of chemotherapy.

## Materials and Methods

### Patients and tumors

All experiments were performed retrospectively and in accordance with the French Bioethics Law 2004-800, the French National Institute of Cancer (INCa) Ethics Charter and after approval by the Institut Curie review board and ethics committee (Comité de Pilotage of the Groupe Sein). In the French legal context, our institutional review board waived the need for written informed consent from the participants. Moreover, women were informed of the research use of their tissues and did not declare any opposition for such researches. Data were analyzed anonymously.

This analysis was confined to ER-positive, node-negative early (T1-T2 tumor smaller than 30 mm), HER2-negative breast cancers because they now represent the majority of breast carcinomas at diagnosis as a result of systematic screening

procedures in western countries [24]. Tumours from the two sets (training and validation) were selected among the 7469 tumor samples from patients treated between 1989 and 1999 with a conservative surgery followed by radiation therapy at the Institut Curie for T1T2 breast carcinoma. The selection criteria for the two sets were small size (

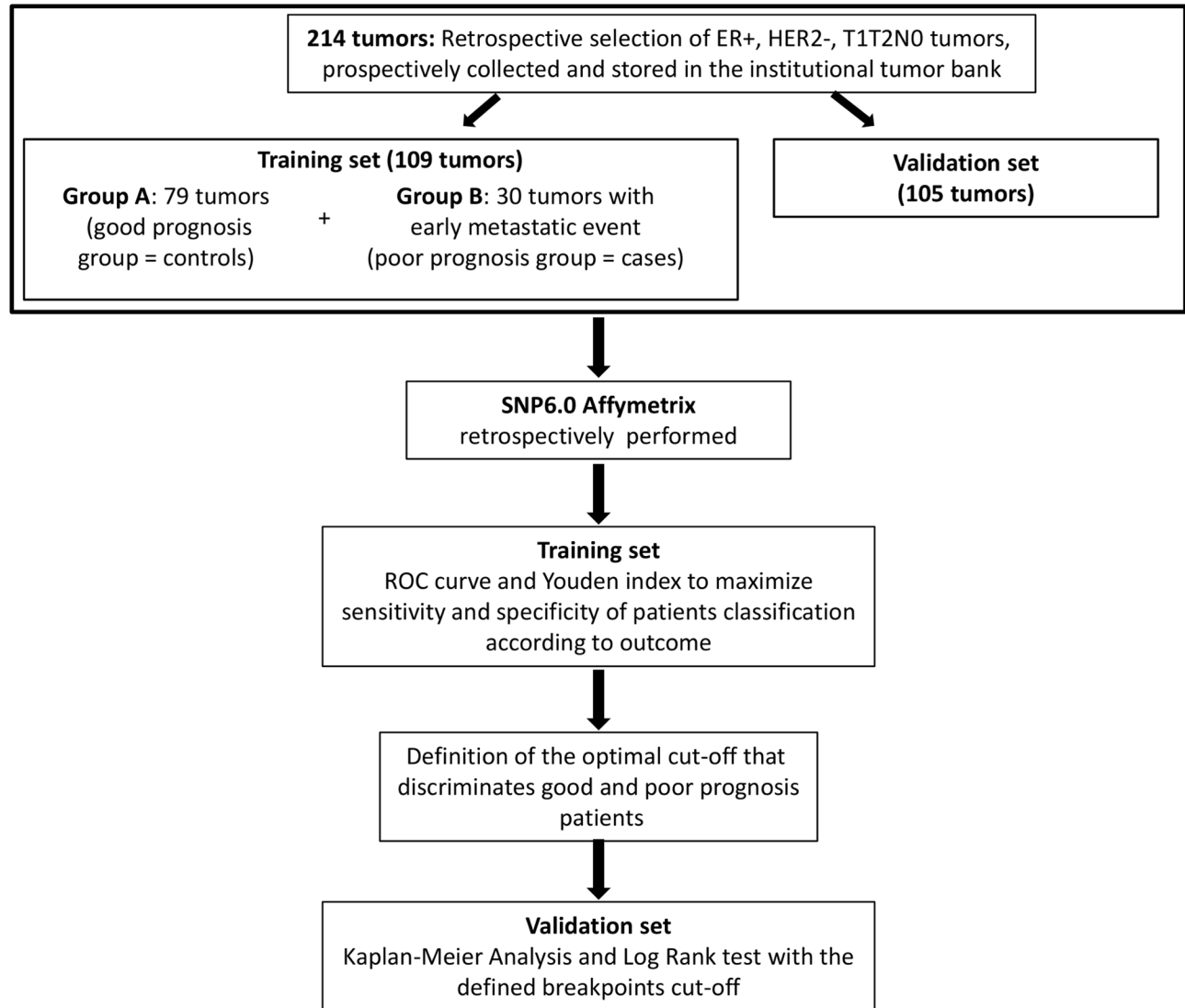
< 3cm), node-negative, invasive ductal breast carcinomas ER + HER2-, node negative. Patients older than 75 years, with previous personal history of cancer, with familial history of breast cancer, with multifocal or bilateral tumours, or with initial metastatic disease were excluded from the study. The final selection was based on the availability and quality of tumor samples in our tumor bank. These samples were collected from fresh surgical specimens by a pathologist within one hour after surgery, flash-frozen in liquid nitrogen and stored at -80°C. The Figure 1 represented the workflow for retrospective tumor selection among available frozen samples collected prospectively between 1989 and 1999. Training set consisted of a case-control group composed of 109 tumors. The median follow-up for the patients of this good prognosis group was 10.9 years, (95% CI: 10.4-11.5). Seventy-nine patients were metastasis-free at five years (A good prognosis group, controls). Thirty patients developed metastasis before four years (Bpoor prognosis group, cases) with a median time to metastasis of 2.3 years (interquartile range: 1.5-3.4). None of the patients had received neoadjuvant or a  
 djuvant chemotherapy, but six patients had received adjuvant hormone therapy.

