Oncogene. Author manuscript; available in PMC 2009 May 28.

Published in final edited form as:
Oncogene. 2006 October 19; 25(49): 6544–6553. doi:10.1038/sj.onc.1209659.

A comprehensive study of chromosome 16q in invasive ductal and lobular breast carcinoma using array CGH

R Roylance[1], **P Gorman**[1], **T Papior**[1], **Y-L Wan**[1], **M Ives**[1], **JE Watson**[2,8], **C Collins**[2], **N Wortham**[1], **C Langford**[3], **H Fiegler**[3], **N Carter**[3], **C Gillett**[4], **P Sasieni**[5], **S Pinder**[6], **A Hanby**[7], and **I Tomlinson**[1]

[1]Molecular and Population Genetics Laboratory, Cancer Research UK, Lincoln's Inn Fields, London, UK

[2]Collins's Lab S151, UCSF Cancer Center, San Francisco, CA, USA

[3]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

[4]Hedley Atkins/CR-UK Breast Pathology Laboratory, Guy's Hospital, London, UK

[5]Mathematics, Statistics and Epidemiology, Bart's and the London Medical School, Charterhouse Square, London, UK

[6]Department of Histopathology, Addenbrooke's Hospital, Cambridge, UK

[7]Academic Unit of Pathology, University of Leeds, c/o Department of Histopathology, St James' University Hospital, Leeds, UK

Abstract

We analysed chromosome 16q in 106 breast cancers using tiling-path array-comparative genomic hybridization (aCGH). About 80% of ductal cancers (IDCs) and all lobular cancers (ILCs) lost at least part of 16q. Grade I (GI) IDCs and ILCs often lost the whole chromosome arm. Grade II (GII) and grade III (GIII) IDCs showed less frequent whole-arm loss, but often had complex changes, typically small regions of gain together with larger regions of loss. The boundaries of gains/losses tended to cluster, common sites being 54.5–55.5 Mb and 57.4–58.8 Mb. Overall, the peak frequency of loss (83% cancers) occurred at 61.9–62.9 Mb. We also found several 'minimal' regions of loss/gain. However, no mutations in candidate genes (*TRADD*, *CDH5*, *CDH8* and *CDH11*) were detected. Cluster analysis based on copy number changes identified a large group of cancers that had lost most of 16q, and two smaller groups (one with few changes, one with a tendency to show copy number gain). Although all morphological types occurred in each cluster group, IDCs (especially GII/GIII) were relatively over-represented in the smaller groups. Cluster groups were not independently associated with survival. Use of tiling-path aCGH prompted re-evaluation of the hypothetical pathways of breast carcinogenesis. ILCs have the simplest changes on 16q and probably diverge from the IDC lineage close to the stage of 16q loss. Higher-grade IDCs probably develop from low-grade lesions in most cases, but there remains evidence that some GII/GIII IDCs arise without a GI precursor.

Correspondence: Dr R Roylance, Molecular and Population Genetics Laboratory, Cancer Research UK, Lincoln's Inn Fields, London WC2A 3PX, UK. E-mail: rebecca.roylance@cancer.org.uk.
[8]Current address: Amgen Inc., 1120 Veterans Blvd, South San Francisco, CA 94080, USA.

Europe PMC Funders Author Manuscripts

Keywords

array CGH; Chromosome 169; breast cancer

Europe PMC Funders Author Manuscripts

Introduction

Loss of the long arm of chromosome 16 is a very common finding in low-grade (grade I or GI) infiltrating ductal carcinomas (IDCs) of the breast, and is reported to occur more commonly than in higher grade (grades II and III (GII and GIII)) ductal cancers (Buerger *et al.*, 1999; Roylance *et al.*, 1999). On the basis that the changes of chromosome 16q target a tumour suppressor gene, there have been many published studies attempting to identify small regions of allele loss (loss of heterozygosity (LOH)). These studies have generally used microsatellite-based analyses and as a result, attention has focused on putative minimal regions of LOH at 16q22.1, 16q23.2–24.1 and 16q24.3 (Dorion-Bonnet *et al.*, 1995; Whitmore *et al.*, 1998; Frengen *et al.*, 2000). Most of these studies have involved mixed populations of invasive breast cancers of various types and grades. As a result of LOH mapping, a number of genes have been screened for mutations (for e.g., Whitmore *et al.*, 1998; Cleton-Jansen *et al.*, 1999; Crawford *et al.*, 1999), but the identity of the 16q gene(s) involved in low-grade breast cancer has remained elusive.

Infiltrating lobular breast carcinomas (ILCs) share a number of gross genetic similarities with GI IDCs, namely a high frequency of loss of 16q and gain of 1q (Nishizaki *et al.*, 1997; Roylance *et al.*, 1999; Hwang *et al.*, 2004). The 16q gene involved in a majority of lobular cancers is known to be E-cadherin (*CDH1*), which maps to 16q22.1 (Berx *et al.*, 1995, 1996) and is inactivated by allelic loss and/or mutation and/or methylation (Droufakou *et al.*, 2001). In view of the genetic – although not morphological - similarities between the lobular and GI IDC subtypes, we have previously screened a cohort of GI tumours to look for inactivation of *CDH1*. However, although some reduced expression was found together with LOH at the *CDH1* locus, we found no mutations or loss of expression, thus suggesting that *CDH1* is not involved in the development of GI IDCs (Roylance *et al.*, 2003). It has therefore been postulated, both by us (Roylance *et al.*, 1999, 2003) and others (Cleton-Jansen, 2002), that GI IDCs develop through mutation of an unknown gene on 16q, followed by LOH, with subsequent mutation or silencing of *CDH1* giving rise to the lobular phenotype.

