

# Unlocking pathology archives for molecular genetic studies: a reliable method to generate probes for chromogenic and fluorescent *in situ* hybridization

Maryou BK Lambros<sup>1</sup>, Pete T Simpson<sup>2</sup>, Chris Jones<sup>3</sup>, Rachael Natrajan<sup>3</sup>, Charlotte Westbury<sup>1</sup>, Dawn Steele<sup>1</sup>, Kay Savage<sup>1</sup>, Alan Mackay<sup>1</sup>, Fernando C Schmitt<sup>4</sup>, Alan Ashworth<sup>1</sup> and Jorge S Reis-Filho<sup>1</sup>

<sup>1</sup>The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London, UK; <sup>2</sup>Molecular and Cellular Pathology, Mayne Medical School, University of Queensland, Queensland Institute of Medical Research and Royal Brisbane and Women's Hospital, Brisbane, Australia; <sup>3</sup>Section of Paediatric Oncology, Institute of Cancer Research, Sutton, UK and <sup>4</sup>IPATIMUP—Institute of Molecular Pathology and Immunology, Portugal and Medical Faculty, University of Porto, Porto, Portugal

**Chromogenic (CISH) and fluorescent (FISH) *in situ* hybridization have emerged as reliable techniques to identify amplifications and chromosomal translocations. CISH provides a spatial distribution of gene copy number changes in tumour tissue and allows a direct correlation between copy number changes and the morphological features of neoplastic cells. However, the limited number of commercially available gene probes has hindered the use of this technique. We have devised a protocol to generate probes for CISH that can be applied to formalin-fixed, paraffin-embedded tissue sections (FFPETS). Bacterial artificial chromosomes (BACs) containing fragments of human DNA which map to specific genomic regions of interest are amplified with  $\phi$ 29 polymerase and random primer labelled with biotin. The genomic location of these can be readily confirmed by BAC end pair sequencing and FISH mapping on normal lymphocyte metaphase spreads. To demonstrate the reliability of the probes generated with this protocol, four strategies were employed: (i) probes mapping to cyclin D1 (*CCND1*) were generated and their performance was compared with that of a commercially available probe for the same gene in a series of 10 FFPETS of breast cancer samples of which five harboured *CCND1* amplification; (ii) probes targeting cyclin-dependent kinase 4 were used to validate an amplification identified by microarray-based comparative genomic hybridization (aCGH) in a pleomorphic adenoma; (iii) probes targeting fibroblast growth factor receptor 1 and *CCND1* were used to validate amplifications mapping to these regions, as defined by aCGH, in an invasive lobular breast carcinoma with FISH and CISH; and (iv) gene-specific probes for *ETV6* and *NTRK3* were used to demonstrate the presence of t(12;15)(p12;q25) translocation in a case of breast secretory carcinoma with dual colour FISH. In summary, this protocol enables the generation of probes mapping to any gene of interest that can be applied to FFPETS, allowing correlation of morphological features with gene copy number.**

*Laboratory Investigation* (2006) 86, 398–408. doi:10.1038/labinvest.3700390; published online 30 January 2006

**Keywords:** chromogenic *in situ* hybridization; fluorescent *in situ* hybridization; comparative genomic hybridization; molecular pathology; protocol

Chromosomal aberrations, such as gains/losses/amplifications of whole chromosomes or chromosome parts, are hallmarks of tumourigenesis and

tumour progression.<sup>1–3</sup> The completion of the human genome project and the advent of high throughput molecular genetic methods for genome-wide analysis have unveiled a new era for cancer research. Array-based comparative genomic hybridization (aCGH) is a powerful screening tool to detect DNA copy changes and hence to identify cancer-associated chromosomal aberrations.<sup>2,4–9</sup> The high resolution of aCGH has been demonstrated in several contexts, particularly in the case of small

Correspondence: Dr JS Reis-Filho, MD, The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK.

