Published in final edited form as:

J Pathol. 2012 May ; 227(1): 42–52. doi:10.1002/path.3990.

Genomic and mutational profiling of ductal carcinomas *in situ* and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection

Lucia Hernandez^{#1,2}, Paul M Wilkerson^{#1}, Maryou B Lambros¹, Adriana Campion-Flora¹, Daniel Nava Rodrigues¹, Arnaud Gauthier^{1,3}, Cecilia Cabral¹, Vidya Pawar¹, Alan Mackay¹, Roger A'Hern⁴, Caterina Marchiò⁵, Jose Palacios⁶, Rachael Natrajan¹, Britta Weigelt^{1,7}, and Jorge S Reis-Filho¹

¹Molecular Pathology Team, The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London SW3 6JB, UK

²Department of Anatomic Pathology, Hospital Universitario 12 de Octubre, Madrid, Spain

³Department of Tumour Biology, Institut Curie, 75005 Paris, France

⁴Cancer Research UK Clinical Trials Unit, The Institute of Cancer Research, Sutton SM2 5NG, UK

⁵Department of Biomedical Sciences and Human Oncology, University of Turin, Turin, Italy

⁶Hospital Universitario Virgen del Rocio, Seville, Spain

⁷Signal Transduction Laboratory, Cancer Research UK London Research Institute, London WC2A 3LY, UK

These authors contributed equally to this work.

Abstract

The mechanisms underlying the progression from ductal carcinoma *in situ* (DCIS) to invasive ductal carcinoma (IDC) of the breast are yet to be fully elucidated. Several hypotheses have been put forward to explain the progression from DCIS to IDC, including the selection of a subpopulation of cancer cells with specific genetic aberrations, the acquisition of new genetic aberrations or non-genetic mechanisms mediated by the tumour microenvironment. To determine whether synchronously diagnosed ipsilateral DCIS and IDCs have modal populations with distinct repertoires of gene copy number aberrations and mutations in common oncogenes, matched frozen samples of DCIS and IDCs were retrieved from 13 patients and subjected to microarray-based

Corresponding author: Jorge S Reis-Filho, MD PhD FRCPath, The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, 237 Fulham Road, London, UK, SW3 6JB, Tel: +44 2071535167; Fax: +44 2071535533; Jorge.Reis-Filho@icr.ac.uk.

Conflict of interest statement: The authors declare no conflicts.

Microarray data: Data and the analysis history, script and code are available at (<http://rock.icr.ac.uk/collaborations/Mackay/DCIS.invasive.progression>).

Statement of Author Contributions

LH, PMW, RN, and JRF conceived and designed the experiments. LH, PMW, MBL, CC, AG, ACF, DNR, VP and RN performed the experiments. PMW, MBL, AM, JP, CM, BW and JRF analysed the data. PMW, BW and JRF wrote the manuscript. All authors revised and approved the final draft.

comparative genomic hybridisation (aCGH), and Sequenom MassARRAY (Oncocarta v1.0 panel). Fluorescence *in situ* hybridisation and Sanger sequencing were employed to validate the aCGH and Sequenom findings, respectively. Although the genomic profiles of matched DCIS and IDCs were similar, in three of 13 matched pairs amplification of distinct loci (i.e. 1q41, 2q24.2, 6q22.31, 7q11.21, 8q21.2 and 9p13.3) was either restricted to, or more prevalent in, the modal population of cancer cells of one of the components. Sequenom MassARRAY identified *PIK3CA* mutations restricted to the DCIS component in two cases, and in a third case, the frequency of the *PIK3CA* mutant allele reduced from 49% in the DCIS to 25% in the IDC component. Despite the genomic similarities between synchronous DCIS and IDC, our data provide strong circumstantial evidence to suggest that in some cases the progression from DCIS to IDC is driven by the selection of non-modal clones that harbour a specific repertoire of genetic aberrations.

Keywords

DCIS; progression; microarray-based comparative genomic hybridisation; fluorescence *in situ* hybridisation; Sequenom MassARRAY; evolution; heterogeneity

Introduction

Breast cancer is perceived as a collection of distinct entities affecting the same anatomical site, which have different risk factors, histological features, clinical behaviour and response to therapy [1,2]. Recent studies have demonstrated that individual breast cancers are composed of a mosaic of non-modal clones, each with distinct constellations of genomic aberrations [3–9].

Since the introduction of mammography in the 1980s, the diagnosis of ductal carcinoma *in situ* (DCIS) has risen faster than any other subtype of breast cancer, accounting for 15–25% of new cases in the United States [10]. DCIS have been shown to constitute *bona fide* non-obligate precursors of invasive breast cancer based on the observations that if untreated, up to 39% of low grade DCIS progress to invasive carcinomas [11], and that DCIS and invasive breast cancers arising in the same quadrant of the breast share similar histological features, gene expression profiles and gene copy number aberrations [12–15]. The progression from DCIS to invasive breast cancer remains an area of great controversy [16,17]. Historically, this process was perceived as a passive phenomenon, where the cancer cells would simply invade through the myoepithelial layer and basement membrane; however, recent studies have suggested that the microenvironment plays an active role in this process [18], with cancer-associated fibroblasts promoting invasion and myoepithelial cells acting as potential invasion suppressors [19,20]. Furthermore, abnormalities in the extracellular matrix and infiltrating leukocytes have been implicated in the process of invasion [21–23].

