

Review

Array-CGH and breast cancer

Erik H van Beers¹ and Petra M Nederlof²

¹Division of Experimental Therapy, Netherlands Cancer Institute NKI-AVL, Amsterdam

²Department of Pathology, Netherlands Cancer Institute NKI-AVL, Amsterdam, The Netherlands

Corresponding author: Petra M Nederlof, p.nederlof@nki.nl

Published: 30 June 2006

This article is online at <http://breast-cancer-research.com/content/8/3/210>

© 2006 BioMed Central Ltd

Breast Cancer Research 2006, **8**:210 (doi:10.1186/bcr1510)

Abstract

The introduction of comparative genomic hybridization (CGH) in 1992 opened new avenues in genomic investigation; in particular, it advanced analysis of solid tumours, including breast cancer, because it obviated the need to culture cells before their chromosomes could be analyzed. The current generation of CGH analysis uses ordered arrays of genomic DNA sequences and is therefore referred to as array-CGH or matrix-CGH. It was introduced in 1998, and further increased the potential of CGH to provide insight into the fundamental processes of chromosomal instability and cancer. This review provides a critical evaluation of the data published on array-CGH and breast cancer, and discusses some of its expected future value and developments.

Introduction

The precise aetiology of the majority of breast cancers is elusive, which contrasts with the estimated 5-10% that are caused by inherited mutations. It is clear that breast cancer presents as a collection of distinct disease types that differ in disease progression, treatment response and disease-free survival. In addition to conventional pathology, emerging technologies, including micro-arrays, have proven to be excellent tools in enhancing our understanding of individual breast cancers and providing assistance in treatment decision making in clinically relevant subgroups [1]. Initial observations, indeed most of our current understanding of chromosome abnormalities came from conventional cytogenetics. The importance of DNA copy number alterations has been demonstrated in many tumours [2]. More recently, the results of array comparative genomic hybridization (array-CGH) analysis of tumour tissues have been described from several perspectives, including identification of subgroups (class discovery); identification of genes that are involved in tumour progression, metastasis and treatment response; identification of candidate oncogenes (in genetic amplifications) or tumour suppressors (in homozygous deletions); and classification of hereditary cancers. A few

studies have also described CGH-based distinctions between sporadic and familial cases of breast cancer. Among familial cases, further classification into *BRCA1*, *BRCA2*, or 'other genetic risk factor(s)' is useful in these families in general but also for potential additional gene discovery.

Array-CGH technology is a fairly recent and important upgrade to the groundbreaking (conventional, metaphase chromosome) CGH technology [3,4], and it is applied to detect chromosomal DNA copy number alterations (CNA) in cells or tissues that occur as a result of genomic instability. CGH technology caused a paradigm shift in (clinical) cytogenetics because it avoids the need to culture (tumour) cells. This is perhaps the main reason why there was considerable over-representation of information on nonsolid tumours in cytogenetic databases. Array-CGH can measure genome-wide copy numbers in an unprecedented and objective manner. More importantly, complex karyotypes – a hallmark of many tumours – could be analyzed, including those too complex for G-banding because CGH readout is done on normal metaphase chromosomes. Last but not least, CGH was spectacular step forward because it preceded and did not require the completion of the human genome (draft) sequence [5]. It was not until 1998, with knowledge of the human sequence, that array-CGH began to replace conventional CGH [6], again with some major improvements.

Array-CGH is superbly reviewed elsewhere [7-12], including its development, initial challenges and ultimate benefits relative to conventional CGH. However, application of array-CGH to breast cancer research has not specifically been addressed. In this review, we discuss some of the prospects of using array-CGH data to elucidate the process of carcinogenesis in breast cells and its possible impact on diagnostic stratification for clinical management. The review addresses the two goals of array-CGH in breast cancer, which can be

BAC = bacterial artificial chromosome; CESH = comparative expressed sequence hybridization; CGH = comparative genomic hybridization; CNA = copy number alterations; FISH = fluorescent *in situ* hybridization; LOH = loss of heterozygosity; PCR = polymerase chain reaction; SNP = single nucleotide polymorphism.

summarized as follows: (novel) gene discovery in relation to subtypes, stage, or prognosis of breast cancer; and building class discovery tools for classification of independent breast cancers, regardless of the primary aim of identifying genes.

Technical issues

The effective resolution, sensitivity and reproducibility of array-CGH are constantly being enhanced. Although the maximal technical resolution of CGH is now approximately 140 kb [13], the resolution of the majority of published array-CGH studies is in the order of 1 or 2 Mb. This is the average spacing of probes along the genome but varies considerably by region. The sensitivity of nearly all platforms is sufficient to detect one-copy gains and losses, provided that at least 70% of the cells in the sample are tumour cells. Thus, the detection limits decrease with a combination of the following parameters; smaller aberration size (in bases), smaller aberration amplitude (copy number) and smaller tumour subclone representation.

Choice of probe sets

Whether to employ probe sets including large insert clones such as bacterial artificial chromosome (BAC), yeast artificial chromosome, cosmid and fosmid, as opposed to oligonucleotide (generally 60-70 mers) probe sets, is a major consideration in the design of array-CGH studies. Most characteristics of various platforms are described in detail elsewhere [11]. Clearly, the advantage of using oligonucleotide provides the greatest flexibility in terms of availability, which is limited only by knowledge of the DNA sequence for the species under investigation. Careful selection of probes is important because the genome contains many pseudogenes and duplicons; this has been done for human and mouse up to about 44k [14-16]. The disadvantage of using short 60-mer probes is the low hybridization signal, which in theory is a factor of about 2000 times lower than when using the much larger insert BAC clones, which in turn suffer from variable but high DNA repeat content that must be quenched by Cot1 (the dominant fraction of human repetitive DNA sequence)-DNA in the hybridization mix. In practice, to deal with the low signal to noise ratio, generally a moving average of about 1-5 Mb is often applied, which in effect decreases resolution. The proprietary technology of Agilent (Palo Alto, CA, USA) to print very high-density arrays enables genome-wide oligo-array-CGH to be conducted on one microscope slide; this contrasts with the genome-wide tiling (32k) BAC array, which uses two slides for the same coverage, doubling the cost for the profile. One way to overcome this problem is to use a 1 Mb resolution for the complete genome followed by tiling resolution for the entire genome [17] or for selected regions of interest [18].

Data analysis

Once an array-CGH profile has been established using any of a range of adequate protocols available [19-27], one may

decide to use segmentation statistics to determine the boundaries between and the copy number levels of the aberrant genomic regions. The underlying biology of copy numbers has been modelled and assumes that CGH data are a combination of underlying DNA copy number with a component of Gaussian noise. One such algorithm uses the maximum likelihood criterion adjusted using a penalization term for taking into account model complexity to define breakpoints [28]. Further, in-depth analyses as well as state-of-the-art implementation of more complex assumptions have been described [28-30]. These methods essentially aim to 'de-noise' CGH data to enhance breakpoint identification, such that candidate regions can be identified with greater precision [18,31-40]. With respect to sensitivity in detecting gains and losses, it is also critical to use reference samples that contain a pool of DNA from at least six, unrelated healthy individuals to suppress the (false) discovery of aberrations in the control sample, caused by the existence of 'normal' inter-individual variation in copy numbers [41].