Although the high frequency of loss of 16q has been well established for both GI IDCs and ILCs, the true frequency of loss in higher grade ductal breast carcinomas has remained more uncertain, with conflicting data. In GIII cancers, conventional comparative genomic hybridization (CGH) has found deletion of 16q in less than 20% of tumours (Buerger *et al.*, 1999; Roylance *et al.*, 1999), whereas allelic loss has been found at a higher frequency (Tsuda *et al.*, 1994; Dorion-Bonnet *et al.*, 1995). In a small study comparing CGH with LOH assessed using microsatellites, the latter occurred at a higher frequency, but there was still a significant difference between GI and GIII IDCs (Roylance *et al.*, 2002). The different frequencies of loss of 16q seen in different grades of IDC led to the hypothesis that generally GI IDCs do not progress to higher grade lesions (Buerger *et al.*, 1999; Roylance *et al.*, 1999). However, there remained a small proportion of GIII tumours with 16q loss that may have arisen by de-differentiation through grade (Korsching *et al.*, 2004).

In contrast, some LOH analyses have failed to find any differences between the frequencies of 16q loss in the different grades (Tsuda *et al.*, 1994; Cleton-Jansen *et al.*, 2001). A recent study examined loss of 16q in breast cancer using different techniques (Cleton-Jansen *et al.*, 2004); a high frequency of loss was found by fluorescence *in situ* hybridization (FISH) and

conventional CGH only in low-grade tumours, but there were similar frequencies of loss in both grades as assessed by LOH analysis. The authors concluded that the most likely explanation for these findings was different mechanisms of 16q loss, that is, physical loss (detectable by both CGH and LOH) in GI IDCs and mitotic recombination (only detectable by LOH) in GIII tumours.

To examine the nature of 16q changes in invasive breast cancer, we have analysed 106 breast cancers using a 16q-specific CGH array, with near-contiguous coverage of the entire chromosome arm. Our aims were several fold: firstly, in GI IDCs to identify, confirm and refine regions of loss, some of which had already been identified by LOH analysis, for the purpose of mapping tumour suppressor gene(s); secondly, to determine whether any ILCs showed minimal regions of deletion which would suggest the involvement of genes other than *CDH1* in their development; and thirdly, to determine the frequency of physical loss of 16q in higher grade ductal tumours (GII/GIII) using a sensitive technique, with the aim of reconciling the differences seen between conventional CGH and allele loss analysis. By comparing patterns of loss on 16q in the different morphological subtypes, we aimed to provide data that were informative regarding the postulated genetic pathways in breast tumorigenesis.

Results

aCGH overall results

Forty GI IDCs, 30 ILCs, 19 GII IDCs and 17 GIII IDCs were analysed using the 16q-specific CGH array. Examples of aCGH profiles are shown in Figure 1. We found apparently erroneously reporting clones at 73 Mb (Watson *et al.*, 2004) and close to the copy number polymorphism at about 68.7 Mb on chromosome 16q, which has previously been reported (Sharp *et al.*, 2005). These changes involved few individual data points within larger regions of gain or loss and those clones were therefore simply disregarded from the point of view of reporting copy number change in the cancers.

aCGH on GI IDCs

Twenty-three (58%) of 40 GI IDCs showed loss of the whole of chromosome 16q, one cancer (T1443) showed gain of the whole arm and four cancers (T1272, T1097, T455, T1547) showed no changes. Of the other 12 GI cancers (Figure 2), one (T175) showed only a region of gain towards the 16q telomere. Two cancers had interstitial regions of loss at 56–63 Mb (T843) and 57.8–61.0 Mb (T1093); a further tumour (T1411) had interstitial loss at 57.5–62.9 Mb and had gained the rest of the chromosome arm. The remaining eight GI tumours showed patterns of alternating loss and gain, with some tumours showing four or more distinct regions of change. Gain in the region around 53 Mb was commonly observed (Figure 3), but cancers tended to lose material distal to about 56 Mb, with at least 30 (75%) of the 40 GI tumours showing loss in this region (Figure 4).

aCGH on ILCs

Nineteen (63%) of 30 ILCs lost the whole chromosome arm. Of the 11 other ILCs (Figure 2), four showed predominant loss combined with a single area of no change, and six showed loss of most clones, but with smaller areas of gain or no change; a single cancer (T865) showed approximately equal numbers of gained and lost clones. At all sites distal to about 57.5 Mb, at least 28 (93%) of the 30 ILCs showed loss (Figure 4). All of the ILCs showed loss within four regions, 58.8–63.9, 74.8–75.4, 77.5–79.3 and 80.7–82 Mb. Only two cancers failed to show loss at *CDH1* (67.3 Mb). Morphological review of the lobular tumours showed that no particular subtype was associated with a particular genetic pattern of aberrations.

GII IDCs

Eight (42%) of 19 GII IDCs lost the whole chromosome arm, one (T1218) gained the whole arm and three showed no changes. Three cancers showed alternating regions of loss and no change (Figure 2), including one tumour (T1039) with interstitial loss at 57.4–61.1 Mb and another (T1333) with loss of the whole arm apart from an interstitial region of no change at 67.1–75.0 Mb. A further cancer (T1314) showed gain only from the centromere to 59.8 Mb. The remaining three cancers showed mixed loss and gain. Distal to 57.3 Mb, at least 11 (58%) of the GII IDCs showed loss at each clone (Figure 4). All but five GII tumours had loss at 57.5–61.1 Mb.

GIII IDCs

Only two (T1055, T1103) of 17 (12%) GIII IDCs had lost the entire chromosome arm; two cancers (T1067, T1146) showed gain of the whole arm and one (T959) had no changes. Two cancers (T801, T1406) had single regions of gain and two (T1010, T1074) had single regions of loss (Figure 2). The remaining eight (47%) tumours had mixed loss and gain, with up to seven separate regions of change. Alone among the four types of cancer, the GIII IDCs showed a distinct peak in the frequency of loss (12 cancers, 71%) between 57.0 and 64.7 Mb. Distal to 57 Mb, at least seven (41%) GIII cancers showed loss at each marker (Figure 4).