Despite the similarities in the genomic profiles of synchronous DCIS and IDC [15], there are clear examples of genetic differences between the *in situ* and invasive components of a cancer. This is perhaps best exemplified by the cases of DCIS harbouring *HER2* gene amplification associated with *HER2*-negative invasive carcinomas [24,25]. These observations may be explained by the fact i) that in these cases, the DCIS analysed was not the precursor of the invasive cancer, ii) that *HER2* amplification was lost during

progression from DCIS to invasive carcinoma, or iii) that the DCIS was composed of a mosaic of clones with distinct genetic aberrations and the cancer cells that eventually invaded did not harbour *HER2* gene amplification.

Given the evidence to suggest that genetic differences may exist between matched DCIS and IDC samples, we hypothesised that in some cases, progression may be mediated by selection of a subpopulation of cancer cells or by acquisition of new aberrations, including gene copy number aberrations or gene mutations. The aims of this study were to determine if matched DCIS and adjacent invasive carcinomas from the same patient (and the same histological block) harbour distinct patterns of copy number aberrations or mutations in known cancer genes.

Material and Methods

Power calculations

This is a hypothesis generating study, as there is limited information on the genome-wide similarities and differences between synchronously diagnosed DCIS and IDC from the same breast. The power calculations were based on the assumption that if progression from DCIS to IDC would be determined by a 'gene copy number aberration' or a 'mutation', this aberration would be expected to be present in the invasive component of the vast majority of carcinomas (e.g. >75%) but not present in the DCIS component (i.e. <5%). If these two levels were exactly 75% and 5% respectively, then nine of 12 samples would be expected to harbour the aberration in the invasive component, which would not be present in the DCIS component, and neither component would harbour the aberration in three cases. With 12 paired samples there is an 80% chance (power) that such an aberration would be detected using a one sided p-value of 1% to allow for multiple testing (McNemar test).

Cases

Fresh frozen breast cancer samples from patients whose tumours were reported to contain foci of both DCIS and invasive breast carcinomas were retrieved from Hospital Universitario 12 de Octubre, Madrid, and reviewed by at least two pathologists (LH, CC, ACF, DNR and/or JRF). Thirteen cases containing *bona fide* areas of DCIS and invasive breast cancer in the same tissue specimen were available for analysis in this study. Samples were anonymised prior to analysis and the study approved by local ethical committees. For each tumour, histological grade was assessed using Nottingham grading system [26], by at least two pathologists (LH, CC, ACF, DNR and/or JRF). The clinicopathological characteristics of the tumours analysed in this study are detailed in Table 1.

Immunohistochemistry

Immunohistochemistry was performed on representative 3µm thick sections containing the DCIS and matched invasive breast carcinomas subjected to microarray-based comparative genomic hybridisations (aCGH), fluorescence *in situ* hybridisation (FISH), and Sequenom MassARRAY. The antibodies, immunohistochemistry protocols and scoring methods used to characterise each tumour are summarised in Supplementary Table S1. In brief, oestrogen receptor (ER) and progesterone receptor (PR) were considered positive if the Allred score

was >2. HER2 was scored according to the current American Society of Clinical Oncology (ASCO)/ College of American Pathologists (CAP) guidelines [27]. Ki67 was scored following a previously described [28] semi-quantitative scoring system where lesions were considered Ki67 low if <10%, intermediate if 10%-30%, and high if >30% of the neoplastic cells displayed nuclear Ki67 expression. p53 was analysed as previously described [29]; cases were considered to have positive p53 nuclear expression (i.e. an imperfect surrogate for *TP53* mutations) if >10% of the neoplastic cells displayed nuclear expression. The results of immunohistochemical analysis were interpreted by at least two pathologists (LH, CC, ACF, DNR and/or JRF), blinded to the results of aCGH, FISH and Sequenom MassARRAY analysis.

Microdissection and DNA extraction

For all 13 cases, 20 representative 10µm thick sections were cut from the frozen tissue blocks. Microdissection was performed with a sterile needle under a stereomicroscope (Olympus SZ61, Tokyo, Japan) to ensure a percentage of tumour cells greater than 70%, as previously described [30,31]. DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit (Hamburg, Germany). Double-stranded DNA concentration was measured using the Picogreen® assay, according to the manufacturer's instructions (Invitrogen, Paisley, UK). DNA quality was assessed using four primer sets in a multiplex PCR, as previously described [32]. Out of the 13 cases, both the DCIS and IDC components of 13 and 12 cases yielded DNA of sufficient quantity and quality for aCGH and Sequenom MassARRAY analysis, respectively (i.e. insufficient DNA was available for Sequenom MassARRAY analysis of Case 7).

Microarray-based comparative genomic hybridisation

The aCGH platform used for this study comprises ~32,000 BAC clones tiled across the genome, which has been shown to be as robust as, and to have comparable resolution with, high-density oligonucleotide arrays [33–35]. DNA labelling, array hybridisation and image acquisition were performed as previously described [3,30,32,36,37]. aCGH data were pre-processed and analysed using the BACE.R R script (R version 2.11.1) as previously described [30,37]. After filtering polymorphic BACs and BACs mapping to chromosome Y, a final dataset of 31,158 clones with unambiguous mapping information according to build hg19 of the human genome (<http://www.ensembl.org>) was smoothed using the circular binary segmentation (cbs) algorithm [3,30,37]. A categorical analysis was applied to the BACs after classifying them as representing amplification (>0.45), gain (>0.08 and <0.45), loss (<-0.08), or no change according to their cbs-smoothed Log₂ ratio values [36]. Threshold values were determined and validated as previously described [36]. Categorical data were subjected to a multi-Fisher's exact test with adjustment for multiple-testing using the step-down permutation procedure maxT, providing strong control of the family-wise type I error rate, as previously described [30,37], to identify statistically significant differences between the genomic profiles of matched DCIS and IDC samples. Unsupervised hierarchical clustering analysis was performed as previously described [30,37]. Briefly, cbs-smoothed ratios were used for clustering, employing Ward's clustering algorithm and correlation distance. Data and the analysis history, script and code are available at (<http://rock.icr.ac.uk/collaborations/Mackay/DCIS.invasive.progression>).