This is in accordance with what others have experienced, namely that advancing to array-CGH leads to much more data, of similar quality. Having more data per sample poses a risk for over-fitting, in other words a risk for finding significant differences between groups by chance alone. One way in which studies are coping with over-fitting is to increase outlier thresholds (e.g. from $\log_2\text{ratio} = 0.2$ to 0.3) or to require a certain minimum proportion (say 33%) of tumours to share a particular aberration. This effectively restricts the number of features (probes) left for comparison but it may not always be the optimal choice. Nevertheless, this is a method for finding the dominant features first. We believe that it may be rather optimistic to set a minimum number of samples for identifying significant aberrations at, for instance, 33% because this would easily ignore (not detect) rare breast cancer subtypes that tend to be under-represented in any particular study.

Array-CGH and expression arrays

Study of gene expression using array technology has garnered enormous popularity, and continues to generate high-impact classifiers of tumour type and disease progression [3,42-44]. One must keep in mind that expression arrays and array-CGH measure different things, which are not always easily correlated. Expression arrays measure relative abundance of specific mRNA transcripts and CGH measures relative copy numbers of genomic DNA regions in a sample. The expression levels per gene vary enormously even in normal cells, depending, for instance, on cell type and cell cycle. The chromosomal fragment (or gene) copy numbers as measured by array-CGH show much less variation, being one or two copies in normal cells, and perhaps ranging between none and 20 copies in tumour cells. It is therefore difficult to correlate expression data with genomic data directly.

Several studies that have correlated DNA copy number data (CGH) with expression profiles have been published

[37,45,46]. They indicate a significant correlation between copy number and gene expression. In regions of high-level amplification, 44–62% of genes were reported to be over-expressed [45,46], compared with just 12% of all genes. The relation may in some cases be less straightforward, probably due to the complexity of gene regulation, including pathway feedback loops. An illustration for this is that breast cancers with *ERBB2* (17q21) amplification almost invariably overexpress the gene, which contrasts with only 55% of cases with a gain of *ERBB1* (7p12) [47]. Arrays that support measurement of DNA as well as RNA include cDNA, oligo-, and single nucleotide polymorphism (SNP) arrays. The benefits of dual use are limited because CGH signals become critically low on short probes; short-probe (high-complexity) arrays tend to be much more expensive than BAC arrays; and the reference sample in expression profiling is not defined but at best held constant across a series of experiments, as opposed to using a completely 2n (normal karyotype) reference for array-CGH.

Another important advantage of CGH compared with expression arrays is that CGH can be conducted reliably using archival formalin-fixed paraffin-embedded material [26,48-50]. Although there is evidence that archival material is also suited to expression analysis, it is only with recently fixed tissues and only with much effort [51].

Perspective from the tumour (cell line) to DNA copy number

One major application of CGH is in the analysis of cancer genomes for genome-wide copy number changes. The resolution of classic chromosomal CGH is limited by the highly condensed state of metaphase chromosomes and further by the optical resolution of the microscopes used to capture the hybridization signals. In addition, chromosome condensation may vary from metaphase spread to metaphase spread. The analysis software normalizes all chromosomes and assumes linear condensation of all chromosomes. As a result, the precise location of aberrations becomes uncertain, which reduces the true effective resolution. In contrast, the resolution of array-CGH is limited only by the density and average length of the probe set printed on the array. Compared with metaphase-CGH, one problem – especially with earlier versions of array-CGH – was the uncertainty regarding the genomic location of the clones, which is dependent on the version of the ‘genome build’ of the human genome. The ability to detect a change of one copy for both technologies is roughly similar but highly influenced by the frequent admixture of nontumour cells in the sample, as well as by heterogeneity of the tumour tissue (i.e. generally not all cells will have the same ploidy for all chromosomal regions). Both factors are important in the analysis of breast cancer, because breast tumours are heterogeneous and may contain significant proportions of normal cells. Therefore, as a rule CGH should be performed only on tissue samples containing 70% or more tumour cells, or one must meet this requirement

through enrichment by macrodissection, (laser capture) microdissection, or flow sorting (fluorescence-activated cell sorting), followed by some means of DNA amplification [27,52].

Contrary to mRNA content, which for a given gene in a cell can range between no copies (not expressed) up to many thousands, the chromosome content of normal (reference) cells is extremely stable (2n, or diploid); when it is unstable, as is frequent in cancer cells, it at least remains quite discrete. This means that extensive regions of the genome are usually still balanced as in the diploid (2n or unchanged) or tetraploid (duplicated) state. Unbalanced regions exist either as homozygous loss (0 copies), heterozygous loss (1 copy or loss of heterozygosity [LOH]), gain (~3 or 4 copies), and amplifications (~5 or more copies). The distinction between diploid and triploid or tetraploid genomes cannot be detected by array-CGH but is best determined by some independent cell-based method such as fluorescent *in situ* hybridization (FISH).

And from DNA copy number back to the tumour

If we reverse the perspective, array-CGH data can also tell us something about tumours. This is similar to how cytogenetic data were used to define subgroups in breast cancer that correlated with histological type, grade and mitotic activity of the tumour [53,54]. CGH has been used to classify breast tumours [55,56]. Because breast cancer is a disease with high levels of chromosome instability, it can readily be studied by CGH. Results from various classification studies have indicated that many gains and losses show recurrences anywhere between 20% and 80% of all tumours in a class, depending on the region and the cancer subtype investigated. Some recurrences exhibit sufficient differential gains or losses between classes to permit their use for classification, as in the case of inherited *BRCA1* and *BRCA2* mutations [55,56]. Certain regions such as 1q and 8q almost exclusively exhibit gain and rarely loss, whereas 16q often shows loss but hardly ever gain.

Robust classification procedures require more than counting frequencies of specific aberrations in various tumour types. It is important to appreciate that even a significant correlation of, for instance, 16q loss, which is more frequent in lobular breast cancer [57], provides limited statistical prediction power for an individual prospective case. This requires more rigorous, iterative methods of feature and model selection, followed by cross-validations and external validations. Provided that the correct methods are used, we believe that array-CGH can be useful for identifying distinct breast tumour subtypes, and can help to define them further, similar to the study reported by Jonsson and coworkers [58].

CGH and cancer

Cancer is the result of a myriad of genetic and epigenetic alterations. Identification of the causal perturbations that

confer malignant transformation is a central goal in cancer biology. CGH is a powerful tool for investigating tumours in a genome-wide manner for such candidate regions. This strategy was successfully used to identify *c-Myc*, *Her2Neu*, *Rab25* and a range of other potential oncogenes [32,59,60]. *Vice versa*, beginning with knowledge of the causal gene for the tumour, CGH has been useful in elucidating the extent of specific as well as recurrent aberrations in tumours such as those associated with mutations in *BRCA1*, *BRCA2*, or *P53* in both mouse [61] and human [55,56,58,62,63]. Genomic instability occurs is still not fully understood. Recently, a number of possible mechanisms focusing on the fidelity of chromosome segregation and/or DNA repair have been proposed [64,65], which may hold some indirect clues as to how breast tumour CGH profiles in *BRCA1* mutation carriers are different from those in sporadic cases, but they still fail to explain why. While progress is made to elucidate further the precise nature of genomic instability, the resulting specificity of CGH profiles has already been put to good use.