Comparison between tumour types

Overall, aCGH showed that both the GI IDCs and ILCs showed frequent deletion of all or part of 16q (Figure 4), with typical frequencies at any site of 75 and 95%, respectively. Gains were rarer (Figure 3), but were still unexpectedly common (up to 20% of all cancers at some sites, with gains especially common in GIII IDCs occurring in 53% at 51.9–53.36 Mb). GII and GIII IDCs showed lower levels of loss overall (typically 65% and 50%, respectively; Figure 4), owing to less frequent deletion of the whole chromosome arm (χ^2_2 trend=10.2, P=0.006). Deletions and gains involving part of the arm were, however, similarly common in all of the cancer types (Table 1). No homozygous deletions were detected in any cancer. Only two cancers (T914, lobular and T1151, GIII) had evidence of gene amplification: T914 (lobular) had a maximum $\log_2$ T:N ratio of 1.02 at 47.4 Mb, but this region contains no known genes, Ensembl genes or human mRNAs, and T1151 (GIII) had a maximum $\log_2$ T:N ratio of 1.37 at and around 52.2 Mb, a region containing about 10 known and predicted genes (http://genome.ucsc.edu/), but no known oncogenes.

For tumours with mixed regions of loss and gain, there was a median of three regions of change per cancer (inter-quartile range=2–4) and (for cancers which showed such a ‘mixed’ pattern) there was no tendency for the changes in the higher grade cancers to be more complex and numerous than those in the lower grade lesions (Kruskal–Wallis test, χ^2_3=2.64, P=0.45). Hierarchical cluster analysis showed that those cancers with loss of the whole chromosome arm consistently clustered into group one (78/106 cancers), those with gain of the whole arm (7/106) into group two and those with no change (21/106) into group three (Figure 5). There was an association between group and cancer morphology, with proportionately more GII and GIII IDCs (16/36) in the groups with a tendency to gain or no change, and almost all (28/30) ILCs clustered into group one (P=0.01, Fisher's exact test). Interestingly, of 42 cancers with mixed loss and gain, 26 clustered with the whole-arm loss group, three clustered with the whole-arm gain group and the remaining 13 were in the no change group. Thus, most of the cancers with mixed gains and losses resembled most closely the cancers which lost the whole chromosome arm.

‘Minimal regions’ of change by aCGH

We have described above for each cancer type several regions of particularly frequent loss or gain. Small regions of change found in individual tumours are shown in Table 2. Ten regions of gain of about 3 Mb or less were found, some in more than one cancer. Given that no high level gains were found in these regions, the changes observed were unlikely to result from amplification of oncogenes. However, given that aCGH reports amplification of a small region in proportion to its overlap with the large insert clone within which the amplicon occurs, and that loss on one or more copies of 16q may partly obscure gains on another homologue, it remains possible that oncogene amplification had occurred in some cases. The known and predicted genes within each discrete region of gain are shown in Table 2. Relatively small, discrete regions of loss were less common and gene-poorer than the discrete gains. Five regions of discrete loss were found, each in only one cancer. Known and predicted genes within each region are shown in Table 2.

Two GI IDCs showed regions of loss at 57.8–61.0 Mb (T1093) and 61.3–66.9 Mb (T715), respectively (Figure 2). The former region is gene-poor proximally and contains only the *CDH8* gene, the function of which is not well described. The latter region contains several potential candidate tumour suppressor loci, including the death-domain gene *TRADD*, and cadherins *CDH5* and *CDH11*. We screened all four genes for mutations in a panel of GI tumours, but no pathogenic changes were found.

When all the data were combined, the ‘minimal’ regions did not suggest a single site, which harboured an uncharacterized tumour suppressor gene or oncogene (details not shown). Given the complex changes found in many cancers, this was not surprising. It is not possible, for example, using aCGH to separate out the changes that occurred on the two (or more) copies of 16q in each cancer. We therefore determined the frequency of loss or gain at each clone in each type of cancer and in the entire tumour set. For gains (Figure 3), the frequency of change showed a similar pattern along the chromosome arm – that is, peaks and troughs of frequency – in all cancer types ($P<0.001$ for all pairwise regression analyses, details not shown). A peak frequency of over 20 gains (19% of cancers) was found between 51.5 and 54.0 Mb. Within this region lie a cluster of iroquois homeobox genes and matrix metalloproteinase 2.

Although the frequency of loss varied relatively little along the chromosome – owing to the predominance of cancers with loss of the whole chromosome arm – the highest frequency of loss (in about 75% cancers overall) occurred between 57.4 and 64.7 Mb (Figure 4); within this region, the peak frequency of loss was 83% between 61.9 and 62.9 Mb. There are no known candidate tumour suppressor genes within the latter region. As was the case for the gains, all tumour types showed the same peaks and troughs of loss frequencies along the chromosome arm; this included the lobular cancers, which had a maximum frequency of loss at the same site as the other tumour types, at 60 Mb, over 5 Mb proximal to *CDH1* (67.3 Mb).

Breakpoints

For cancers that did not lose or gain the whole chromosome arm, we examined regions of gain or loss for consistent boundaries, presumed to result originally from a chromosome breakpoint or similar event, which gave rise to the gain or loss observed. This analysis must be hedged with the *caveat* that not all apparent breakpoints are real (consider, for example, a region of single-copy loss on one 16q homologue and a partially overlapping region of single-copy gain on the other). Nevertheless, we scored each boundary of a region of loss or gain as a ‘breakpoint’ and analysed clustering of breakpoints by grouping into bins of 10 adjacent clones (approximately 1 Mb). The distribution of breakpoint frequency differed

significantly from the expected Poisson distribution (P=0.002), demonstrating nonrandom breakpoint clustering. Two locations (54.5–55.5 Mb and 57.4–58.8 Mb) were particularly common breakpoint sites, each occurring in nine (8.5%) cancers. The first of these regions is just distal to a low copy number repeat containing members of the carboxylesterase gene family. The second is gene-poor and has no apparent special features.

Clinico-pathological–molecular associations

We tested for associations between the three cluster groups of cancers (based on 16q losses and gains) and the following clinico-pathological variables: morphology and grade (as above), age at presentation, menopausal status (pre-, peri- or post-), size of primary tumour, node status (0 or 1), Oestrogen receptor (ER) expression (0 or 1) (Barnes *et al.*, 2004), progesterone receptor (PR) expression (0 or 1) (Cooper *et al.*, 1999), time to relapse and overall survival. Median follow-up was 114 months (range 7–196 months). Although the three cluster groups were associated with differential survival (Figure 6), this was not independent of ER status. Further analysis showed that the tumours with poor survival tended to be ER-, GII and GIII IDCs, largely as expected. The association of survival with cluster group probably arose as a secondary effect of the association between cluster group and morphology, and hence, ER status. There was no other association between the three cluster groups and any of the other clinico-pathological variables (details not shown).