Fluorescence *in situ* hybridisation (FISH)

FISH was employed to validate regions of amplification identified by aCGH as previously described [38]. In-house Digoxigenin- or Biotin-labelled probes mapping to the smallest region of amplification in each locus showing enrichment of amplification in either the DCIS or invasive breast cancer component of cases analysed in this study (i.e. 1q41, 2q24.2, 6q22.31, 7q11.21, 8q21.2 and 9p13.3, Supplementary Table S2) were constructed as previously described [39]. FISH using these probes was carried out as previously described [39]. Copy number signals were counted in the nuclei of at least 100 morphologically unequivocal neoplastic cells from each component, and amplification was defined as 5 or more copies per nucleus, large clusters or a combination of both, in more than 50% of nuclei analysed, as previously described. [32,37]. The proportion of nuclei defined as harbouring amplification for each locus in each component was independently analysed by two observers (PMW and ACF) blinded to the results of the aCGH analyses, and used to compare the frequency of amplification among the populations of tumour cells analysed in each component. For comparative analysis, a semi-quantitative scoring system was adopted, where morphologically unequivocal neoplastic cells were classified as having 1 copy, 2-4 copies, or 5 or more copies per nucleus, large clusters or a combination of both.

Sequenom MassARRAY OncoCarta

DNA from each component of 12 out of 13 breast cancer samples included in this study was subjected to mutation screening to detect 238 mutations in 19 oncogenes using the OncoCarta Panel v1.0 (Sequenom Inc., San Diego, CA) as previously described [30]. The prevalence of mutant alleles was estimated by calculating the ratio of the area of the raw spectra of the mutant allele to its wild type. Mutations were validated using Sanger sequencing as previously described [30], where primers were designed on Primer3 (<http://frodo.wi.mit.edu/primer3/>). Primer sequences are detailed in Supplementary Table S3. Sequences were visualised using 4Peaks (<http://4peaks.en.softonic.com/>) and analysed using Mutation Surveyor (Softgenetics, PA, USA).

Statistical analysis

With the exception of aCGH analysis, which was performed using the R package (version 2.11.1) as described above, all statistical analyses were performed using Prism v5.04 (Graphpad Software Inc, La Jolla, CA, USA). Comparisons of the proportions of amplified and non-amplified nuclei were performed using a Chi-square test. All *p* values were two-tailed and 95% confidence intervals were adopted. A *p* value < 0.05 was considered significant.

Results

Matched DCIS and adjacent invasive carcinomas have similar genomic profiles

To characterise the genomic profiles of matched DCIS and adjacent invasive carcinomas, we performed aCGH of 13 pairs of microdissected DCIS and adjacent invasive carcinomas. After exclusion of regions mapping to known copy number polymorphisms (<http://projects.tcag.ca/variation/>), aCGH analysis demonstrated that *in situ* and invasive

components harboured similar numbers of copy number aberrations (mean of 12.6%, range of 1.0%-28.4% vs mean of 10.1%, range of 1.0%-27.1%, respectively, of BACs showing either gains, losses or amplifications; Student's *t* test, *p* = 0.498). The most frequently amplified genomic regions (three out of 13 cases) mapped to 17q12, encompassing *HER2*, and 11q3-q14, encompassing *CCND1*. These regions were amplified in both the DCIS and adjacent invasive carcinomas components of each of the three cases (Supplementary Figure S1 and Supplementary Table S4).

We have previously demonstrated that statistically significant differences in copy number profiles between two groups of tumours can be determined by performing grouped analyses of aCGH copy number data using a multi-Fisher's exact test [30,32,37]. We therefore applied the same principle in this study and performed a grouped analysis using a multi-Fisher's exact test to compare the copy number profiles of DCIS and adjacent invasive breast carcinoma components. No statistically significant differences in copy number profile between the two groups were identified (Figure 1A). This observation was further corroborated by unsupervised hierarchical clustering, which revealed that each DCIS more closely resembled its respective adjacent invasive carcinoma than any other sample (Figure 1B). Taken together, our findings confirm previous observations that synchronously diagnosed matched DCIS and invasive breast cancers display similar patterns of molecular aberrations [12–15,40,41].