Classification

For instance, CGH profiles have been used in the classification of breast tumours of unknown causality into *BRCA1* mutation carriers and noncarriers [56]; to delineate the relationship between synchronous, recurrent and/or metastatic tumours [55,66-71]; and to define the recurrent aberrations that appear to be associated with certain clinical types of breast cancer (e.g. ductal tumours) or with prognosis or clinical course, or both [17,57,72-79] (Table 1). The study by Rennstam and coworkers [72] was performed using metaphase-CGH, but it clearly demonstrated differential 5-year survival statistics (56% versus 96%) for distinct CNA tumour types that were independent of more conventional markers such as grade, and progesterone receptor or node status. Jones and coworkers [73] used CGH to subclassify 86 breast tumours of grade III and basal type into groups with shorter (3.5 years) and longer (15 years) survival.

Others have successfully used metaphase and/or array-CGH profiles for classification and for mutation pre-screening. Both Wessels and coworkers [56] and Alvarez and colleagues [77] conducted studies that were effective in identifying *BRCA1*-associated tumours based on CGH profiles of the tumours. Wessels and coworkers [56] identified 33 out of 34 proven *BRCA1* cases and assigned 10 false-positives among bilateral breast cancer cases (enriched for elevated risk), four of which have since been proven to be true *BRCA1* mutation carriers (personal communication). This CGH driven classification has been repeated by Van Beers and coworkers [55], who reported specific CGH aberrations for *BRCA2*-related breast tumours. Also along these lines, Alvarez and colleagues [77] built a predictor for *BRCA1* and found nearly half of the '*BRCA1*-like' CGH profiles in their '*BRCAX*' breast cancer type to be hypermethylated on the *BRCA1* promoter, suggesting that loss of *BRCA1* could have been an initiating event in these tumours.

Candidate gene searches

A number of results of array-CGH analysis of breast tumours and breast cancer cell lines have been published [19,21,22,25,27,80-88] (Table 1). Some of these results, together with large amounts of conventional CGH, FISH and SKY data of breast cancer, are freely available through the Progenetix repository [89,90]; NCI and NCBI's SKY/M-FISH and CGH Database (2001) [91]; Charité, Berlin University [92]; and in the supplementary data of many individual publications. Cytogenetic data has been described [53,54] and is available for more than 1000 cases at NCBI. The following discussion illustrates the use and usefulness of array-CGH for various breast cancer genomes.

A detailed study of 31 advanced archival breast tumours conducted by Nessling and coworkers [47] elegantly demonstrates how specificity, sensitivity and resolution are increased in (matrix) array-CGH compared with conventional CGH by using the same samples on both platforms. They identified 44 genes by array-CGH and verified all of them by PCR in 31 breast tumours. Many of these genes are located in common altered regions in breast cancer, such as *A/B1* (amplified in breast cancer-1) at 20q12 (68% of their cases). A novel and significant finding was gain of 6p21, containing *CCND3* and *p21/WAF1*, in 45% of their cases. There is ample evidence implicating these genes in cancer, and so a gain of 6q21 suggests a role for *CCND3* and *p21/WAF1* in these cancers also.

Cowell and coworkers [93], using FISH, observed a single translocation t(3;9) in the MCF10 'normal breast' cell line associated with an immortalized phenotype. During further passages this immortalized cell lineage (MCF10A) acquired additional alterations including t(6;19) and gain of 1q, detected by CGH. This may suggest that the 1q gain in breast cancer is an early event and thus may explain why it is so common (>40% of all breast cancers). It is quite remarkable for a metaphase-CGH study, even after verification of the predominant rearrangements by FISH, to close in on potential tumour genes, such as *TAF1* and *p16* in this study. A similar approach looking for recurrent alterations among selected (familiar) breast cancers by metaphase-CGH did find such a candidate region but it failed to identify predisposing mutations. It is reasonable to assume that array-CGH outperforms conventional CGH in this respect, and is more efficient as a first step toward evaluating candidate regions, mainly because of increased resolution. Such candidate gene searches by array-CGH will be valuable in identifying regions and ultimately genes that are associated with specific phenotypes such as cell growth, anchorage-independent survival and metastasis capacity, which are relevant to and could translate into differential clinical treatment. Although counterintuitive, it might be important to recognize that regions of recurrent aberrations may not contain the mutations sought at all. This is exemplified by the recurrent aberrations in *BRCA1* mutant tumours occurring on

Table 1**Studies using array-CGH on breast cancer samples or breast cancer cell lines**

Patient source or cell line	Array platform	Resolution (Mb)	Ref.
2 Human cell lines, including breast (BT474)	18,664 oligo nucleotides	~0.15	[16]
BC (<i>n</i> = 5)	32k tiling BAC	Sub-Mb	[17]
Primary BC (<i>n</i> = 33), BC cell line (<i>n</i> = 27)	Genome-wide BAC and high resolution on 8p11-12	~1.5	[18]
FFPE BC (<i>n</i> = 44)	4153 BAC	<1	[24]
Comparison of FFPE with frozen BC	2464 BAC	1.5	[27]
Gain in 47% BC (unpublished results)	See Pinkel and coworkers [2]		[32]
FFPE BC (<i>n</i> = 31)	471 BAC/PAC	15	[47]
<i>BRCA1</i> (<i>n</i> = 14), <i>BRCA2</i> (<i>n</i> = 12), sporadic (<i>n</i> = 26)	3.5k BAC	1	[58]
1 Case report	1777 BAC	~1.5	[66]
24 pairs LCIS + ILC	~2400 BAC	1.5	[68]
4 synchronous tumours (2 BC and 2 leiomyosarcomas)	2.5k BAC	1.5	[70]
BC cell line (<i>n</i> = 1)	6k BAC	<1	[93]
11 BC cell lines	2464 BAC (Hum-Array1.14)	1.5	[95]
52 node-negative invasive BC	57 oncogene loci (AmpliOnc™)		[96]
Premenopausal BC and radiation (<i>n</i> = 47)	3.5k BAC	1	[107]
16 cell lines (8 of them BC)	9862 SNPs		[108]
2 BC cell lines	2560 BAC	1.5	[109]
47 fresh-frozen IDC BC	4134 BAC	0.9	[110]

BAC, bacterial artificial chromosome; BC, breast cancer; CGH, comparative genomic hybridization; FFPE, formalin-fixed paraffin-embedded; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; LCIS, lobular carcinoma *in situ*; SNP, single nucleotide polymorphism.

chromosomes 3 and 5 and 10, which are more frequent than aberrations at the *BRCA1* locus itself [56]. Thus, CGH has helped to uncover some of the intrinsic complexity of tumour chromosome behaviour, which is still poorly understood.

There are several effective ways to avoid some of the risks involved in searching complexly rearranged genomes for candidate genes. One is to map global gene expression onto genomic positions using comparative expressed sequence hybridization (CESH) [94]. Another is to focus on known candidate regions and simultaneously monitor gene expression as a filtering step to exclude genes. This was the approach used by Garcia and coworkers [18], who were prompted by results reported by Ray and colleagues [95] to study gene expression and CGH at near-tiling resolution of 33 primary breast cancers, 27 breast cancer cell lines and 20 primary ovarian cancers at chromosome 8p11-12 (about 10 Mb), which is a gene-dense region that has been implicated in various tumour types. By cross-comparison they were able to define a minimal region of common amplification that contained four overexpressed and therefore candidate oncogenes, namely *FLJ14299*, *C8org2*, *BRF2* and *RAB11FIP*. This important finding requires confirmation in independent

series of breast cancer, but the study clearly demonstrates the power of the CGH approach in combination with other assays.