The validation set was composed of 105 tumors samples from 105 patients. Median follow-up was 10.5 years, (95% CI: 10.1-10.9). Tumors were invasive ductal carcinomas smaller than 30 mm, ER-positive, HER2-negative, and node-negative. None of the 105 patients had received neoadjuvant or adjuvant chemotherapy and ten patients had received adjuvant hormone therapy. Thirty out of the 105 patients developed loco-regional or contralateral relapses or experienced a metastatic event (8 out of the 30 patients).

All tumors were retrospectively reviewed by pathologists experienced in breast pathology (XSG and AVS) for the purpose of this study. In particular, histoprognostic grade was assessed retrospectively according to the recommendations of Elston and Ellis [25].

### Immunostaining for Estrogen-receptor (ER), progesterone-receptor (PR), HER2 and KI67 determination

ER, progesterone receptor, HER2 receptor and KI67 status were assessed by immunohistochemistry on representative formalin-fixed tumor blocks, according to previously published protocols [26] (see Data S1 for antibody details). The semiquantitative KI67 assessment was performed as previously published [27] and as recommended [28]. A cut-off of 14% was used to define tumors with a high KI67 score (according to St Gallen recommendations [29] and cut-off for molecular classification [30]. Internal (normal glands surrounding the carcinoma) and external controls (for ER, PR and HER2: tissue-microarrays composed of tumors with known ER, PR status, and known numbers of *HER2* gene copies



**Figure 1. Flowchart of patients and tumor's selection and design of the study.**

doi: 10.1371/journal.pone.0076496.g001

together with normal mammary tissue; for Ki67: normal lymph node with germinal centers as positive controls) were included in all immunostaining experiments.

Following Cuzick's recommendations, the IHC4 was determined and adapted for HER2 negative tumors (named "IHC3") and, as Ki67 was assessed manually, in the formulae, the Ki67 value was rescaled [13].

#### DNA and RNA analyses

All tumor samples contained more than 50% of cancer cells, as assessed by Hematein & Eosin staining of histologic sections of the frozen samples used for nucleic acid extraction. DNA and RNA were extracted according to a previously described phenol-chloroform protocol [27,31].

#### SNP6.0 analyses for tumor genomic complexity evaluation

The Affymetrix GeneChip Human Mapping 6.0<sup>®</sup> was normalized using Affymetrix Genotyping Console<sup>®</sup> (version GTC 3.0.1) (.cel files and data available under accession number GSE48064 in GEO database). The signal from SNP probes was then segmented using Colibri [32]. A first segmentation round was performed to identify and discard outlier probes giving rise to singleton segments. The number of breakpoints was estimated according to the Zhang and Siegmund's BIC-based model selection criterion (Data S2) [33].

The optimal segmentation in terms of least-square fit was recovered via dynamic programming. The pruned version proposed in [[32]] can be used to achieve this optimization step with almost linear complexity. We used the implementation available in the cghseg package [34] via the function

**Table 1.** Clinico-biological characteristics of patients and tumors.

Characteristics	Training set (n=109)	Validation set (n=105)
<b>Age (years)</b>		
Means (min-max)	55.6 (36-75)	54.2 (40-69)
>50	74 (68%, 95%CI : 59%-77%)	68 (65%, 95%CI : 56%-74%)
≤50	35 (32%, 95%CI : 23%-41%)	37 (35%, 95%CI : 26%-44%)
<b>Size</b>		
T1	77 (71%, 95%CI :62%-79%)	77 (73%, 95%CI : 65%-82%)
T2	32 (29%, 95%CI : 20%-38%)	28 (27%, 95%CI : 19%-35%)
<b>Elston-Ellis grade</b>		
I	63 (58%, 95% CI:49%-67%)	48 (46%, 95% CI : 36%-55%)
II	35 (32%, 95%CI:23%-41%)	41 (39%, 95%CI : 30%-48%)
III	9 (8%, 95%CI :3%-13%)	16 (15, 95%CI : 8%-22%)
Missing values	2 (2%)	0
<b>Vascular invasion</b>		
Yes	10 (9%, 95%CI:4%-14%)	20 (19%, 95%CI :12%-26%)
<b>ER status (IHC)</b>		
Positive	109 (100%)	105 (100%)
<b>PR status (IHC)</b>		
Positive	83 (76%, 95%CI:68%-84%)	80 (76%, 95%CI:68%-84%)
<b>KI67</b>		
≤ 14%	nd	47 (45%, 95%CI:36%-55%)
> 14%	nd	58 (55%, 95%CI:46%-64%)
<b>Genomic grade</b>		
GGI 1	nd	56 (60%, 95%CI:50%-69%)
GGI 3	nd	38 (40%, 95%CI:30%-49%)
Undetermined	nd	11 (10.5%)