Detailed pair-wise analysis of DCIS and its adjacent invasive carcinoma reveals the presence of gene copy number aberrations restricted to one of the components

It is possible that the lack of significant differences between DCIS and their respective invasive carcinomas, when analysed as a group, stem from i) the lack of a common genetic aberration that drives the progression from *in situ* to invasive breast cancer and ii) the possibility that the differences in gene copy numbers may be quantitative rather than qualitative. A pair-wise comparison of the genomic profiles of the matched samples of DCIS and adjacent invasive carcinomas revealed differences in the presence and level of specific amplifications between the two components in three cases (Cases 2, 4, and 5; Figure 2). The cbs-smoothed Log₂ aCGH ratio of four amplicons (i.e. 1q41, 6q22.31, 7q11.21, and 8q21.2) was higher in the invasive than in the DCIS component (i.e. Case 2, amplicons 1q14 and 7q11.21; Case 4, amplicon 8q21.2; Case 5, amplicon 6q22.31) and in two amplicons (i.e. 9p13.3 and 2q24.2), the aCGH Log₂ ratios were lower in the invasive than in the DCIS component (i.e. Case 4, amplicon 2q24.2; Case 5, amplicon 9p13.3). Given that the samples were microdissected to ensure tumour cell content of >70%, and that the direction of the change in Log₂ ratio was not unidirectional, normal cell DNA contamination of the DNA samples subjected to aCGH is unlikely to account for the observed differences. Of note, the three cases harbouring differences in copy number at the above loci between the DCIS and invasive components were of histological grade 3, displayed *HER2* overexpression and harboured *HER2* gene amplification (Table 1). To determine if the increase in copy number at these loci was associated with an enrichment of cells harbouring specific amplifications in each component, each case harbouring an amplification with distinct copy numbers in the DCIS or adjacent invasive carcinoma was subjected to FISH using in-house probes. In each case, FISH revealed that the components with the lower cbs-smoothed Log₂ aCGH ratios of

the amplified loci were composed of mosaics of cells with or devoid of the amplification, and the population of cancer cells harbouring the amplification usually accounted for <60% of the cancer cell population. In fact, the differences in the proportion of cells harbouring amplification in between two components of the same case closely mirrored the corresponding differences in the cbs-smoothed Log₂ aCGH ratios (Figure 3). In all cases, the differences in the distribution of copy number for each locus between each component were statistically significant (Chi-square test, $p < 0.0002$). These findings provide strong circumstantial evidence of clonal selection in the progression from *in situ* to invasive cancer [12].

Patterns of mutations in cancer oncogenes in DCIS and adjacent invasive carcinomas

To determine if *in situ* and invasive breast cancers harbour distinct patterns of mutations in known cancer genes, 12 paired samples were subjected to Sequenom mutational profiling using the Oncocarta v1.0 panel as previously described [30]. Out of all genes tested, only *PIK3CA* was found to be somatically mutated in the samples investigated. *PIK3CA* mutations were identified in 6/12 cases (50%), consistent with previous reports that *PIK3CA* mutations are common in ER-positive luminal cancers [42] and in DCIS [43] (Supplementary Table S5). In a grouped comparison, considering only if a component was positive or negative for *PIK3CA* mutations (rather than the frequency of the mutant allele), no significant difference was identified in the frequency of mutations between DCIS and adjacent invasive carcinomas (Fisher's exact test, $p=0.68$). Pair-wise analysis of the six cases harbouring *PIK3CA* mutations in either component identified differences in mutant allele frequency between components. In three of these cases (Cases 3, 8, and 10), the mutant allele frequency was similar in the *in situ* and invasive components (c.1633G>A, p.545E>K, range 19.2-54.2%, Figure 4A). In the remaining three cases, the mutant allele frequency was higher in the *in situ* than in the invasive component (Cases 2 and 12 with c.1624G>A, p.542E>K, and Case 5 with c.3140A>G, p.1047H>R, range 11.5-48.8%, Figures 4B and 4C). In fact, in two of these cases (Case 12 harbouring an E542K mutation, and Case 5 an H1047R mutation), the *PIK3CA* mutation was absent from the invasive component. Sanger sequencing analysis of the same DNA samples confirmed all E545K *PIK3CA* mutations and 2 out of 3 E542K *PIK3CA* mutations. One of the E542K *PIK3CA* mutations and the H1047R *PIK3CA* mutation were not validated by Sanger sequencing (Supplementary Table S5). This is not surprising, given that in these samples, the frequency of the mutated allele was estimated by Sequenom to be either at lower limit of detection by Sanger sequencing (i.e. 21% E542K *PIK3CA* mutation in Case 12) or below the threshold of mutation detection (i.e. 11.5% H1047R *PIK3CA* mutation in Case 5). Despite the very low allelic frequency of the H1047R mutation in Case 5, a small peak representing an A to a G base change consistent with the H1047R mutation can be identified (Figure 4C). Taken together, these data confirm previous observations [43] that *PIK3CA* mutations are an early event in breast cancer, are frequently found in ER-positive DCIS, and suggest *PIK3CA* mutations are unlikely to play a role in progression from DCIS to invasive breast cancer.

Discussion

This hypothesis-generating pilot study provides direct evidence to reconcile some of the discrepant observations on the molecular characteristics of synchronously diagnosed matched DCIS and adjacent invasive breast carcinomas. First, in agreement with previous studies [12,14,40,41], we demonstrate that overall, the genomic profiles of matched DCIS and adjacent invasive breast carcinomas are remarkably similar in terms of their copy number profiles. These data demonstrate that DCIS is as advanced as its invasive counterpart in terms of the genomic characteristics of their modal populations. We extended these observations to the mutational repertoire of matched DCIS and invasive breast cancers and demonstrated that from a qualitative viewpoint, these lesions are remarkably similar. Second, we present direct evidence using orthogonal methods (i.e. aCGH and FISH, Sequenom and Sanger sequencing) that at least some matched DCIS and synchronously diagnosed adjacent invasive carcinomas are composed of modal populations of cancer cells that differ in their repertoire of gene copy number aberrations and gene mutations. These findings provide direct evidence to support the notion that intra-tumour genetic heterogeneity is not uncommon in breast cancer [3,5–8] and that it exists at the early stages of tumour development (i.e. DCIS).