Discussion

Our aCGH data has provided the most comprehensive analysis to date of the changes on 16q in different morphological subtypes of invasive breast cancer. Almost all ILCs show loss of chromosome 16 distal to about 58.8 Mb, as do about three-quarters of the GI IDCs. These frequencies are much higher than reported previously (Nishizaki *et al.*, 1997; Buerger *et al.*, 1999; Roylance *et al.*, 1999), and involve much larger regions than expected. The demonstration of such a high prevalence of loss in predominantly diploid tumours (Pandis *et al.*, 1996) with few other large genetic aberrations (Roylance *et al.*, 1999) supports the hypothesis that loss of 16q is an important change in both these morphological subtypes. In GII and GIII IDCs, loss of 16q was typically found at each clone in one-half to two-thirds of tumours, higher than expected from previous CGH studies (Buerger *et al.*, 1999; Roylance *et al.*, 1999), but significantly lower than in the GI lesions and ILCs. Smaller regions of 16q change, and multiple regions of change per tumour, were found in all cancer types, but were more frequent in the higher grade IDCs. The identification of these complex changes and gain of material has not been reported using other, less sensitive methods (for e.g., Hwang *et al.*, 2004). It is possible that the complex changes were the result of segmental gains subsequent to large-scale losses. The significance of the regions of gain is unclear, but the multiple changes on 16q in some tumours do suggest inherent instability, whether global or limited to 16q.

On the basis of cancers with small, discrete changes and the locations of the maximum frequencies of loss or gain, we have identified several regions that may harbour the elusive 16q tumour suppressor, or even new oncogenes. Given that many small regions of gain overlapped with regions of loss in other tumours – and *vice versa* – it seems likely that many of the small changes detected were structural or secondary rather than primarily targeting genes involved in tumorigenesis. Furthermore, the fact that gains and losses were often contiguous, rather than separated by regions of no change, also suggests that many changes resulted from some form of local instability within the chromosome arm. This contention is supported by the nonrandom occurrence of ‘breakpoints’ flanking gains and losses. These breakpoint events may be the initial event which is being selected for. It follows that identifying genes targeted by deletions on 16q will be problematic. It is anticipated that as so many tumours lose the entire arm, expression analysis may be helpful in elucidating the

gene(s) involved. Otherwise, the necessary alternative approach of mutation/methylation screening of all candidate genes will be a very time-consuming effort. In addition, with so many tumours losing the whole arm, we cannot exclude the possibility that haploinsufficiency plays a role in the pathogenesis of some cancers (Quon and Berns, 2001; Santarosa and Ashworth, 2004).

The increased frequency of partial or whole arm loss compared with previous reports was unexpected, but is highly unlikely to be artefactual. Complete loss of 16q in 100% of ILCs has been recently reported using 1 Mb aCGH (Hwang *et al.*, 2004). Previous work reporting the lower frequency of 16q loss has used either conventional CGH or microsatellite-derived allele loss analysis. The former is not as sensitive as aCGH (for e.g., Douglas *et al.*, 2004); the latter is prone to false-positives and -negatives (Devilee *et al.*, 2001; Kern, 2002; Tomlinson *et al.*, 2002), cannot distinguish gain from deletion and cannot detect 'symmetrical' changes affecting both alleles. Comparing our previous conventional CGH data (Roylance *et al.*, 1999) with the current work, there was good agreement where changes were large and simple. However, as might have been expected, aCGH had markedly greater ability to detect smaller regions of change, to detect losses and gains involving adjacent regions and to detect changes of apparently lower copy number (probably owing to contaminating normal cells within the tumour).

Using microsatellites, Cleton-Jansen *et al.* (2004) reported that 16q LOH in higher grade breast cancers occurred about as frequently as in GI IDCs, but that deletions (as assessed by conventional CGH or FISH) were much less frequent in the former, suggesting that the mechanism of LOH in higher grade tumours was mitotic recombination. Our data based on the higher sensitivity of tiling-path aCGH suggest the following explanation. Notwithstanding the inherent unreliability of microsatellite-based LOH analysis – a view which Cleton-Jansen and we share (Devilee *et al.*, 2001; Tomlinson *et al.*, 2002), our aCGH data show that 16q deletion in GII/GIII lesions is actually frequent, but it rarely involves the whole arm and commonly occurs together with gain. The copy number changes in GII/GIII tumours frequently involve relatively small regions below the resolution of conventional CGH and methods such as FISH, unless performed at high density. Overall, our data show that the copy number changes on 16q in GII/GIII tumours are only slightly less common than in GI IDCs and ILCs (Table 1). Therefore, there is no need to invoke mitotic recombination in GII/GIII cancers, because the gap between the frequencies of changes found by aCGH and LOH is both similar and small.

Is it possible to make any inferences regarding genetic pathways of tumorigenesis by comparing patterns of 16q loss in different types of breast cancer? Conventional CGH suggested that 16q deletions were much less common in GII/GIII IDCs than in GI IDCs and ILCs. This, together with morphological data (Millis *et al.*, 1998), led to the hypothesis that for the majority of breast tumours there were parallel pathways of tumour development, with GI tumours having a separate pathway from the higher grade tumours (Buerger *et al.*, 1999; Roylance *et al.*, 1999). For a minority of high-grade tumours with 16q loss, a progression from low-grade tumours was thought likely and mathematical modelling seemed to support this hypothesis (Korsching *et al.*, 2004). It was further hypothesized, from the pattern of genetic changes in GI IDCs and ILCs, that ILCs developed along a similar pathway to GI IDCs with loss of E-cadherin leading to the ILC phenotype (Roylance *et al.*, 1999, 2003; Cleton-Jansen, 2002). Using the 16q-specific array, although we still found a differential loss of 16q, there was a higher frequency of changes on 16q in the higher grade lesions than expected. Furthermore, there were more complex changes seen in all cancers, but especially in the higher grade ductal lesions. Interestingly, the peaks and troughs of the frequencies of gains and losses tended to occur at the same sites in each of the four tumour types. Cluster analysis showed that most cancers with predominant loss on 16q clustered into a single

group, but that, in the other two groups, ILCs were under-represented compared with IDCs. We suspect, therefore, that the relationship between different morphological types of breast cancer is complex and that the previous hypothetical model needs refining. Our data remain consistent with a model in which ILCs and GI IDCs have a common progenitor and in which the lineages diverge at or soon after 16q loss owing to E-cadherin mutation or silencing causing the lobular phenotype to develop. However, unlike in the previous model, our aCGH data are more consistent with a significant number of IDCs showing progression through grades, with subsequent accumulation of segmental gains (and, perhaps, further losses) in the higher grade lesions. The fact that aCGH continues to find a lower frequency of any 16q loss in higher grade IDCs compared to GI tumours continues to support the hypothesis that some GII/GIII IDCs develop along a pathogenic pathway which does not involve a precursor stage as a GI cancer.