At the same time, Prentice and coworkers [17] reported the same amplified region in 24% of all 382 cases of breast cancer examined by FISH on tissue micro-arrays followed by BAC array-CGH in five cases. They reported a minimal common amplified region in all five cases centromeric to *NRG* and *FGFR1* (which are also frequently involved in breast cancer) that contain just three genes, namely *FLJ14299*, *SPFH2* (*C8org2*) and *PROSC*. Interestingly, only *FLJ14299* and *SPFH2* overlap with findings reported by Garcia and coworkers, suggesting that amplification of one of these, or both, could represent functional alterations in breast cancer. The only information available for *FLJ14299* is that it resembles a C2H2 zinc-finger type transcription factor that is conserved in *Drosophila* and zebra fish (*Danio rerio*). For *SPFH2*, a potential membrane association is predicted but unproven (GeneCards; www.genecards.org). Although high-level amplification in this region correlated with poor survival in both studies, further studies will be necessary to determine whether these genes have the capacity of an oncogene.

Another candidate gene, namely *Rab25*, was identified through array-CGH by Cheng and coworkers [32]. *Rab25* was found to stimulate anchorage-independent cell survival and was thus characterized as a potential driver of ovarian and breast tumour development. Those investigators showed that the copy number of this 1q22 region correlated with differential disease-free survival in ovarian cancer patients, further suggesting that 1q22 is associated with tumour aggressiveness.

The above examples illustrate the usefulness of array-CGH in cancer genome research and justify further investigation involving mapping more candidate cancer genes in a relatively straightforward and high-throughput manner.

Hereditary breast (and ovarian) cancer and array-CGH

Jonsson and coworkers [58] reported array-CGH findings in 26 cases of hereditary breast cancer and 26 cases of sporadic breast cancer. They reported one recurrent amplicon (3q27.1-3) in 71% of 14 *BRCA1* breast tumours and one recurrent amplicon (17q23.3-24.2) in more than 75% of 12 *BRCA2* breast tumours. They further identified a set of 169 BAC clones out of their total 3.6k probe set, with sufficiently differential log ratios between *BRCA1*, *BRCA2* and sporadic cancer to permit its use for classification of hereditary cases. However, this classifier will only gain validity after independent CGH profiles (a validation set) can verify it.

Wessels and colleagues [56] built a molecular classifier with a performance of 84% for detecting *BRCA1* tumours. This classifier was made using 36 breast tumours from proven *BRCA1* mutation carriers and 30 breast tumours from 30 independent bilateral (elevated risk) breast cancer patients. This classifier was built from metaphase-CGH but performs well on array CGH data from *BRCA1*, *BRCA2*, sporadic and BRCAX cases (van Beers EH, unpublished data). A classifier for *BRCA2* has been more problematic. Using metaphase-CGH, a number of significant and recurrent genomic aberrations in 25 *BRCA2* tumours were described [55] but the collective predictive power fell short of building a reliable classifier. We believe that the improved resolution together with the improved reproducibility of array-CGH might generate the data necessary for a *BRCA2* classifier and possibly also for other subtypes in the future.

Prognostic information from CGH data

Callagy and coworkers [96] have made an important contribution to clinical prognostication using an array-CGH based study. They asked how array-CGH was different between short-term survivors (<5 years) and long-term survivors (>10 years). As a result, they found a trio of *TOP2a*, *ERBB2* and *EMS1* to identify statistically significant ($P=0.01$) differential prognosis among these subgroups. The differential expression for these genes was further substantiated by tissue microarray FISH. Quite surprisingly, their good and bad prognosis groups did not exhibit

differences in grade, size, or oestrogen receptor status – features that are currently most widely accepted for clinical prognostication. Because results were not stratified by tumour type (40 of 52 tumours were invasive ductal carcinoma), this study does not permit a definitive conclusion to be drawn about whether this prognostic set of three genes will be equally valuable in all breast cancer types. It is worth mentioning that their CGH probe set consisted of just 57 cancer-gene selected loci. This suggests that larger probe sets could identify more markers, and more relevant ones.

Cingoz and colleagues [57] reported a number of recurrent CGH observations that appear to correlate with certain disease characteristics. Their list of most frequent gains includes 1q (55%), 8q (52%) and 20q (29%), which is similar to many other breast cancer CGH studies and can be considered general breast cancer copy number gains. Other regions were distinct to subgroups, such as 16q loss in seven out of nine (78%) cases of invasive lobular carcinoma compared with five out of 18 (28%) cases of invasive ductal carcinoma. Since invasive lobular and ductal carcinoma are easily distinguished by pathologists, their findings are probably most important to our understanding of the intrinsic biology of these different tumour types rather than being beneficial for diagnosis or clinical decision making.

Pathways of instability

It has long been recognized that several different pathways of genomic instability exist. Probably most apparent is the difference between chromosomal instability and microsatellite instability, the former being more frequent in breast cancer and the latter being found more often in relation to *MLH1* mutations that occur frequently in colon cancers [97]. It seems likely that certain tumours arrive at an aneuploid state through a series of events starting with telomere dysfunction, followed by polyploidization as a cellular rescue/survival event. Then, as a direct result chromosomes suffer numerous additional breakage-fusion events, resulting in seemingly chaotic accumulation of chromosome gains and losses [98]. This scenario is thought to occur frequently in adenocarcinomas, including breast cancer, with high-level aneuploidy. With respect to breast cancer, at least two distinct genome instability pathways exist and are described in detail in a recent review by Reis-Filho and coworkers [7].

Some related array-based methods

Any review of the analysis of chromosomal copy numbers in breast cancer by array-CGH would be incomplete without mention of some related technologies (e.g. genome-wide LOH analysis, SNP [haplotype] analysis, CESH [94,99] [also called expressive genomic hybridization [100]], large-scale PCR [101] and expression array studies) that often have similar or overlapping goals, including classification, clinical prognostication, defining treatment subgroups, finding novel genes, and so on.

Several studies have already demonstrated the feasibility of performing genome-wide LOH and SNP analysis, for example by using Illumina bead-array technology on formalin-fixed paraffin-embedded tumour tissues [102-104]. These technologies have a clear advantage over (array) CGH in that they extract haplotype information in addition to chromosome copy number. This is especially relevant in chromosomal uniparental isodisomy, inherited or acquired through mitotic recombination resulting in variably extended stretches of homozygosity (i.e. LOH) in the absence of copy number changes. This has been observed in breast cancer and may be more frequent than is generally acknowledged [105].

Conclusion

Array-CGH is a reliable, sensitive and high-resolution method that is highly automated compared with metaphase-CGH, which it is now replacing; thus, array-CGH is expected to generate an enormous amount of data in the coming years. One limitation of all current breast cancer array-CGH studies is the limited number of samples used per study. It poses the classic risk for over-fitting for as long as the average study contains two orders of magnitude more features (i.e. probes) than samples. One way to tackle this limitation is to have available the combined data sets or a repository for meta-analyses of CGH data, histological data, immunostaining data, clinical data, and so forth.