The observations that DCIS and synchronously diagnosed adjacent invasive carcinomas do not show any significant differences in terms of copy number aberrations and mutations in known cancer genes when analysed as a group, while pair-wise analysis of each matched pair occasionally revealed important differences are not contradictory. In fact, the lack of significant differences is likely to stem from the fact that even if gene copy number aberrations drive the process of progression from DCIS to invasive carcinoma, progression may be a convergent phenotype and different genetic aberrations may result in invasion depending on the epistatic interactions resulting from the genomic instability found in these lesions [8,9]. In the absence of a common denominator in the form of gene copy number aberrations or mutations in known cancer genes, the analyses performed would not be able to identify statistically significant changes. It is also possible that genetic aberrations other than gene copy number changes or mutations in known cancer genes (e.g. somatic point mutations or small insertions and deletions in other genes, and somatic rearrangements), or epigenetic changes [44,45] may drive this phenomenon. Further analysis based on massively parallel sequencing of synchronously diagnosed matched DCIS and invasive breast cancers are warranted.

Here we performed the most comprehensive profiling for mutations in known cancer genes in matched DCIS and adjacent invasive breast carcinomas. We have confirmed that activating *PIK3CA* mutations are a frequent phenomenon in DCIS [43]. In agreement with Miron et al. [43], we observed differences in the presence of *PIK3CA* mutations between the DCIS and adjacent invasive carcinoma. Our analyses using orthogonal methods demonstrate that although good concordance in the qualitative *PIK3CA* mutation status (i.e. mutant vs non-mutant) between DCIS and adjacent invasive breast cancer was observed as previously reported [43,46,47], quantitative differences were observed in that the clones harbouring a *PIK3CA* mutation were more prevalent in the DCIS than in the adjacent invasive carcinoma. These observations suggest that activating mutations in this gene are more likely to play a

role in breast tumour initiation than in invasive progression [43], and do not support the contention that *PIK3CA* mutations occur early and are selected for in breast cancer progression [46]. In fact, in two of the cases reported by Kalinsky et al. [46], quantitative changes in the prevalence of *PIK3CA* mutations from DCIS to invasive carcinoma were also observed [46]. Alternative hypotheses for our findings include the possibility i) that in some breast cancers, *PIK3CA* mutations may not be driver events due to other epistatic interactions (e.g. other genetic or epigenetic alterations resulting in activation of the PI3K pathway); and ii) that although *PIK3CA* mutations occur relatively early, they are not initiating events, given that in some DCIS samples they were shown to be restricted to a subpopulation of cancer cells. In this context, other genetic or epigenetic alterations may confer an advantage for a subclone of cancer cells to be able to progress to invasive breast cancer.

Although phenotypic heterogeneity in regards to the expression of multiple immunohistochemical markers in DCIS [48,49] and genetic differences between unmatched DCIS and invasive breast cancers [14] have been previously documented, the finding of genetic heterogeneity for selected amplifications between synchronously diagnosed matched DCIS and invasive breast cancers is novel, but not entirely surprising. Previous studies have either examined DCIS samples alone [15,50], studied unmatched groups of DCIS and IDC samples [14,51], or surveyed only specific loci in the genome [40]. These approaches are not equipped to identify low frequency events involving subtle changes in copy number profile between two cases, as we have demonstrated in this study (Figures 1, 2 and 3). The findings of this study highlight the difficulties in studying tumour progression, namely that it appears that subsets of tumours harbour distinct mechanisms driving progression. Therefore, future studies of tumour progression should be designed with this in mind, using only matched pairs of tumours, and performing analyses in a pair-wise fashion.

The limitations of this study include the small sample size, rendering this pilot study hypothesis-generating. Eleven of the 13 cases were ER-positive, making the generalisability of the findings to ER-negative breast cancer limited. Given the lack of adequate *in vitro* models of ER-positive DCIS, no mechanistic/ functional validation of the identified aberrations (e.g. *PIK3CA* mutations) could be performed. In fact, the most widely used *in vitro* models of DCIS (e.g. the MCF10A progression series) are representative of ER-negative disease [52].

In conclusion, here we demonstrate that although the modal populations of synchronously diagnosed matched DCIS and adjacent invasive carcinomas are similar at the genetic level, intra-tumour genetic heterogeneity does exist from the early stages of breast cancer progression. We have also provided evidence to demonstrate that at least some DCIS and invasive breast carcinomas are composed of mosaics of neoplastic cells that harbour additional genetic aberrations to the founder genetic events and the selection of populations with specific genomic aberrations take place in the progression from DCIS to invasive breast cancer. Based on our observations, it is likely that progression from DCIS to invasive cancer is likely to constitute a complex biological phenomenon that follows a Darwinian evolution model and a convergent phenotype (i.e. progression from DCIS to invasive cancer may be caused by a large constellation of genetic and/ or epigenetic aberrations, and/ or be mediated

by the microenvironment) [9]. Therefore, further studies are warranted to characterise the entire repertoire of genetic and epigenetic changes associated with progression through massively parallel sequencing analysis of matched samples of DCIS and IDC.

Supplementary Materials

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was in part funded by Breakthrough Breast Cancer. LH is funded by a grant from the Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación. PMW is funded by a Wellcome Trust clinical fellowship grant. BW is funded by a CRUK postdoctoral fellowship. We acknowledge NHS funding to the NIHR Biomedical Research Centre. The study sponsors had no involvement in the design of this perspective, the literature review, data interpretation, writing of the manuscript or the decision to submit it for publication.