Materials and methods

Tissue samples

Tumours were collected from breast cancer patients treated by the Breast Unit at Guy's Hospital, London between 1988 and 2000. All information regarding the tumours was recorded by the Guy's Breast Pathology Laboratory. IDCs were graded using the Nottingham modified criteria of Scarff, Bloom and Richardson (Elston and Ellis, 1991). The ILCs were further subdivided according to their morphology into classical, pleomorphic and mixed types. Samples were reviewed by two pathologists (AH and SP) to confirm morphological subtype and tumour grade and to determine the proportion of admixed normal cells; all tumours had greater than 60% cancer cells. DNA was extracted from frozen tumour, normal lymphoblastoid cell lines and blood using standard methods. The work was approved by Guy's Hospital Local Research Ethics Committee.

BAC array construction

A minimal overlapping tiling set of 433 clones (bacterial artificial chromosomes (BACs), P1-derived artificial chromosome (PACs) and cosmids) spanning the entire long arm of chromosome 16 (16q11.2–16q24.3) was selected, together with an additional 32 clones mapping to the short arm of chromosome 16. Three hundred and twenty-six of these clones formed the basis of a previously described region-specific array (Watson *et al.*, 2004). Additional clones were added to improve coverage, which was 95% complete, the largest gap being 417 kb between 86.6 and 87 Mb. The location of the clones was confirmed using the University of California at Santa Cruz genome browser (http://www.genome.ucsc.edu) May 2004 freeze. A set of 256 large insert clones that mapped randomly to 20 different chromosomes was used as internal controls to detect hybridization artefacts and to allow normalization. Six *Drosophila* clones were also included on the array to control for non-specific hybridization. DNA from each clone was amplified using degenerate oligonuclotide-primed (DOP)-polymerase chain reaction (PCR) as described previously (Fiegler *et al.*, 2003). Following the initial DOP-PCR, a second amino-linking PCR was performed and the clones were then spotted in duplicate or triplicate onto amine binding slides (3-D link activated slides, GE Healthcare Amersham, UK) using a MicroGrid II arrayer (BioRobotics, Cambridge, UK). The clones were arranged to ensure that adjacent clones on the tiling path were not adjacent on the array, and the clones from the rest of the genome were randomly distributed throughout the array.

Genomic DNA labelling, array hybridization and image analysis

These techniques were essentially performed as described by Fiegler *et al.* (2003). Briefly, 350 ng of tumour and reference DNA (pooled normal female DNA from lymphoblastoid cell lines) were labelled with Cy5 and Cy3, respectively, by random priming. The quality of the

labelled DNAs (test/control) was assessed using a 2% agarose gel before hybridizing to the arrays and washing. The array was scanned using a confocal scanner (GSI Lumonics, Perkin Elmer, Beaconsfield, UK) and 'Spot' software (Jain *et al.*, 2002) was used to process the images. After rejecting poorly hybridized clones and correcting for background, the $\log_2$ ratio of the fluorescence intensities of test (tumour, T) to reference (normal, N) was calculated, after normalization to the remaining non-chromosome 16 clones in the block.

Data analysis

A series of nine normal versus normal hybridizations was used to define variability for each clone. Clones that showed large variation (T:N $\log_2$ ratio ±0.2) in four or more hybridizations were excluded from the remainder of the analysis. In addition, clones found on test hybridizations to behave aberrantly compared to adjacent clones were analysed using FISH; almost all of these were found to map to locations additional to or other than that expected on 16q and they were excluded. A T:N $\log_2$ ratio of zero implied no copy number change at that clone, a significantly increased positive value indicated an increase in copy number and a significantly reduced negative value, a decrease in copy number – all of these changes being relative to the rest of the genome in that tumour. For formal assessment of copy number changes on 16q, independent observers (PG, IT) identified continuous regions of gain or loss. We then used a Student's *t*-test with threshold $P<0.001$ to determine whether the dosage in each region differed significantly from the normal:normal hybridizations. We formally required three adjacent clones to report the same copy number change for a region to be considered as gained or lost.

We initially tested that the array could detect copy number changes using the colorectal cancer cell line HCT116, which harbours a homozygous deletion of *WWOX* at 16q23.1 (details not shown). The breast cancer cell line MPE-600 was then used to test the sensitivity of the array. It is a diploid cell line with known loss of an entire copy of 16q and very few other large-scale changes. We used MPE-600 DNA 'spiked' with increasing amounts of normal DNA to determine the sensitivity of the aCGH. We found that with up to 67% normal DNA present, it was still possible to identify 16q loss (details not shown). In order to exclude artefactual gain or loss on 16q resulting from dosage changes elsewhere in the genome, we examined the profiles of each tumour using data from a genome-wide 1 Mb density CGH array, to check that 16q clones reported deletion or gain concordant with the 16q-specific array, whereas other chromosomes showed $\log_2$ ratios of zero. In all cases, the 16q changes were confirmed as genuine (details available from authors).

Hierarchical cluster analysis

For every tumour, each clone was treated as one data point and encoded as follows: gain (+1), no change (0) or loss (−1). Hierarchical cluster analysis using the STATA 'cluster averagelinkage' and 'cluster generate' commands was used to place all cancers into three groups (termed groups one, two and three). The default parameters for continuous data (the Euclidean/L2 distance) were used (http://www.stata.com).