This review has described the two separate goals of CGH in breast cancer: gene discovery and class discovery. For both we have given successful examples. An example of gene identification is *TAF11*, the significance of which in cell transformation was found by array-CGH and was verified extensively with other methods [93]. These data clearly prioritize this locus for further functional studies of this relatively uncharacterized gene. An example of class prediction by array-CGH is the study Jonsson and coworkers [58], who built classifiers for *BRCA1* and *BRCA2* breast tumours.

One further conclusion of our review of the array-CGH literature is that the known heterogeneity of breast cancer also seems reflected in array-CGH data, such that multiple types of profiles have been reported, sometimes with clear but sometimes with less clear associations with known types of breast cancer. The fact is that the number of described subtypes of breast cancer increases with increasing sensitivity and resolution of newer methods. Although novel technology has no *a priori* knowledge of the problem analyzed, one would generally be satisfied if a new technology would permit enhanced stratification of disease entities. However, in breast cancer we see a trend toward discovery of more subgroups that sometimes appear independent of more traditional classification features, such as grade, size and immunochemistry (e.g. *P53*, oestrogen receptor, progesterone receptor, *ERBB2*) [96]. One interpretation is that the within group variability, for instance,

among all grade III cases or all oestrogen receptor negative cases is still quite large. In our opinion, this could account for the independent clustering when using CGH data compared with other tumour characteristics, even when the same tumour set is studied.

In the near future, when increasing numbers of studies generate higher resolution data in conjunction with allele-specific information, such as produced by SNP arrays or Illumina bead arrays, it may become possible also to elucidate some effects of genetic backgrounds or genotypes on CGH as a proxy for genomic instability. Nevertheless, genetic background appears to impact on CGH profiles [106], and this type of information should hopefully teach us more about the biology of chromosomal instability. This touches on a possible relation between CGH profiles and human genetic diversity that can best be studied on genome-wide SNP/CGH arrays and could be extremely useful in locating putative (median and low risk) breast cancer genes.

Despite the current trend in which oligo array-CGH is gradually replacing BAC array-CGH because of its flexibility (clone-less, fast and custom print-on-demand, and probably the most powerful advantage of assessing copy number and allele information simultaneously), it seems that BAC arrays will remain important for profiling DNA from formalin-fixed material.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank the Marie-Curie Foundation (meetings and course) that helped us to share CGH experiences with many researchers in this field.

References

1. 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, *et al.*: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**:530-536.
2. Heim S, Mitelman F: **Primary chromosome abnormalities in human neoplasia.** *Adv Cancer Res* 1989, **52**:1-43.
3. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D: **Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.** *Science* 1992, **258**:818-821.
4. Kallioniemi OP, Kallioniemi A, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D: **Comparative genomic hybridization: a rapid new method for detecting and mapping DNA amplification in tumors.** *Semin Cancer Biol* 1993, **4**:41-46.
5. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, *et al.*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.
6. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, *et al.*: **High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.** *Nat Genet* 1998, **20**:207-211.
7. Reis-Filho JS, Simpson PT, Gale T, Lakhani SR: **The molecular genetics of breast cancer: the contribution of comparative genomic hybridization.** *Pathol Res Pract* 2005, **201**:713-725.
8. Mantripragada KK, Buckley PG, de Stahl TD, Dumanski JP: **Genomic microarrays in the spotlight.** *Trends Genet* 2004, **20**: 87-94.