References

1. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet*. 2011; 378:1812–1823. [PubMed: 22098854]
2. Weigelt B, Pusztai L, Ashworth A, et al. Challenges translating breast cancer gene signatures into the clinic. *Nat Rev Clin Oncol*. 2011; 9:58–64. [PubMed: 21878891]
3. Geyer FC, Weigelt B, Natrajan R, et al. Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. *J Pathol*. 2010; 220:562–573. [PubMed: 20099298]
4. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011; 472:90–94. [PubMed: 21399628]
5. Navin N, Krasnitz A, Rodgers L, et al. Inferring tumor progression from genomic heterogeneity. *Genome Res*. 2010; 20:68–80. [PubMed: 19903760]
6. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature*. 2010; 464:999–1005. [PubMed: 20393555]
7. Shah SP, Morin RD, Khattra J, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature*. 2009; 461:809–813. [PubMed: 19812674]
8. Gerlinger M, Swanton C. How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br J Cancer*. 2010; 103:1139–1143. [PubMed: 20877357]
9. Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. *Cell*. 2011; 145:30–38. [PubMed: 21458666]
10. Burstein HJ, Polyak K, Wong JS, et al. Ductal carcinoma in situ of the breast. *N Engl J Med*. 2004; 350:1430–1441. [PubMed: 15070793]
11. Sanders ME, Schuyler PA, Dupont WD, et al. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer*. 2005; 103:2481–2484. [PubMed: 15884091]
12. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, et al. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology*. 2010; 57:171–192. [PubMed: 20500230]
13. Ma XJ, Salunga R, Tuggle JT, et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A*. 2003; 100:5974–5979. [PubMed: 12714683]
14. Robanus-Maandag EC, Bosch CA, Kristel PM, et al. Association of C-MYC amplification with progression from the in situ to the invasive stage in C-MYC-amplified breast carcinomas. *J Pathol*. 2003; 201:75–82. [PubMed: 12950019]
15. Vincent-Salomon A, Lucchesi C, Gruel N, et al. Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. *Clin Cancer Res*. 2008; 14:1956–1965. [PubMed: 18381933]
16. Polyak K, Hu M. Do myoepithelial cells hold the key for breast tumor progression? *J Mammary Gland Biol Neoplasia*. 2005; 10:231–247. [PubMed: 16807803]

17. Schnitt SJ. The transition from ductal carcinoma in situ to invasive breast cancer: the other side of the coin. *Breast Cancer Res.* 2009; 11:101. [PubMed: 19291276]
18. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144:646–674. [PubMed: 21376230]
19. Hu M, Yao J, Carroll DK, et al. Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell.* 2008; 13:394–406. [PubMed: 18455123]
20. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. *Breast Cancer Res.* 2011; 13:227. [PubMed: 22078026]
21. Chabottaux V, Noel A. Breast cancer progression: insights into multifaceted matrix metalloproteinases. *Clin Exp Metastasis.* 2007; 24:647–656. [PubMed: 17968664]
22. Levental KR, Yu H, Kass L, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell.* 2009; 139:891–906. [PubMed: 19931152]
23. Barker HE, Chang J, Cox TR, et al. LOXL2-mediated matrix remodeling in metastasis and mammary gland involution. *Cancer Res.* 2011; 71:1561–1572. [PubMed: 21233336]
24. Latta EK, Tjan S, Parkes RK, et al. The role of HER2/neu overexpression/amplification in the progression of ductal carcinoma in situ to invasive carcinoma of the breast. *Mod Pathol.* 2002; 15:1318–1325. [PubMed: 12481013]
25. Park K, Han S, Kim HJ, et al. HER2 status in pure ductal carcinoma in situ and in the intraductal and invasive components of invasive ductal carcinoma determined by fluorescence in situ hybridization and immunohistochemistry. *Histopathology.* 2006; 48:702–707. [PubMed: 16681686]
26. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology.* 1991; 19:403–410. [PubMed: 1757079]
27. Wolff AC, Hammond ME, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol.* 2007; 25:118–145. [PubMed: 17159189]
28. Tan DS, Marchio C, Jones RL, et al. Triple negative breast cancer: molecular profiling and prognostic impact in adjuvant anthracycline-treated patients. *Breast Cancer Res Treat.* 2008; 111:27–44. [PubMed: 17922188]
29. Abdel-Fatah TM, Powe DG, Agboola J, et al. The biological, clinical and prognostic implications of p53 transcriptional pathways in breast cancers. *J Pathol.* 2010; 220:419–434. [PubMed: 20044801]
30. Duprez R, Wilkerson PM, Lacroix-Triki M, et al. Immunophenotypic and genomic characterisation of papillary carcinomas of the breast. *J Pathol.* 2011 epub ahead of print.
31. Wetterkog D, Lopez-Garcia MA, Lambros MB, et al. Adenoid cystic carcinomas constitute a genomically distinct subgroup of triple-negative and basal-like breast cancers. *J Pathol.* 2012; 226:84–96. [PubMed: 22015727]
32. Marchio C, Iravani M, Natrajan R, et al. Mixed micropapillary-ductal carcinomas of the breast: a genomic and immunohistochemical analysis of morphologically distinct components. *J Pathol.* 2009; 218:301–315. [PubMed: 19479727]
33. Coe BP, Ylstra B, Carvalho B, et al. Resolving the resolution of array CGH. *Genomics.* 2007; 89:647–653. [PubMed: 17276656]
34. Tan DS, Lambros MB, Natrajan R, et al. Getting it right: designing microarray (and not 'microarray') comparative genomic hybridization studies for cancer research. *Lab Invest.* 2007; 87:737–754. [PubMed: 17558419]
35. Gunnarsson R, Staaf J, Jansson M, et al. Screening for copy-number alterations and loss of heterozygosity in chronic lymphocytic leukemia—a comparative study of four differently designed, high resolution microarray platforms. *Genes Chromosomes Cancer.* 2008; 47:697–711. [PubMed: 18484635]
36. Natrajan R, Lambros MB, Rodriguez-Pinilla SM, et al. Tiling path genomic profiling of grade 3 invasive ductal breast cancers. *Clin Cancer Res.* 2009; 15:2711–2722. [PubMed: 19318498]