Mutation analysis

Four candidate genes, *TRADD*, *CDH5*, *CDH8* and *CDH11*, were screened for mutations. All exons for each gene were PCR amplified (details available from authors) and analysed using either single-strand conformational polymorphism analysis or denaturing high-performance liquid chromatography depending upon the size of the amplified product as described previously. PCR products displaying bandshifts were sequenced in duplicate on an ABI Prism 3100 sequencer (PE Biosystems, Warrington, UK).

Acknowledgments

We are very grateful to colleagues at the Sanger Institute Microarray Facility for array printing.

References

Barnes DM, Millis RR, Gillett CE, Ryder K, Skilton D, Fentiman IS, et al. Endocr Relat Cancer. 2004; 11:85–96. [PubMed: 15027887]

Berx G, Cleton-Jansen AM, Nollet F, de Leeuw WJ, van de Vijver M, Cornelisse C, et al. EMBO J. 1995; 14:6107–6115. [PubMed: 8557030]

Berx G, Cleton-Jansen AM, Strumane K, de Leeuw WJ, Nollet F, van Roy F, et al. Oncogene. 1996; 13:1919–1925. [PubMed: 8934538]

Buerger H, Otterbach F, Simon R, Schafer KL, Poremba C, Diallo R, et al. J Pathol. 1999; 189:521–526. [PubMed: 10629552]

Cleton-Jansen AM. Breast Cancer Res. 2002; 4:5–8. [PubMed: 11879552]

Cleton-Jansen AM, Buerger H, Haar N, Philippo K, van de Vijver MJ, Boecker W, et al. Genes Chromosomes Cancer. 2004; 41:109–116. [PubMed: 15287023]

Cleton-Jansen AM, Callen DF, Seshadri R, Goldup S, McCallum B, Crawford J, et al. Cancer Res. 2001; 61:1171–1177. [PubMed: 11221848]

Cleton-Jansen AM, Moerland EW, Pronk JC, van Berkel CG, Apostolou S, Crawford J, et al. Br J Cancer. 1999; 79:1049–1052. [PubMed: 10098735]

Cooper LS, Gillett CE, Patel NK, Barnes DM, Fentiman IS. Cancer. 1999; 86:2053–2058. [PubMed: 10570431]

Crawford J, Ianzano L, Savino M, Whitmore S, Cleton-Jansen AM, Settasatian C, et al. Genomics. 1999; 56:90–97. [PubMed: 10036189]

Devilee P, Cleton-Jansen AM, Cornelisse CJ. Trends Genet. 2001; 17:569–573. [PubMed: 11585662]

Dorion-Bonnet F, Mautalen S, Hostein I, Longy M. Genes Chromosomes Cancer. 1995; 14:171–181. [PubMed: 8589033]

Douglas EJ, Fiegler H, Rowan A, Halford S, Bicknell DC, Bodmer W, et al. Cancer Res. 2004; 64:4817–4825. [PubMed: 15256451]

Droufakou S, Deshmane V, Roylance R, Hanby A, Tomlinson I, Hart IR. Int J Cancer. 2001; 92:404–408. [PubMed: 11291078]

Elston CW, Ellis IO. Histopathology. 1991; 19:403–410. [PubMed: 1757079]

Fiegler H, Carr P, Douglas EJ, Burford DC, Hunt S, Scott CE, et al. Genes Chromosomes Cancer. 2003; 36:361–374. [PubMed: 12619160]

Frengen E, Rocca-Serra P, Shaposhnikov S, Taine L, Thorsen J, Bepoldin C, et al. Genomics. 2000; 70:273–285. [PubMed: 11161777]

Hwang ES, Nyante SJ, Yi Chen Y, Moore D, DeVries S, Korkola JE, et al. Cancer. 2004; 100:2562–2572. [PubMed: 15197797]

Jain AN, Tokuyasu TA, Snijders AM, Segraves R, Albertson DG, Pinkel D. Genome Res. 2002; 12:325–332. [PubMed: 11827952]

Kern SE. Cancer Biol Ther. 2002; 1:571–581. [PubMed: 12496492]

Korsching E, Packeisen J, Helms MW, Kersting C, Voss R, van Diest PJ, et al. Br J Cancer. 2004; 90:1422–1428. [PubMed: 15054466]

Millis RR, Barnes DM, Lampejo OT, Egan MK, Smith P. Eur J Cancer. 1998; 34:548–553. [PubMed: 9713307]

Nishizaki T, Chew K, Chu L, Isola J, Kallioniemi A, Weidner N, et al. Int J Cancer. 1997; 74:513–517. [PubMed: 9355973]

Pandis N, Idvall I, Bardi G, Jin Y, Gorunova L, Mertens F, et al. Int J Cancer. 1996; 66:191–196. [PubMed: 8603810]

Quon KC, Berns A. Genes Dev. 2001; 15:2917–2921. [PubMed: 11711426]

Roylance R, Droufakou S, Gorman P, Gillett C, Hart IR, Hanby A, et al. J Pathol. 2003; 200:53–58. [PubMed: 12692841]

Roylance R, Gorman P, Hanby A, Tomlinson I. J Pathol. 2002; 196:32–36. [PubMed: 11748639]

Roylance R, Gorman P, Harris W, Liebmann R, Barnes D, Hanby A, et al. Cancer Res. 1999; 59:1433–1436. [PubMed: 10197608]

Santarosa M, Ashworth A. Biochim Biophys Acta. 2004; 1654:105–122. [PubMed: 15172699]

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, et al. Am J Hum Genet. 2005; 77:78–88. [PubMed: 15918152]

Tomlinson IP, Lambros MB, Roylance RR. Genes Chromosomes Cancer. 2002; 34:349–353. [PubMed: 12112523]

Tsuda H, Callen DF, Fukutomi T, Nakamura Y, Hirohashi S. Cancer Res. 1994; 54:513–517. [PubMed: 8275489]

Watson JE, Doggett NA, Albertson DG, Andaya A, Chinnaiyan A, van Dekken H, et al. Oncogene. 2004; 23:3487–3494. [PubMed: 15007382]