9. Bertucci F, Viens P, Hingamp P, Nasser V, Houlgatte R, Birnbaum D: **Breast cancer revisited using DNA array-based gene expression profiling.** *Int J Cancer* 2003, **103**:565-571.
10. Davies JJ, Wilson IM, Lam WL: **Array CGH technologies and their applications to cancer genomes.** *Chromosome Res* 2005, **13**:237-248.
11. Albertson DG, Pinkel D: **Genomic microarrays in human genetic disease and cancer.** *Hum Mol Genet* 2003, **12 Spec No 2**:R145-R152.
12. Albertson DG: **Profiling breast cancer by array CGH.** *Breast Cancer Res Treat* 2003, **78**:289-298.
13. Selzer RR, Richmond TA, Pofahl NJ, Green RD, Eis PS, Nair P, Brothman AR, Stallings RL: **Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH.** *Genes Chromosomes Cancer* 2005, **44**:305-319.
14. van den Ijssel P, Tijssen M, Chin SF, Eijk P, Carvalho B, Hopmans E, Holstege H, Bangarusamy DK, Jonkers J, Meijer GA, Caldas C, Ylstra B: **Human and mouse oligonucleotide-based array CGH.** *Nucleic Acids Res* 2005, **33**:e192.
15. van d, I, Tijssen M, Chin SF, Eijk P, Carvalho B, Hopmans E, Holstege H, Bangarusamy DK, Jonkers J, Meijer GA, et al.: **Human and mouse oligonucleotide-based array CGH.** *Nucleic Acids Res* 2005, **33**:e192.
16. Carvalho B, Ouwerkerk E, Meijer GA, Ylstra B: **High resolution microarray comparative genomic hybridisation analysis using spotted oligonucleotides.** *J Clin Pathol* 2004, **57**:644-646.
17. Prentice LM, Shadeo A, Lestou VS, Miller MA, deLeeuw RJ, Makretsov N, Turbin D, Brown LA, Macpherson N, Yorida E, et al.: **NRG1 gene rearrangements in clinical breast cancer: identification of an adjacent novel amplicon associated with poor prognosis.** *Oncogene* 2005, **24**:7281-7289.
18. Garcia MJ, Pole JC, Chin SF, Teschendorff A, Naderi A, Ozdag H, Vias M, Kranjac T, Subkhankulova T, Paish C, et al.: **A 1 Mb minimal amplicon at 8p11-12 in breast cancer identifies new candidate oncogenes.** *Oncogene* 2005, **24**:5235-5245.
19. Albertson DG, Collins C, McCormick F, Gray JW: **Chromosome aberrations in solid tumors.** *Nat Genet* 2003, **34**:369-376.
20. Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigoroava M, Jones KW, Wei W, Stratton MR, et al.: **High-resolution analysis of DNA copy number using oligonucleotide microarrays.** *Genome Res* 2004, **14**:287-295.
21. Brennan C, Zhang Y, Leo C, Feng B, Cauwels C, Aguirre AJ, Kim M, Protopopov A, Chin L: **High-resolution global profiling of genomic alterations with long oligonucleotide microarray.** *Cancer Res* 2004, **64**:4744-4748.
22. Cai WW, Mao JH, Chow CW, Damani S, Balmain A, Bradley A: **Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays.** *Nat Biotechnol* 2002, **20**:393-396.
23. Daigo Y, Chin SF, Goringe KL, Bobrow LG, Ponder BA, Pharoah PD, Caldas C: **Degenerate oligonucleotide primed-polymerase chain reaction-based array comparative genomic hybridization for extensive amplicon profiling of breast cancers: a new approach for the molecular analysis of paraffin-embedded cancer tissue.** *Am J Pathol* 2001, **158**:1623-1631.
24. Loo LW, Grove DI, Williams EM, Neal CL, Cousens LA, Schubert EL, Holcomb IN, Massa HF, Glogovac J, Li Cl, et al.: **Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes.** *Cancer Res* 2004, **64**:8541-8549.
25. Shadeo A, Lam WL: **Comprehensive copy number profiles of breast cancer cell model genomes.** *Breast Cancer Res* 2006, **8**:R9.
26. van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoef S, Nederlof PM: **A multiplex PCR predictor for aCGH success of FFPE samples.** *Br J Cancer* 2006, **94**:333-337.
27. DeVries S, Nyante S, Korkola J, Segraves R, Nakao K, Moore D, Bae H, Wilhelm M, Hwang S, Waldman F: **Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors.** *J Mol Diagn* 2005, **7**:65-71.
28. Willenbrock H, Fridlyand J: **A comparison study: applying segmentation to array CGH data for downstream analyses.** *Bioinformatics* 2005, **21**:4084-4091.
29. Myers CL, Dunham MJ, Kung SY, Troyanskaya OG: **Accurate detection of aneuploidies in array CGH and gene expression microarray data.** *Bioinformatics* 2004, **20**:3533-3543.
30. Jong YJ, Li LH, Tsou MH, Chen YJ, Cheng SH, Wang-Wuu S, Tsai SF, Chen CM, Huang AT, Hsu MT, et al.: **Chromosomal comparative genomic hybridization abnormalities in early- and late-onset human breast cancers: correlation with disease progression and TP53 mutations.** *Cancer Genet Cytogenet* 2004, **148**:55-65.
31. Agata S, Dalla Palma M, Callegaro M, Scaini MC, Menin C, Ghiotto C, Nicoletto O, Zavagno G, Chieco-Bianchi L, D'Andrea E, et al.: **Large genomic deletions inactivate the BRCA2 gene in breast cancer families.** *J Med Genet* 2005, **42**:e64.
32. Cheng KW, Lahad JP, Kuo WL, Lapuk A, Yamada K, Auersperg N, Liu J, Smith-McCune K, Lu KH, Fishman D, et al.: **The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers.** *Nat Med* 2004, **10**:1251-1256.
33. Heiskanen M, Kononen J, Barlund M, Torhorst J, Sauter G, Kallioniemi A, Kallioniemi O: **CGH, cDNA and tissue microarray analyses implicate FGFR2 amplification in a small subset of breast tumors.** *Anal Cell Pathol* 2001, **22**:229-234.
34. Huang HE, Chin SF, Ginestier C, Bardou VJ, Adelaide J, Iyer NG, Garcia MJ, Pole JC, Callagy GM, Hewitt SM, et al.: **A recurrent chromosome breakpoint in breast cancer at the NRG1/neuregulin 1/herregulin gene.** *Cancer Res* 2004, **64**:6840-6844.
35. Kauraniemi P, Barlund M, Monni O, Kallioniemi A: **New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays.** *Cancer Res* 2001, **61**:8235-8240.
36. Maguire P, Holmberg K, Kost-Alimova M, Imreh S, Skoog L, Lindblom A: **CGH analysis of familial non-BRCA1/BRCA2 breast tumors and mutation screening of a candidate locus on chromosome 17q11.2-12.** *Int J Mol Med* 2005, **16**:135-141.
37. Monni O, Barlund M, Mousses S, Kononen J, Sauter G, Heiskanen M, Paaola P, Avela K, Chen Y, Bittner ML, et al.: **Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer.** *Proc Natl Acad Sci USA* 2001, **98**:5711-5716.
38. Montagna C, Lyu MS, Hunter K, Lukes L, Lowther W, Reppert T, Hissong B, Weaver Z, Ried T: **The Septin 9 (MSF) gene is amplified and overexpressed in mouse mammary gland adenocarcinomas and human breast cancer cell lines.** *Cancer Res* 2003, **63**:2179-2187.
39. Nathanson KL, Shugart YY, Omaruddin R, Szabo C, Goldgar D, Rebbeck TR, Weber BL: **CGH-targeted linkage analysis reveals a possible BRCA1 modifier locus on chromosome 5q.** *Hum Mol Genet* 2002, **11**:1327-1332.
40. Orsetti B, Nugoli M, Cervera N, Lasorsa L, Chuchana P, Ursule L, Nguyen C, Redon R, du Manoir S, Rodriguez C, Theillet C: **Genomic and expression profiling of chromosome 17 in breast cancer reveals complex patterns of alterations and novel candidate genes.** *Cancer Res* 2004, **64**:6453-6460.
41. Buckley PG, Mantripragada KK, Piotrowski A, Diaz de Stahl T, Dumanski JP: **Copy-number polymorphisms: mining the tip of an iceberg.** *Trends Genet* 2005, **21**:315-317.
42. 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Bernards R, Friend SH: **Expression profiling predicts outcome in breast cancer.** *Breast Cancer Res* 2003, **5**:57-58.
43. Lonning PE, Sorlie T, Borresen-Dale AL: **Genomics in breast cancer-therapeutic implications.** *Nat Clin Pract Oncol* 2005, **2**:26-33.
44. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, et al.: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci USA* 2003, **100**:8418-8423.
45. Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, Ringner M, Sauter G, Monni O, Elkhouloun A, et al.: **Impact of DNA amplification on gene expression patterns in breast cancer.** *Cancer Res* 2002, **62**:6240-6245.
46. Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, Brown PO: **Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors.** *Proc Natl Acad Sci USA* 2002, **99**:12963-12968.
47. Nessling M, Richter K, Schwaenen C, Roerig P, Wrobel G, Wessendorf S, Fritz B, Bentz M, Sinn HP, Radlwimmer B, et al.: **Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue.** *Cancer Res* 2005, **65**:439-447.