37. Lacroix-Triki M, Suarez PH, MacKay A, et al. Mucinous carcinoma of the breast is genomically distinct from invasive ductal carcinomas of no special type. *J Pathol.* 2010; 222:282–298. [PubMed: 20815046]
38. Lambros MB, Natrajan R, Reis-Filho JS. Chromogenic and fluorescent in situ hybridization in breast cancer. *Hum Pathol.* 2007; 38:1105–1122. [PubMed: 17640550]
39. Lambros MB, Simpson PT, Jones C, et al. Unlocking pathology archives for molecular genetic studies: a reliable method to generate probes for chromogenic and fluorescent in situ hybridization. *Lab Invest.* 2006; 86:398–408. [PubMed: 16446704]
40. Moelans CB, de Wegers RA, Monsuurs HN, et al. Molecular differences between ductal carcinoma in situ and adjacent invasive breast carcinoma: a multiplex ligation-dependent probe amplification study. *Cell Oncol (Dordr).* 2011; 34:475–482. [PubMed: 21547576]
41. Buerger H, Otterbach F, Simon R, et al. Comparative genomic hybridization of ductal carcinoma in situ of the breast-evidence of multiple genetic pathways. *J Pathol.* 1999; 187:396–402. [PubMed: 10398097]
42. Di Cosimo S, Baselga J. Phosphoinositide 3-kinase mutations in breast cancer: a “good” activating mutation? *Clin Cancer Res.* 2009; 15:5017–5019. [PubMed: 19671849]
43. Miron A, Varadi M, Carrasco D, et al. PIK3CA mutations in in situ and invasive breast carcinomas. *Cancer Res.* 2010; 70:5674–5678. [PubMed: 20551053]
44. Qi L, Bart J, Tan LP, et al. Expression of miR-21 and its targets (PTEN, PDCD4, TM1) in flat epithelial atypia of the breast in relation to ductal carcinoma in situ and invasive carcinoma. *BMC Cancer.* 2009; 9:163. [PubMed: 19473551]
45. Hu M, Yao J, Cai L, et al. Distinct epigenetic changes in the stromal cells of breast cancers. *Nat Genet.* 2005; 37:899–905. [PubMed: 16007089]
46. Kalinsky K, Heguy A, Bhanot UK, et al. PIK3CA mutations rarely demonstrate genotypic intratumoral heterogeneity and are selected for in breast cancer progression. *Breast Cancer Res Treat.* 2011; 129:635–643. [PubMed: 21617917]
47. Dunlap J, Le C, Shukla A, et al. Phosphatidylinositol-3-kinase and AKT1 mutations occur early in breast carcinoma. *Breast Cancer Res Treat.* 2010; 120:409–418. [PubMed: 19418217]
48. Allred DC, Wu Y, Mao S, et al. Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. *Clin Cancer Res.* 2008; 14:370–378. [PubMed: 18223211]
49. Park SY, Gonen M, Kim HJ, et al. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *J Clin Invest.* 2010; 120:636–644. [PubMed: 20101094]
50. Hwang ES, DeVries S, Chew KL, et al. Patterns of chromosomal alterations in breast ductal carcinoma in situ. *Clin Cancer Res.* 2004; 10:5160–5167. [PubMed: 15297420]
51. Yao J, Weremowicz S, Feng B, et al. Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression. *Cancer Res.* 2006; 66:4065–4078. [PubMed: 16618726]
52. Miller FR, Santner SJ, Tait L, et al. MCF10DCIS.com xenograft model of human comedo ductal carcinoma in situ. *J Natl Cancer Inst.* 2000; 92:1185–1186. [PubMed: 10904098]

Figure 1. DCIS and adjacent invasive breast carcinomas have similar patterns of gene copy number aberrations.

(A) Frequency plot of copy number gains and losses (top) or amplifications (bottom) in 13 matched *in situ* and adjacent invasive breast carcinomas. The proportion of tumours in which each bacterial artificial chromosome (BAC) clone is gained (green bars) or lost (red bars) is plotted (Y-axis) for each BAC clone according to its genomic position (X-axis). No significant differences were identified between the two components. (B) Hierarchical cluster analysis performed with microarray comparative genomic hybridisation (aCGH) cbs-smoothed ratios using a correlation distance metric and the Ward's algorithm of the 13

matched samples. The components from each pair of synchronously diagnosed matched DCIS and invasive breast cancer preferentially clustered together. (C) Representative micrographs of the DCIS and its adjacent invasive component from the frozen section of Case 9 subjected to microdissection (left) and their respective genome plots (right). In the genome plots, the genomic position is plotted along the X-axis and cbs-smoothed Log_2 ratio on the Y-axis; amplifications are shown in bright green, gains in dark green, losses in dark red, and normal copy number in black. Amp: amplification; ER: oestrogen receptor; Gain: copy number gain; Loss: copy number loss; NC: no copy number change; PR: progesterone receptor.