Whitmore SA, Crawford J, Apostolou S, Eyre H, Baker E, Lower KM, et al. Genomics. 1998; 50:1–8. [PubMed: 9628816]

Europe PMC Funders Author Manuscripts

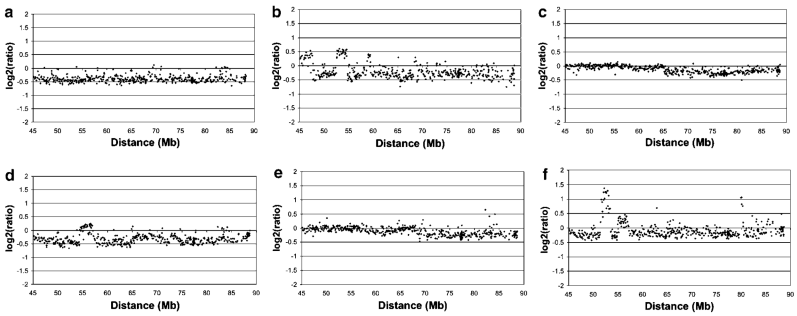

Figure 1.

Specimen aCGH profiles. Graphs showing the $\log_2$ ratio profiles for the long arm of chromosome 16 for a selection of all morphological subtypes of invasive breast cancer. (**a**) GI IDC with complete loss of 16q. (**b**) GI IDC with complex changes of both gain and loss. (**c**) ILC with a 'breakpoint' at 65.0 Mb before loss of the rest of the arm. (**d**) GII IDC with complete loss of 16q apart from a region between 54.4 and 57.0 Mb. (**e**) GIII IDC with a breakpoint at 68 Mb followed by loss. (**f**) GIII IDC with regions of loss and of relatively high-level gain.

Europe PMC Funders Author Manuscripts

Figure 2.
'Heat map' of aCGH changes in cancers with changes involving part of the 16q chromosome arm. Gain (green), loss (red) and no change (white) are shown at each clone for the GI IDCs, ILCs, GII IDCs and GIII IDCs (tumour ID numbers shown). Position in Mb along the chromosome arm is shown left. Occasional, individual clones reporting a ratio discordant from other clones within a larger region of consistent gain or loss are not shown.

Europe PMC Funders Author Manuscripts

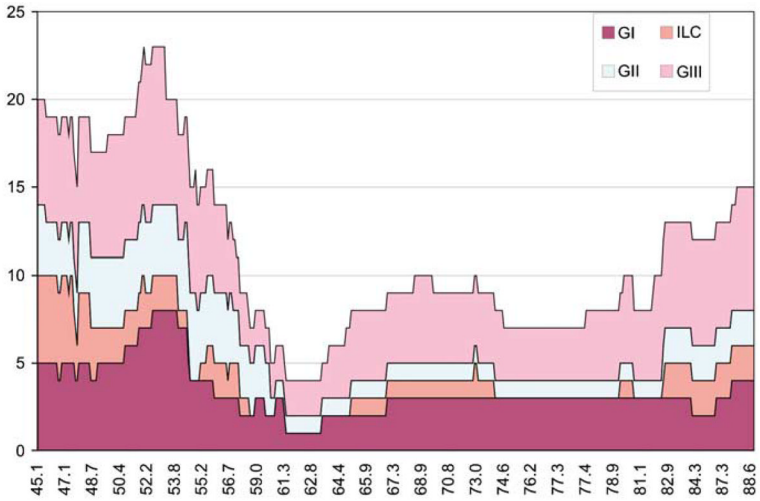

Figure 3.
Frequencies of gain along the chromosome arm as proportions of each of the four tumour types. Cumulative frequencies at each clone are shown (*y* axis) by clone position (Mb) (*x* axis).

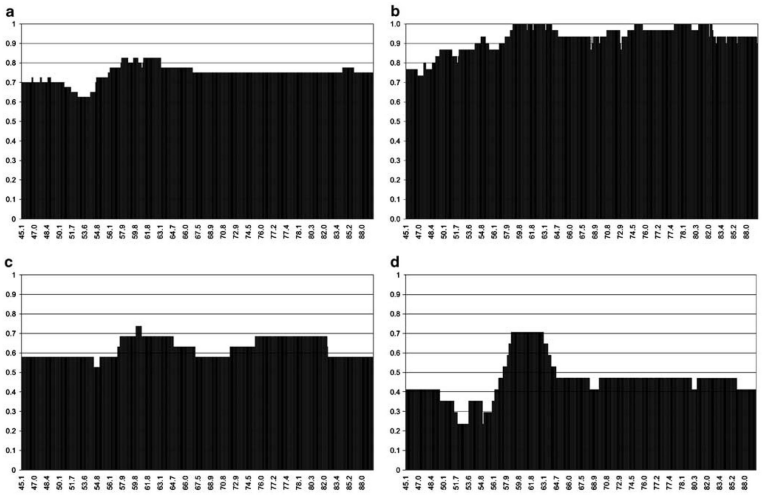

Figure 4.

Frequencies of loss at each clone as a proportion of all tumours of each of the four types. (**a**) GI IDCs, (**b**) ILCs, (**c**) GII IDCs, and (**d**) GIII IDCs. *y* axis shows frequency of loss; *x*-axis shows clone position (Mb).

Europe PMC Funders Author Manuscripts

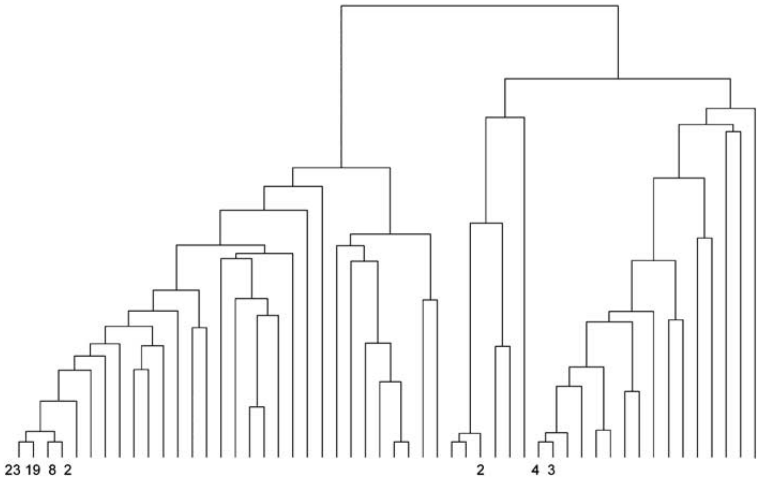

Figure 5.
Cluster dendrogram showing three groups of tumours. Each terminal stem represents one tumour, except where indicated by the numbers under a branch.