Min: minimum; max: maximum. GG: genomic grade; ER: estrogen receptors; PR: progesterone receptor; 95%CI: Confidence Interval.

doi: 10.1371/journal.pone.0076496.t001

segmeanCO. The training set was used to assess the link between the number of breakpoints and the tumor group (A: late onset versus B: early onset metastatic or relapse event). The breakpoint number (assessing the number of chromosomal segmental alterations) cut-off able to accurately identify B tumors (poor prognosis tumors) was determined using the Youden index. This index is defined as the threshold that maximizes the sum of sensitivity and specificity. The number of breakpoints was used as a marker of genomic complexity.

### Expression analyses and determination of genomic grade index

After RNA quality control and according to a previously published protocol [27], 94 of the 105 validation cases meeting our selection criteria for node status, tumor size, ER and HER2 status, were hybridized onto U133 plus 2.0 Affymetrix® chips. Transcriptomic data were normalized using RMA (under R software 2.13.2 version). The genomic grade index (GGI), a continuous variable, was calculated using the MapQuant DxM protocol, according to Sotiriou et al [12] and as previously described [27], and was defined as  $GGI = \text{scale} [\text{sum} (\text{Probe Sets up in Grade 3 tumors}-112 \text{ probe sets}) - \text{sum} (\text{Probe Sets up in Grade 1 tumors}-16 \text{ probe sets}) - \text{offset}]$ . Scale and offset are transformation parameters to standardize the genomic

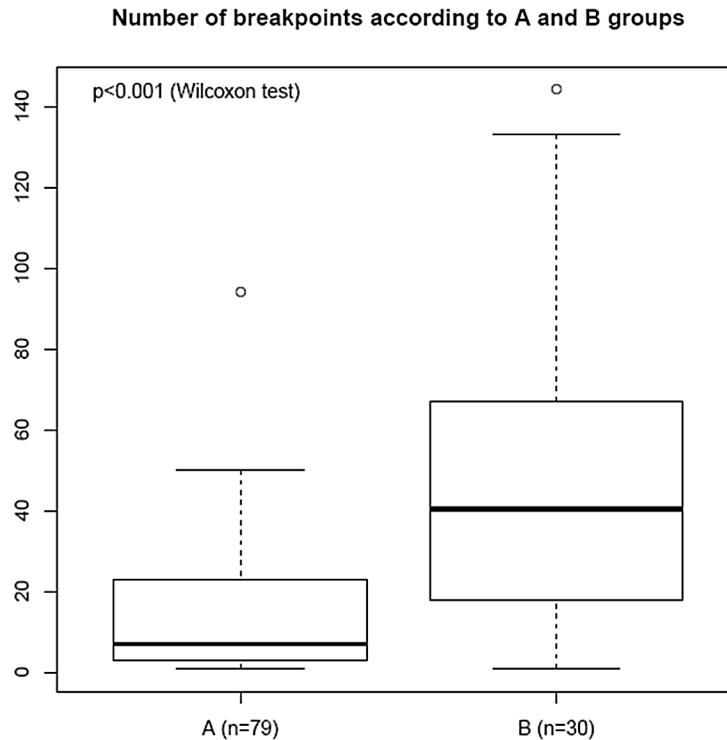
grade index values. The MapQuant® GGI was then standardized by setting the scale and offset parameters so that the mean GGI of histologic Grade 1 tumors was -1 and that of histologic grade 3 tumors was +1. The cut-off was set at 0. Based on the value of the GGI, a genomic grade (Genomic Grade 1 or Genomic Grade 3) was then attributed to each tumor sample.

## Results

### Patients

Detailed patient and tumor characteristics are displayed in Table 1. Training and validation sets comprised a large majority of T1, grade 1 and 2 tumors.

In addition to the Elston and Ellis histoprognostic grade, proliferation was also evaluated by KI67, IHC3 and genomic grade index for 94 of the 105 tumors in the validation set. As shown in Table 1, 45% and 60% of these tumors presented a KI67 index less than 14% of positive cells and a genomic grade index of 1, respectively.