Figure 2. Pair-wise analysis of matched DCIS and adjacent invasive breast cancer.

Representative micrographs of the DCIS and its adjacent invasive breast cancer component from frozen sections of (A) Case 2, (B) Case 4, and (C) Case 5 and their respective genome plots. Genomic position is plotted along the X-axis and cbs-smoothed Log₂ ratio on the Y-axis, amplifications are shown in bright green, gains in dark green, losses in dark red, and normal copy number in black. A red star denotes differences in the copy number profiles between matched DCIS and invasive components.

Figure 3. Fluorescence *in situ* hybridisation (FISH) demonstrates intra-tumour genetic heterogeneity and genetic differences between matched DCIS and adjacent invasive breast carcinomas.

FISH was performed with in-house probes for genetic loci displaying differences in copy number as defined by microarray-based comparative genomic hybridisation in three cases (A-F). For each locus, a representative micrograph of each component, a chromosome plot denoting the region probed (marked by a red star in the component with the highest Log_2 ratio), representative FISH images and a bar plot showing the proportions of nuclei containing 1, 2-4 or 5+ copies (at least 100 nuclei were scored in each component) are shown. In the chromosome plots, genomic position is plotted along the X-axis and cbs-smoothed Log_2 ratio on the Y-axis, amplification and gains are shown in bright and dark green, while losses are shown in dark red, and normal copy number in black. FISH images show nuclei counterstained with DAPI, and FISH probes labelled either with Digoxigenin (green signals) or Biotin (red signals). Inset images of single amplified cells are depicted in the bottom right corner of each FISH image. Bar graphs contain red bars denoting regions with 5+ copies, grey bars with 2-4 copies and black bars with 1 copy. Significant differences were determined using the Chi-square test.

Figure 4. Mutational profiling using the Sequenom MassARRAY platform identifies differences in the proportion of cells harbouring *PIK3CA* mutations between *in situ* and invasive breast cancer components.

Using the Oncocarta v1.0 panel on the Sequenom platform, mutations in *PIK3CA* were identified in 6 cases. (A) In three of these cases, the mutations were present at equal frequency in both components (Cases 3, 8 and 10). In the remaining three cases, the mutation was either present at lower frequency in the invasive component (Case 2) or absent in the invasive component (Cases 12 and 5) (B and C). Sanger sequencing was used to validate these findings. In each panel, the mass spectrometry profile is shown demonstrating the allele called, with the corresponding Sanger sequencing trace below, centred on the mutation site (red arrow). The top case in each panel is one with wild-type *PIK3CA* (i.e. Cases 11, 13 and 6).

Table 1

Clinicopathological features of matched DCIS and invasive breast cancer samples.

Case	Component	Subtype	Histological grade	Nuclear grade	ER	PR	HER2	Ki67	p53	TNM stage	Age
Case 1	DCIS	Solid	2	3	+	+	-	Low	-	T2N1	65
	Invasive	IDC-NST		3	+	+	-	Low	-		
Case 2	DCIS	Solid	3	3	-	-	+	High	+	T2N0	56
	Invasive	IDC-NST		3	-	-	+	High	+		
Case 3	DCIS	Solid	3	3	+	+	-	Low	+	T4N0	78
	Invasive	IDC-NST		3	+	+	-	Low	+		
Case 4	DCIS	Solid	3	3	+	-	+	Moderate	-	T2N3	66
	Invasive	IDC-NST		3	+	+	+	Moderate	-		
Case 5	DCIS	Solid	3	3	-	-	+	Moderate	-	T2N1	86
	Invasive	IDC-NST		3	-	-	+	High	-		
Case 6	DCIS	Cribriform	3	2	+	+	-	Low	+	T1N0	71
	Invasive	IDC-NST		2	+	+	-	Low	+		
Case 7	DCIS	Solid	3	3	+	+	-	Low	-	T1N0	74
	Invasive	IDC-NST		3	+	+	-	Low	-		
Case 8	DCIS	Cribriform	2	2	+	+	-	Low	+	T1N0	46
	Invasive	IDC-NST		2	+	+	-	Moderate	+		
Case 9	DCIS	Solid	3	3	+	+	-	Low	+	T1N0	67
	Invasive	IDC-NST		3	+	+	-	Moderate	+		
Case 10	DCIS	Solid	3	2	+	+	-	Moderate	+	T3N1	42
	Invasive	IDC-NST		2	+	+	-	High	+		
Case 11	DCIS	Solid	3	3	+	+	-	Low	+	T2N0	46
	Invasive	IDC-NST		3	+	+	-	Moderate	+		
Case 12	DCIS	Solid	2	2	+	+	-	Low	-	T3N0	76

Case	Component	Subtype	Histological grade	Nuclear grade	ER	PR	HER2	Ki67	p53	TNM stage	Age
	Invasive	Neuroendocrine		2	+	+	-	Low	-		
Case 13	DCIS	Solid	3	3	+	+	-	Moderate	+	T2N0	54
	Invasive	IDC-NST		2	+	+	-	Moderate	+		

ER – oestrogen receptor, PR – progesterone receptor, DCIS – ductal carcinoma *in situ*, IDC-NST – invasive ductal carcinoma of no special type, Ki67 expression – Low: <10%, Moderate: 10-30%, High: >30%, Neuroendocrine – neuroendocrine carcinoma.