Europe PMC Funders Author Manuscripts

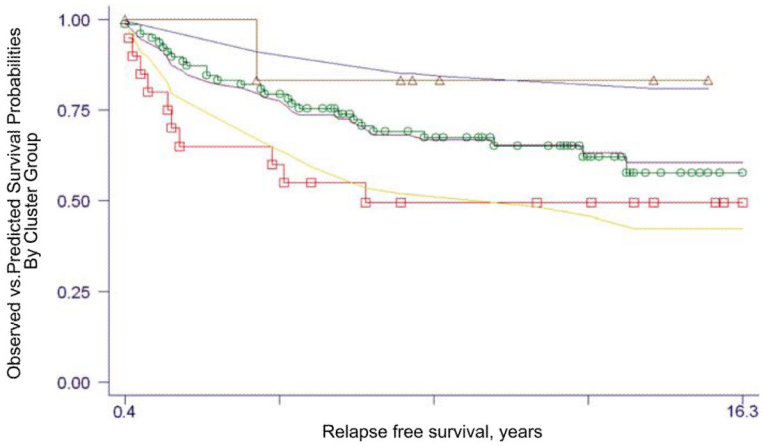

Figure 6.

Proportion of patients with relapse-free survival in each of the three cluster groups. Kaplan–Meier curves and the fitted curves under the Cox proportional hazards model are shown, with the data stratified by morphological type (GI, GII and GIII IDC and ILC). Group 1 (green) comprises principally cancers with loss of the whole chromosome arm, Group 2 (brown), cancers principally with gain and Group 3 (red), cancers principally with no change. The fitted curves for each group are purple, blue and yellow, respectively. Cancers with mixed gains, losses and no change are split among the three groups. The Cox regression model estimates a relative hazard ratio of 1.70 (95% CI 1.08–2.66, P=0.020).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Summary of aCGH changes found in each of the four tumour types

	Armloss	NC	Armgain	Mixed (%)	Total	Any loss (%)
GI IDC	23 (58)	4 (10)	1 (3)	12 (30)	40	34 (85)
ILC	19 (63)	0 (0)	0 (0)	11 (37)	30	30 (100)
GII IDC	8 (42)	3 (16)	1 (5)	7 (37)	19	14 (74)
GIII IDC	2 (12)	1 (6)	2 (12)	12 (71)	17	12 (71)

Abbreviations: Armgain, gain of whole arm; Armloss, deletion of whole chromosome arm; Anyloss, loss of any clone; GI IDC, grade I infiltrating ductal carcinoma; ILC, infiltrating lobular carcinoma; Mixed, two or more regions of gain and/or loss and/or no change; NC, no change Number of cancers (%) are shown.

Table 2

Small discrete regions of gain and loss detected in individual tumours by aCGH

Distance (Mb)	Flanking clones	Tumour no.	Gain or loss	Genes
45–45.5	RP11-46D6-RP11-283C7	T1044,T1226, T1498	Gain	AK125968, SHCBP1, VPS35, ORC6L, LOC91807, BC056676, GPT2
45–46.7	RP11-46D6-RP11-264A16	T1758	Loss	AK125968, SHCBP1, VPS35, ORC6L, LOC91807, BC056676, GPT2, DNAJA2, NETO2, CDA08, PHKB
47.9–48.4	RP11-358G8-RP11-305A4	T978	Gain	CBLN1, AK126710, MGC33367, OAZ
51.5–53.4	RP11-467J12-RP11-338P8	T1151	Gain	AK092208, AF150735, FLJ12178, BX537874, RBL2, FTS, AB023222, AB051539, IRX3
53.6–54.8	RP11-453K2-RP11-441F2	T1151	Loss	IRX6, MMP2, FLJ20481, CAPNS2, SLC6A2, CES1, CES4, FLJ31547, GNAO1
55.6–56.5	RP11-322D14-RP11-332G1	T1151,T1513	Gain	CETP, NOD27, CPNE2, NIP30, KIAA1972, TM4SF11, CCL22, CX3CL1, CCL17, LOC57019, DKFZP434K046, POLR2C, DOK4, GPR114, GPR56, GPR97, DKFZp434I099, KATNB1, KIFC3, AF452719, CNGB1
56.7–58.2	RP11-459F6-RP11-105C20	T1151,T1772	Gain	GTL3, CSNK2A2, HSPC065, BC057843, FLJ13912, NDRG4, KIAA1007, BCGF1, FLJ10815, GOT2, AK057513
57.8–61	RP11-55H18-RP11-267H9	T1093	Loss	CDH8
58.8–59.4	RP11-457D20-RP11-90K18	T1498	Gain	None
60.4–61.1	RP11-157H19-RP11-467O15	T1066	Gain	CDH8
72.9–73	RP11-252A24-RP11-572F4	T1511	Gain	PSMD7, AK124154
74.5–75.7	RP11-707E24-RP11-357G16	T865	Loss	CASPR4
77.4–79.3	RP1-2-RP11-314O13	T865	Loss	AF211943, AF447709, MAF, DNCL2B, CDYL2
79.9–80.4	RP11-340B7-RP11-465H19	T865,T1151	Gain	GAN, CMIP, PLCG2
88.3–88.7	RP11-104N10-cosmid425E4	T175,T1005	Gain	ZFP276, FANCA, AJ422077, KIAA1049, MC1R, TUBB4, AK096485, AF289561, MGC13198, AFG3L1, AK127143, MGC3101, GAS8, PRDM7

Distances are in Mb from May 2004 Human Genome Freeze, flanking markers are shown. Cancers with a discrete change involving the region are shown, although only one cancer defines the minimal region in each case.