48. Isola J, DeVries S, Chu L, Ghazvini S, Waldman F: **Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples.** *Am J Pathol* 1994, **145**:1301-1308.
49. Ried T, Just KE, Holtgreve-Grez H, du Manoir S, Speicher MR, Schrock E, Latham C, Blegen H, Zetterberg A, Cremer T: **Comparative genomic hybridization of formalin-fixed, paraffin-embedded breast tumors reveals different patterns of chromosomal gains and losses in fibroadenomas and diploid and aneuploid carcinomas.** *Cancer Res* 1995, **55**:5415-5423.
50. Speicher MR, du Manoir S, Schrock E, Holtgreve-Grez H, Schoell B, Lengauer C, Cremer T, Ried T: **Molecular cytogenetic analysis of formalin-fixed, paraffin-embedded solid tumors by comparative genomic hybridization after universal DNA-amplification.** *Hum Mol Genet* 1993, **2**:1907-1914.
51. Paik S, Kim CY, Song YK, Kim WS: **Technology insight: Application of molecular techniques to formalin-fixed paraffin-embedded tissues from breast cancer.** *Nat Clin Pract Oncol* 2005, **2**:246-254.
52. Lage JM, Leamon JH, Pejovic T, Hamann S, Lacey M, Dillon D, Segraves R, Vossbrinck B, Gonzalez A, Pinkel D, et al.: **Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-CGH.** *Genome Res* 2003, **13**:294-307.
53. Pandis N, Jin Y, Gorunova L, Petersson C, Bardi G, Idvall I, Johansson B, Ingvar C, Mandahl N, Mitelman F: **Chromosome analysis of 97 primary breast carcinomas: identification of eight karyotypic subgroups.** *Genes Chromosomes Cancer* 1995, **12**:173-185.
54. Pandis N, Idvall I, Bardi G, Jin Y, Gorunova L, Mertens F, Olsson H, Ingvar C, Beroukas K, Mitelman F, et al.: **Correlation between karyotypic pattern and clinicopathologic features in 125 breast cancer cases.** *Int J Cancer* 1996, **66**:191-196.
55. van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ, et al.: **Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations.** *Cancer Res* 2005, **65**:822-827.
56. Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM: **Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors.** *Cancer Res* 2002, **62**:7110-7117.
57. Cingoz S, Altungoz O, Canda T, Saydam S, Aksakoglu G, Sakizli M: **DNA copy number changes detected by comparative genomic hybridization and their association with clinicopathologic parameters in breast tumors.** *Cancer Genet Cytogenet* 2003, **145**:108-114.
58. Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR, Greshock JD, Luts L, Olsson H, Rahman N, et al.: **Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization.** *Cancer Res* 2005, **65**:7612-7621.
59. Reis-Filho JS, Simpson PT, Jones C, Steele D, Mackay A, Irvani M, Fenwick K, Valgeirsson H, Lambros M, Ashworth A, et al.: **Pleomorphic lobular carcinoma of the breast: role of comprehensive molecular pathology in characterization of an entity.** *J Pathol* 2005, **207**:1-13.
60. Somiari SB, Shriver CD, He J, Parikh K, Jordan R, Hooke J, Hu H, Deyarmin B, Lubert S, Malicki L, et al.: **Global search for chromosomal abnormalities in infiltrating ductal carcinoma of the breast using array-comparative genomic hybridization.** *Cancer Genet Cytogenet* 2004, **155**:108-118.
61. Weaver Z, Montagna C, Xu X, Howard T, Gadina M, Brodie SG, Deng CX, Ried T: **Mammary tumors in mice conditionally mutant for Brca1 exhibit gross genomic instability and centrosome amplification yet display a recurring distribution of genomic imbalances that is similar to human breast cancer.** *Oncogene* 2002, **21**:5097-5107.
62. Gronwald J, Jauch A, Cybulski C, Schoell B, Bohm-Steuer B, Lener M, Grabowska E, Gorski B, Jakubowska A, Domagala W, et al.: **Comparison of genomic abnormalities between BRCAx and sporadic breast cancers studied by comparative genomic hybridization.** *Int J Cancer* 2005, **114**:230-236.
63. Tirkkonen M, Kainu T, Loman N, Johannsson OT, Olsson H, Barkardottir RB, Kallioniemi OP, Borg A: **Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer.** *Genes Chromosomes Cancer* 1999, **24**:56-61.
64. Goepfert TM, Adigun YE, Zhong L, Gay J, Medina D, Brinkley WR: **Centrosome amplification and overexpression of aurora A are early events in rat mammary carcinogenesis.** *Cancer Res* 2002, **62**:4115-4122.
65. Lo YL, Yu JC, Chen ST, Yang HC, Fann CS, Mau YC, Shen CY: **Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK.** *Int J Cancer* 2005, **115**:276-283.
66. Nyante SJ, DeVries S, Chen YY, Hwang ES: **Array-based comparative genomic hybridization of ductal carcinoma in situ and synchronous invasive lobular cancer.** *Hum Pathol* 2004, **35**:759-763.
67. Seo MY, Rha SY, Yang SH, Kim SC, Lee GY, Park CH, Yang WI, Ahn JB, Park BW, Chung HC: **The pattern of gene copy number changes in bilateral breast cancer surveyed by cDNA microarray-based comparative genomic hybridization.** *Int J Mol Med* 2004, **13**:17-24.
68. Shelley Hwang E, Nyante SJ, Yi Chen Y, Moore D, DeVries S, Korkola JE, Esserman LJ, Waldman FM: **Clonality of lobular carcinoma in situ and synchronous invasive lobular carcinoma.** *Cancer* 2004, **100**:2562-2572.
69. Teixeira MR, Ribeiro FR, Torres L, Pandis N, Andersen JA, Lothe RA, Heim S: **Assessment of clonal relationships in ipsilateral and bilateral multiple breast carcinomas by comparative genomic hybridisation and hierarchical clustering analysis.** *Br J Cancer* 2004, **91**:775-782.
70. Wa CV, DeVries S, Chen YY, Waldman FM, Hwang ES: **Clinical application of array-based comparative genomic hybridization to define the relationship between multiple synchronous tumors.** *Mod Pathol* 2005, **18**:591-597.
71. Waldman FM, DeVries S, Chew KL, Moore DH, Kerlikowske K, Ljung BM: **Chromosomal alterations in ductal carcinomas in situ and their in situ recurrences.** *J Natl Cancer Inst* 2000, **92**:313-320.
72. Rennstam K, Ahlstedt-Soini M, Baldetorp B, Bendahl PO, Borg A, Karhu R, Tanner M, Tirkkonen M, Isola J: **Patterns of chromosomal imbalances defines subgroups of breast cancer with distinct clinical features and prognosis. A study of 305 tumors by comparative genomic hybridization.** *Cancer Res* 2003, **63**:8861-8868.
73. Jones C, Ford E, Gillett C, Ryder K, Merrett S, Reis-Filho JS, Fulford LG, Hanby A, Lakhani SR: **Molecular cytogenetic identification of subgroups of grade III invasive ductal breast carcinomas with different clinical outcomes.** *Clin Cancer Res* 2004, **10**:5988-5997.
74. Blegen H, Will JS, Ghadimi BM, Nash HP, Zetterberg A, Auer G, Ried T: **DNA amplifications and aneuploidy, high proliferative activity and impaired cell cycle control characterize breast carcinomas with poor prognosis.** *Anal Cell Pathol* 2003, **25**:103-114.
75. Barlund M, Tirkkonen M, Forozan F, Tanner MM, Kallioniemi O, Kallioniemi A: **Increased copy number at 17q22-q24 by CGH in breast cancer is due to high-level amplification of two separate regions.** *Genes Chromosomes Cancer* 1997, **20**:372-376.
76. Martinez A, Walker RA, Shaw JA, Dearing SJ, Maher ER, Latif F: **Chromosome 3p allele loss in early invasive breast cancer: detailed mapping and association with clinicopathological features.** *Mol Pathol* 2001, **54**:300-306.
77. Alvarez S, Diaz-Uriarte R, Osorio A, Barroso A, Melchor L, Paz MF, Honrado E, Rodriguez R, Urioste M, Valle L, et al.: **A predictor based on the somatic genomic changes of the BRCA1/BRCA2 breast cancer tumors identifies the non-BRCA1/BRCA2 tumors with BRCA1 promoter hypermethylation.** *Clin Cancer Res* 2005, **11**:1146-1153.
78. Climent J, Martinez-Climent JA, Blesa D, Garcia-Barchino MJ, Saez R, Sanchez-Izquierdo D, Azagra P, Lluch A, Garcia-Conde J: **Genomic loss of 18p predicts an adverse clinical outcome in patients with high-risk breast cancer.** *Clin Cancer Res* 2002, **8**:3863-3869.
79. Sotiropoulos C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET: **Breast cancer classification and prognosis based on gene expression profiles from a population-based study.** *Proc Natl Acad Sci USA* 2003, **100**:10393-10398.
80. Rhiem K, Klein A, Munch M, Kreutzfeld R, Ramser J, Wardelmann E, Schackert G, Von Deimling A, Wiestler OD, Schmutzler RK: **Chromosomal region 15q21.1 is a frequent target of allelic**