**Figure 2. Boxplot of the number of breakpoints within groups A and B from the training set.** The black line represents the breakpoints median values (median = 7 in A group; median = 40.5 in B group). The bottom and top of the boxes represent the 25th and 75th percentile respectively, whereas the box represents the interquartile range. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. The single point at the top of each boxplot represents the maximum number of breakpoints within each group A and B respectively.

doi: 10.1371/journal.pone.0076496.g002

**Table 2. Number of tumors with less or more than 34 breakpoints (Threshold) within A and B groups, specificity and sensibility of tumor classification in the training set.**

Groups	Number of breakpoints		Specificity	Sensitivity
	≤ 34 (%)	> 34 (%)		
A	74 (94)	5 (6)	0.94	
B	13 (43)	17 (57)		0.57

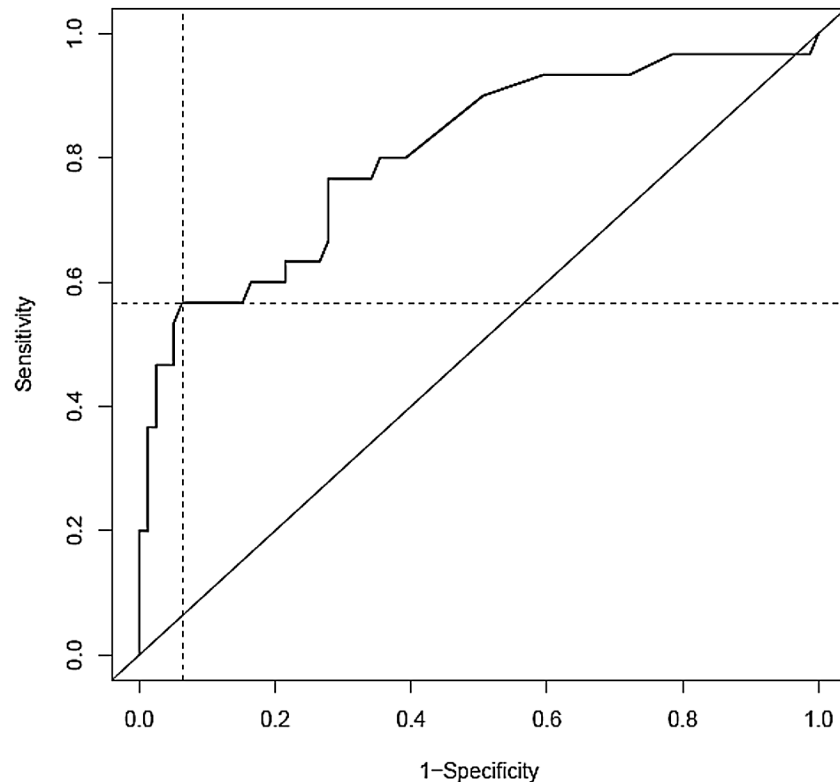
doi: 10.1371/journal.pone.0076496.t002

### Assessment of genomic complexity with the number of chromosomal breakpoints and definition of a classifier within the training set

According to the segmentation protocol described above, the numbers of breakpoints was determined in group A and group B tumors of the training set and were compared. As shown in Figure 2, the median number of breakpoints in group A tumors was significantly lower (median value=7) than the median number of breakpoints in group B (median value = 40.5)(95% CI: -39.00 ; -14.00,  $p < 0.001$ , Wilcoxon test).

An optimal cut-off of thirty-four breakpoints was defined to assign each patient to a risk class, as the number of breakpoints ensuring maximum sum of sensitivity (0.57 95% CI: 39%-74%) and specificity (0.94 95% CI: 88%-99%), was

34. A tumor with more than 34 breakpoints was classified in the high-risk complexity group, otherwise it was assigned to the low-risk complexity group. In good prognosis group (no metastatic event within five years), 74 tumors (74/79, 94% 95% CI: 88%-99%) were considered to be low-risk complexity tumors and 5 were considered to be high-risk complexity tumors (5/79, 6% 95% CI: 0%-12%) whereas in poor-prognosis group (metastatic relapse in less than four years), 13 tumors were considered to be low-risk complexity tumors (13/30, 43% 95% CI: 26%-61%) and 17 tumors (17/30, 57% 95% CI: 39%-74%) were considered to be high-risk complexity tumors (Table 2).



**Figure 3. Receiving Operating Curve (ROC) curve for the number of breakpoints threshold determination.** Intersection of dotted-lines = Youden index.

doi: 10.1371/journal.pone.0076496.g003

Analysis of the complexity classifier produced a Receiving Operating Characteristics (ROC) curve with a significant area of 0.81 (95% CI: 0.71-0.91) (Figure 3).

### Number of breakpoints is linked with histoprognostic grade, genomic grade and KI67 labeling

In the validation set of this series of cases, a link between grade ( $p < 0.001$ ; Kruskal-Wallis test), genomic grade ( $p < 0.001$ , Wilcoxon test), KI67 ( $p < 0.001$ , Wilcoxon test) and the number of breakpoints was observed.

The median number of breakpoints in the group of tumors with a KI67 smaller than 14% and greater than or equal to 14% were 5 and 17, respectively, (95% CI: -19-4) (Figure S1).

The Spearman correlation coefficient between genomic grade index and number of breakpoints was 0.54 ( $p < 0.001$ ) (Figure S2).

### Link between the number of breakpoints according to the cut-off chosen and disease-free interval in the validation set

To test the robustness of this cut-off, we determined disease-free interval in the validation set and tested the correlation with the number of breakpoints in univariate and multivariate analyses taking into account clinicopathological and

proliferation markers such as grade, KI67 GGI and IHC3 (Table 3).

Whereas IHC3 score was not found statistically significant in this series of tumors, the three proliferation markers, histoprognostic grade, KI67 and genomic grade index together with the number of breakpoints were significantly associated with disease free interval at a significance level of 10% in univariate Cox analysis (Table 3). Interestingly, the number of breakpoints remained the only significant parameter when tested with grade, KI67 and genomic grade index in multivariate Cox analysis (relative risk of 3.47, 95% CI: 1.29–9.31,  $p = 0.014$ , Wald test). Patients with more than 34 breakpoints experienced a statistically shorter disease-free interval (Log rank test  $p < 0.001$ ), as determined by the Cox model and Kaplan-Meier analyses (Figure 4a).

The number of breakpoints was significantly predictive of metastasis-free interval

(  
< 34 breakpoints: 4 events in 83 patients; > 34 breakpoints: 4 events in 22 patients ( $p = 0.009$ , logrank test; RR: 5.29 [1.32; 21.26]) (Figure 4b).

### Discussion

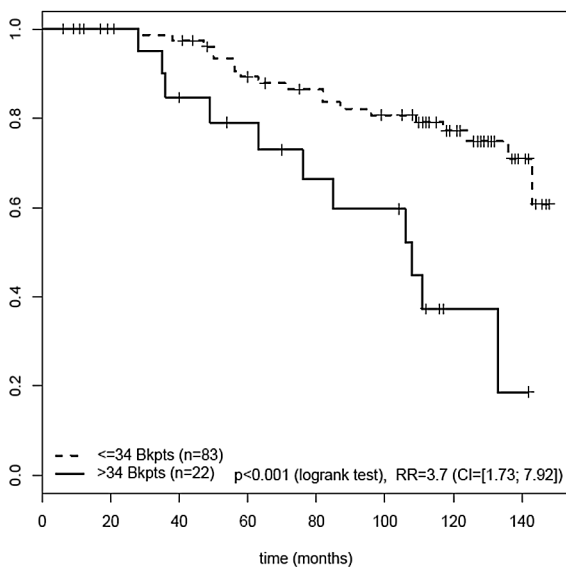
In this study, we assessed and validated the prognostic power of a genomic complexity signature based on the number

**Table 3.** Univariate and multivariate analyses for disease free-interval in the validation set.

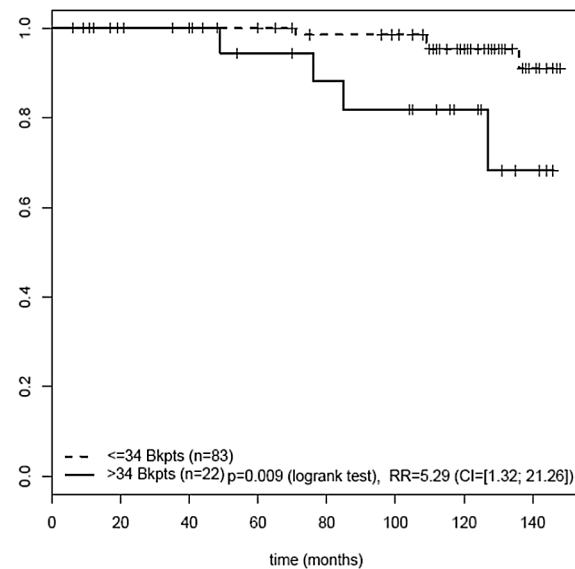
Variables	Univariate analysis		Multivariate analysis	
	Relative Risk (95% CI)	p	Relative Risk (95% CI)	p
Age <=50 vs >50 years	0.94 (0.45-1.98)	0.873		
Pathological size >20mm vs <=20mm	1.11 (0.42-2.92)	0.827		
Grade II/III vs I	2.12 (0.99-4.5)	0.052	1.39 (0.52-3.70)	0.515
PR status Negative vs Positive	0.9 (0.38-2.1)	0.805		
Vascular invasion Yes vs No	1.36 (0.55-3.36)	0.512		
KI67(IHC) >14 vs ≤ 14%	2.57 (1.16-5.68)	0.020	1.39 (0.48-4.08)	0.544
Genomic Grade GG3 vs GG1	2.59 (1.17-5.73)	0.019	1.16 (0.39-3.48)	0.788
IHC3 score (continuous)	1.01 (1.00-1.01)	0.204		
Number of Breakpoints >34 vs ≤34	3.7 (1.73-7.92)	<0.001	3.47 (1.29-9.31)	0.014

vs: versus.

doi: 10.1371/journal.pone.0076496.t003

**Disease Free Interval according to the number of breakpoints (n=105)**

a

**Metastasis free interval according to the number of breakpoints (n=105)**

b

**Figure 4.** Disease free-interval (a) and metastasis free-interval (b) analyses (Kaplan Meier analysis and log rank test) in the validation set.

doi: 10.1371/journal.pone.0076496.g004

of breakpoints determined with a pan-genomic SNP array in early stage node-negative luminal HER2-negative invasive ductal breast cancers. The number of breakpoints in this series outperformed the prognostic value of histoprognostic grade, KI67 index and genomic grade of the tumors.