- imbalance in advanced breast carcinomas. *Int J Cancer* 2003, **106**:74-77.
81. Perez-Losada J, Mao JH, Balmain A: **Control of genomic instability and epithelial tumor development by the p53-Fbxw7/Cdc4 pathway.** *Cancer Res* 2005, **65**:6488-6492.
 82. Balmain A, Gray J, Ponder B: **The genetics and genomics of cancer.** *Nat Genet* 2003, **Suppl**:238-244.
 83. Garnis C, Coe BP, Lam SL, MacAulay C, Lam WL: **High-resolution array CGH increases heterogeneity tolerance in the analysis of clinical samples.** *Genomics* 2005, **85**:790-793.
 84. Barrett MT, Scheffer A, Ben Dor A, Sampas N, Lipsen D, Kincaid R, Tsang P, Curry B, Baird K, Meltzer PS, et al.: **Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA.** *Proc Natl Acad Sci USA* 2004, **101**:17765-17770.
 85. Ramaswamy S, Golub TR: **DNA microarrays in clinical oncology.** *J Clin Oncol* 2002, **20**:1932-1941.
 86. Snijders AM, Nowak N, Segreaves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, et al.: **Assembly of microarrays for genome-wide measurement of DNA copy number.** *Nat Genet* 2001, **29**:263-264.
 87. Greshock J, Naylor TL, Margolin A, Diskin S, Cleaver SH, Futreal PA, deJong PJ, Zhao S, Liebman M, Weber BL: **1-Mb resolution array-based comparative genomic hybridization using a BAC clone set optimized for cancer gene analysis.** *Genome Res* 2004, **14**:179-187.
 88. Ishkanian AS, Malloff CA, Watson SK, deLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, et al.: **A tiling resolution DNA microarray with complete coverage of the human genome.** *Nat Genet* 2004, **36**:299-303.
 89. Baudis M, Cleary ML: **Progenetix.net: an online repository for molecular cytogenetic aberration data.** *Bioinformatics* 2001, **17**:1228-1229.
 90. [progenetix] - cytogenetic abnormalities in human cancer [<http://www.progenetix.net/progenetix/Aboutprogenetix.html>]
 91. **SKY/M-FISH and CGH Database** [<http://www.ncbi.nlm.nih.gov/sky/skyweb.cgi>]
 92. **Online CGH Tumor Database** [<http://amba.charite.de/~ksch/cghdatabase/index.htm>]
 93. Cowell JK, LaDuca J, Rossi MR, Burkhardt T, Nowak NJ, Matsui S: **Molecular characterization of the t(3, 9) associated with immortalization in the MCF10A cell line.** *Cancer Genet Cytogenet* 2005, **163**:23-29.
 94. Vanden Bempt I, Vanhentenrijk V, Drijkoningen M, Wolf-Peters CD: **Comparative expressed sequence hybridization reveals differential gene expression in morphological breast cancer subtypes.** *J Pathol* 2006, **208**:486-494.
 95. Ray ME, Yang ZQ, Albertson D, Kleer CG, Washburn JG, Macoska JA, Ethier SP: **Genomic and expression analysis of the 8p11-12 amplicon in human breast cancer cell lines.** *Cancer Res* 2004, **64**:40-47.
 96. Callagy G, Pharoah P, Chin SF, Sangan T, Daigo Y, Jackson L, Caldas C: **Identification and validation of prognostic markers in breast cancer with the complementary use of array-CGH and tissue microarrays.** *J Pathol* 2005, **205**:388-396.
 97. Yiu R, Qiu H, Lee SH, Garcia-Aguilar J: **Mechanisms of microsatellite instability in colorectal cancer patients in different age groups.** *Dis Colon Rectum* 2005, **48**:2061-2069.
 98. Pathak S, Multani AS, Furlong CL, Sohn SH: **Telomere dynamics, aneuploidy, stem cells, and cancer (review).** *Int J Oncol* 2002, **20**:637-641.
 99. Lu YJ, Williamson D, Wang R, Summersgill B, Rodriguez S, Rogers S, Pritchard-Jones K, Campbell C, Shipley J: **Expression profiling targeting chromosomes for tumor classification and prediction of clinical behavior.** *Genes Chromosomes Cancer* 2003, **38**:207-214.
 100. Al Mulla F, Al Maghrebi M, Varadharaj G: **Expressive genomic hybridisation: gene expression profiling at the cytogenetic level.** *Mol Pathol* 2003, **56**:210-217.
 101. Dabbs DJ, Carter G, Fudge M, Peng Y, Swalsky P, Finkelstein S: **Molecular alterations in columnar cell lesions of the breast.** *Mod Pathol* 2006, **19**:344-349.
 102. Lips EH, Dierssen JW, van Eijk R, Oosting J, Eilers PH, Tollenaar RA, de Graaf EJ, van't Slot R, Wijmenga C, Morreau H, et al.: **Reliable high-throughput genotyping and loss-of-heterozygosity detection in formalin-fixed, paraffin-embedded tumors using single nucleotide polymorphism arrays.** *Cancer Res* 2005, **65**:10188-10191.
 103. Ting JC, Ye Y, Thomas GH, Ruczinski I, Pevsner J: **Analysis and visualization of chromosomal abnormalities in SNP data with SNPscan.** *BMC Bioinformatics* 2006, **7**:25.
 104. Huang J, Wei W, Zhang J, Liu G, Bignell GR, Stratton MR, Futreal PA, Wooster R, Jones KW, Shapero MH: **Whole genome DNA copy number changes identified by high density oligonucleotide arrays.** *Hum Genomics* 2004, **1**:287-299.
 105. Murthy SK, DiFrancesco LM, Ogilvie RT, Demetrick DJ: **Loss of heterozygosity associated with uniparental disomy in breast carcinoma.** *Mod Pathol* 2002, **15**:1241-1250.
 106. Packeisen J, Nakachi K, Boecker W, Brandt B, Buerger H: **Cytogenetic differences in breast cancer samples between German and Japanese patients.** *J Clin Pathol* 2005, **58**:1101-1103.
 107. Varma G, Varma R, Huang H, Pryshchepava A, Groth J, Fleming D, Nowak NJ, McQuaid D, Conroy J, Mahoney M, et al.: **Array comparative genomic hybridisation (aCGH) analysis of premenopausal breast cancers from a nuclear fallout area and matched cases from Western New York.** *Br J Cancer* 2005, **93**:699-708.
 108. Lai Y, Zhao H: **A statistical method to detect chromosomal regions with DNA copy number alterations using SNP-array-based CGH data.** *Comput Biol Chem* 2005, **29**:47-54.
 109. Willmarth NE, Albertson DG, Ethier SP: **Chromosomal instability and lack of cyclin E regulation in hCdc4 mutant human breast cancer cells.** *Breast Cancer Res* 2004, **6**:R531-R539.
 110. Naylor TL, Greshock J, Wang Y, Colligon T, Yu QC, Clemmer V, Zaks TZ, Weber BL: **High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization.** *Breast Cancer Res* 2005, **7**:R1186-R1198.