Using high resolution genome-wide tools from different platforms, Russnes and colleagues [22] showed that, within the different molecular subtypes, qualitative genomic alterations and complex arm aberration index identified tumors associated with breast cancer-specific death. Others groups have shown the prognostic value of chromosomal instability score evaluated as the number of gains and losses assessed on SNP arrays,

either in ER positive or negative tumors [16,20,23]. However, these analyses did not focus on early stage (T1 and small T2 N0) luminal tumors.

In our previous work, we have identified a DNA-based genomic signature [21] on T1-T2 breast carcinomas but that encompassed ER and HER2 positive and negative cases. In the present study focused specifically on early stage luminal breast carcinomas, we have further demonstrated the prognostic value of the number of chromosomal breakpoints to assess chromosomal instability using a SNP-based array. ER-positive and HER2-negative, node-negative early stage breast carcinomas are the commonest form of breast cancers in the

context of systematic mammography screening. Tailoring treatment for these patients is mandatory but challenging [35]. The challenge is to identify poor prognosis patients when classical anatomical parameters such as size and node status are irrelevant and when the tumors are luminal [36]. The specificity of this new DNA-based signature was increased from 84% to 94%, which means that at initial diagnosis, this signature can contribute to accurate identification of good prognosis breast cancers, i.e. those with 34 breakpoints or less. This signature takes into account the number of breakpoints and not regions of gains or losses. Indeed, luminal invasive carcinomas not otherwise specified demonstrate highly heterogeneous genomic profiles [5] with numerous chromosomal gains and losses, but which are never observed in all cases in the same combination [14,22,34]. We therefore preferred to develop a quantitative signature than a qualitative assessment of the alterations. This approach is more easily reproducible using a standardized bioinformatic tool that can be applied in a customized fashion and is readily accessible [34].

Other authors and we have proposed the use of proliferation as a powerful prognostic tool to identify those tumors with poor outcome [27,28,36]. However, the present study showed that the number of chromosomal breakpoints in multivariate analysis remained the only statistically significant parameter to identify poor prognosis early invasive luminal HER2-negative and node-negative breast cancers, and was more reliable than proliferation assessed by either grade, KI67, IHC3or genomic grade.

Signature such as Oncotype DX® [37] has also been proposed as a robust tool to classify luminal breast carcinomas into three different prognostic groups. However, although the proliferation set of genes is robust and based on five genes encompassing KI67, the HER2 set of genes is more controversial [38], it is of no value in HER2-negative carcinomas. In addition, this test dramatically increases the cost of the diagnostic procedure [39] and it adds little prognostic information compared to the recently proposed cheaper IHC4 test [13,40,41].

In that context, we propose that determination of the number of breakpoints should help to refine the prognosis of early stage breast tumors [21,22]. Even though, the DNA-based signature is correlated with proliferation, it also assesses a different level of biological alterations and provides the global genomic profile of the tumor in a single experiment. It therefore allows an identification of the regions of amplifications, some of which encompass drivers genes that encode for druggable targets [42]. Furthermore, recent developments of Affymetrix®

arrays performed with DNA extracted from paraffin-embedded tissue provide promising and robust results like it has been obtained with other arrays [43].

## Conclusion

This study demonstrates that genomic complexity assessed by SNPs arrays and determined as the number of breakpoints can be used to predict the outcome of early stage luminal HER2-negative invasive breast carcinomas. But the use of chromosomal breakpoints number in clinical practice will require a prospective validation of its prognostic value in larger and multicentric series of breast cancer patients.

## Supporting Information

**Data S1. Antibodies used for tumor characterization.**  
(DOCX)

**Data S2. Mathematical analysis for SNP6.0 data segmentation.**  
(DOCX)

**Figure S1. Number of breakpoints according to KI67.**  
(PDF)

**Figure S2. Gene expression grade according to number of breakpoints.**  
(PDF)

## Acknowledgements

The authors thank Dr Anthony Saul for his help in editing the English of our manuscript, Jaydutt Bhalshankhar for his help to deposit transcriptomic and genomic data in a public depository.

## Author Contributions

Conceived and designed the experiments: AVS OD GP FR PC RR AF XSG. Performed the experiments: VB NG EG DG. Analyzed the data: AVS VB EG YdR SR GR GP. Contributed reagents/materials/analysis tools: AVS XSG AF RR OM FR PC DG. Wrote the manuscript: AVS OD EG. Designed the software used in the analysis to perform the SNP6.0 segmentation analysis: GR SR.

## References

1. Bedard PL, Cardoso F (2011) Can some patients avoid adjuvant chemotherapy for early-stage breast cancer? *Nat Rev Clin Oncol* 8: 272-279. doi:10.1038/nrclinonc.2011.19. PubMed: 21364525.
2. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S et al. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98: 10869-10874. doi:10.1073/pnas.191367098. PubMed: 11553815.
3. Foulkes WD, Reis-Filho JS, Narod SA (2010) Tumor size and survival in breast cancer--a reappraisal. *Nat Rev Clin Oncol* 7: 348-353. doi: 10.1038/nrclinonc.2010.39. PubMed: 20309006.
4. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M et al. (2005) Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy. *Clin Cancer Res* 11: 5678-5685. doi: 10.1158/1078-0432.CCR-04-2421. PubMed: 16115903.
5. Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL et al. (2012) Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* 486: 405-409. doi:10.1038/nature11154. PubMed: 22722202.
6. Network TCGA (2012) Comprehensive molecular portraits of human breast tumours. *Nature* 490: 61-70.



7. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C et al. (2012) The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486: 400-404. PubMed: 22722201.
8. Chin SF, Wang Y, Thorne NP, Teschendorff AE, Pinder SE et al. (2007) Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. *Oncogene* 26: 1959-1970. doi:10.1038/sj.onc.1209985. PubMed: 17001317.
9. Guedj M, Marisa L, de Reynies A, Orsetti B, Schiappa R et al. (2011) A refined molecular taxonomy of breast cancer. *Oncogene* 31: 1196-1206. PubMed: 21785460.
10. Ellis IO, Galea M, Broughton N, Locker A, Blamey RW et al. (1992) Pathological prognostic factors in breast cancer. II - Histological type. Relationship with survival in a large study with long-term follow-up. *Histopathology* 20: 479-489. PubMed: 1607149.
11. Luporsi E, André F, Spyrtos F, Martin PM, Jacquemier J et al. (2011) Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review. *Breast Cancer Res Treat* 132: 895-915. PubMed: 22048814.
12. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S et al. (2006) Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 98: 262-272. doi:10.1093/jnci/dj052. PubMed: 16478745.
13. Cuzick J, Dowsett M, Pineda S, Wale C, Salter J et al. (2011) Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol* 29: 4273-4278. doi:10.1200/JCO.2010.31.2835. PubMed: 21990413.
14. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R et al. (2006) Genomic and transcriptional aberrations linked to breast cancer pathophysiology. *Cancer Cell* 10: 529-541. doi:10.1016/j.ccr.2006.10.009. PubMed: 17157792.
15. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM et al. (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486: 346-352. PubMed: 22522925.
16. Birkbak NJ, Eklund AC, Li Q, McClelland SE, Endesfelder D et al. (2011) Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res* 71: 3447-3452. doi: 10.1158/0008-5472.CAN-10-3667. PubMed: 21270108.
17. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z (2006) A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet* 38: 1043-1048. doi:10.1038/ng1861. PubMed: 16921376.
18. Chibon F, Lagarde P, Salas S, Pérot G, Broust V et al. (2010) Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat Med* 16: 781-787. doi:10.1038/nm.2174. PubMed: 20581836.
19. Habermann JK, Doering J, Hautaniemi S, Roblick UJ, Bündgen NK et al. (2009) The gene expression signature of genomic instability in breast cancer is an independent predictor of clinical outcome. *Int J Cancer* 124: 1552-1564. doi:10.1002/ijc.24017. PubMed: 19101988.
20. Roylance R, Endesfelder D, Gorman P, Burrell RA, Sander J et al. (2011) Relationship of extreme chromosomal instability with long-term survival in a retrospective analysis of primary breast cancer. *Cancer Epidemiol Biomarkers Prev* 20: 2183-2194. doi: 10.1158/1055-9965.EPI-11-0343. PubMed: 21784954.
21. Gravier E, Pierron G, Vincent-Salomon A, Gruel N, Raynal V et al. (2010) A prognostic DNA signature for T1T2 node-negative breast cancer patients. *Genes Chromosomes Cancer* 49: 1125-1134. doi: 10.1002/gcc.20820. PubMed: 20842727.
22. Russnes HG, Vollan HK, Lingjaerde OC, Krasnitz A, Lundin P et al. (2010) Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med* 2: 38ra47. PubMed: 20592421.
23. Smid M, Hoes M, Sieuwerts AM, Sleijfer S, Zhang Y et al. (2010) Patterns and incidence of chromosomal instability and their prognostic relevance in breast cancer subtypes. *Breast Cancer Res Treat* 128: 23-30. PubMed: 20632083.
24. Blamey RW, Hornmark-Stenstam B, Ball G, Blichert-Toft M, Cataliotti L et al. (2009) ONCOPOOL - a European database for 16,944 cases of breast cancer. *Eur J Cancer* 46: 56-71.
25. Elston CW, Ellis IO (1991) Pathological Prognostic Factors in Breast cancer. The value of histological Grade in Breast cancer : experience from a large study with long-term follow-up. *Histopathology* 19: 403-410. doi:10.1111/j.1365-2559.1991.tb00229.x. PubMed: 1757079.
26. Vincent-Salomon A, Lucchesi C, Gruel N, Raynal V, Pierron G et al. (2008) Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. *Clin Cancer Res* 14: 1956-1965. doi: 10.1158/1078-0432.CCR-07-1465. PubMed: 18381933.
27. Reyat F, Bollet MA, Caly M, Gentien D, Carpentier S et al. (2012) Respective prognostic value of genomic grade and histological proliferation markers in early stage (pN0) breast carcinoma. *PLOS ONE* 7: e35184. doi:10.1371/journal.pone.0035184. PubMed: 22529987.
28. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC et al. (2011) Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst* 103: 1656-1664. doi:10.1093/jnci/djr393. PubMed: 21960707.
29. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B et al. (2011) Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 22: 1736-1747. doi:10.1093/annonc/mdr304. PubMed: 21709140.
30. Prat A, Cheang MC, Martin M, Parker JS, Carrasco E et al. (2013) Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. *J Clin Oncol* 31: 203-209. doi:10.1200/JCO.2012.43.8150. PubMed: 23233704.
31. Vincent-Salomon A, Raynal V, Lucchesi C, Gruel N, Delattre O (2008) ESR1 gene amplification in breast cancer: a common phenomenon? *Nat Genet* 40: 802; author reply: 802. doi:10.1038/ng0708-810. PubMed: 18583967.
32. Rigall G (2010). runed dynamic programming for optimal multiple change-point detection. arXiv:10040887v1
33. Zhang NR, Siegmund DO (2007) A modified Bayes information criterion with applications to the analysis of comparative genomic hybridization data. *Biometrics* 63: 22-32. doi:10.1111/j.1541-0420.2006.00662.x. PubMed: 17447926.
34. Picard F, Lebarbier E, Hoebeker M, Rigall G, Thiam B et al. (2011) Joint segmentation, calling, and normalization of multiple CGH profiles. *Biostatistics* 12: 413-428. doi:10.1093/biostatistics/kxq076. PubMed: 21209153.
35. Piccart-Gebhart MJ, Sotiriou C (2007) Adjuvant chemotherapy--yes or no? Prognostic markers in early breast cancer. *Ann Oncol* 18 Suppl 12: xii2-xii7. doi:10.1093/annonc/mdm532. PubMed: 18083699.
36. Stuart-Harris R, Caldas C, Pinder SE, Pharoah P (2008) Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients. *Breast* 17: 323-334. doi: 10.1016/j.breast.2008.02.002. PubMed: 18455396.
37. Paik S, Shak S, Tang G, Kim C, Baker J et al. (2004) A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *N Engl J Med* 351: 2817-2826. doi:10.1056/NEJMoa041588. PubMed: 15591335.
38. Ignatiadis M, Sotiriou C (2011) Breast cancer: Should we assess HER2 status by Oncotype DX? *Nat Rev Clin Oncol* 9: 12-14. doi:10.1038/nrclinonc.2011.188. PubMed: 22143144.
39. Dabbs DJ, Klein ME, Mohsin SK, Tubbs RR, Shuai Y et al. (2011) High false-negative rate of HER2 quantitative reverse transcription polymerase chain reaction of the Oncotype DX test: an independent quality assurance study. *J Clin Oncol* 29: 4279-4285. doi:10.1200/JCO.2011.34.7963. PubMed: 21990395.
40. Barton S, Zabaglio L, A'Hern R, Turner N, Ferguson T et al. (2012) Assessment of the contribution of the IHC4+C score to decision making in clinical practice in early breast cancer. *Br J Cancer* 106: 1760-1765. doi:10.1038/bjc.2012.166. PubMed: 22531639.
41. Tang G, Cuzick J, Costantino JP, Dowsett M, Forbes JF et al. (2011) Risk of recurrence and chemotherapy benefit for patients with node-negative, estrogen receptor-positive breast cancer: recurrence score alone and integrated with pathologic and clinical factors. *J Clin Oncol* 29: 4365-4372. doi:10.1200/JCO.2011.35.3714. PubMed: 22010013.
42. Arnedos M, Scott V, Job B, De La Cruz J, Commo F et al. (2012) Array CGH and PIK3CA/AKT1 mutations to drive patients to specific targeted agents: a clinical experience in 108 patients with metastatic breast cancer. *Eur J Cancer* 48: 2293-2299. doi:10.1016/j.ejca.2012.06.014. PubMed: 22840369.
43. Krijgsman O, Israeli D, Haan JC, van Essen HF, Smeets SJ et al. (2011) CGH arrays compared for DNA isolated from formalin-fixed, paraffin-embedded material. *Genes Chromosomes Cancer* 51: 344-352. PubMed: 22162309.