E-mail: jorgerf@icr.ac.uk

Received 14 November 2005; revised 19 December 2005; accepted 20 December 2005; published online 30 January 2006

amplicons, previously undetected by chromosomal CGH (cCGH), which have been identified in a number of cell lines and human tumours.<sup>2,4–9</sup> Although the principles of aCGH are rather simple, it is an expensive and labour-intensive method. Furthermore, while our group and others have successfully performed aCGH with DNA extracted from formalin-fixed archival material,<sup>2,4–6,8,9</sup> this is associated with considerable difficulties. Therefore, it is unlikely that this technique will be easily applied to large cohorts of archival samples.

The genomic changes identified by cCGH and aCGH reflect the pattern of changes observed in the dominant clonal population of a given tumour. Hence, changes found in less prevalent clonal populations of tumour cells are less likely to be identified with these techniques. Given that there is evidence to suggest that most tumours are heterogeneous,<sup>10–12</sup> even with the increased resolution of aCGH, some genetic abnormalities that may drive tumour behaviour may go undetected, for example, when a given change is restricted to <10–20% of neoplastic cells.

*In situ* hybridization (ISH) is a useful alternative for demonstrating the presence of gene-specific copy number gains in tumours and to validate aCGH results.<sup>4,9</sup> Although this methodology allows only one or two chromosomal regions to be studied at a time, it does permit a thorough assessment of the spatial distribution of a given copy number change within the tumour bulk and a correlation with morphological and immunohistochemical findings. In addition, with the advent of tissue microarrays, ISH experiments can now be performed in a high throughput fashion.<sup>13–16</sup> This has been elegantly demonstrated by several groups using both fluorescent (FISH)<sup>13,14</sup> and chromogenic ISH (CISH).<sup>15,16</sup> Although FISH provides accurate copy number assessment, this technique is not ideally suited for TMAs, as fluorescent microscopy does not allow an optimal correlation between copy numbers and cytomorphological aspects of the cells and it proves difficult and time consuming to navigate through the tissue microarray section under a fluorescent microscope.

CISH has proven to be as sensitive and specific as FISH in detecting amplification and high level gene copy number gains.<sup>17–19</sup> However, only a few CISH probes are currently commercially available, restricting the use of this technique to analyse different genetic loci. In this study, we describe a protocol for generating CISH probes from a pool of up to three overlapping bacterial artificial chromosome (BAC) clones. These probes can be hybridized to cell line interphase nuclei and metaphase spreads, and most importantly to both frozen and formalin-fixed, paraffin-embedded tumour samples. Coupled with a quick dual labelling FISH method or a dual colour CISH detection system, these probes can also be applied to formalin-fixed, paraffin-embedded tissue sections (FFPETS) to assess ba-

lanced chromosomal translocations and co-amplifications.

## Materials and methods

### Samples

In all, 10 cases of invasive ductal breast carcinoma of no special type and one case of invasive secretory carcinoma of the breast were retrieved from the files of the Royal Marsden Hospital, with approval of the local Ethics Committee. A case of invasive lobular breast carcinoma and a case of pleomorphic adenoma were retrieved from the consultation files of one of the authors (JSRF).

### BAC Selection

Publicly available BAC clones were selected from Ensembl genome browser ([http://www.ensembl.org/Homo\\_sapiens/](http://www.ensembl.org/Homo_sapiens/)) mapped to the latest human genome assembly (May 2004 build of the human genome sequence, hg17) and obtained from the FISH mapped 32K rearray set of BAC clones from Children's Hospital Oakland Research Institute (<http://bacpac.chori.org/>).

### BAC DNA Extraction

BAC DNA was extracted using Qiagen Plasmid Mini Kit (Qiagen, Crawley, West Sussex, UK) according to the manufacturer's instructions. The QIAGEN Anion-Exchange Resin column (Qiagen-tip) was used to purify the extracted DNA, which was then desalted in isopropanol and washed with 70% ethanol. Finally, the purified DNA was briefly air dried and resuspended in 22  $\mu$ l of prewarmed Tris-Cl pH 8.5. The final concentration of the BAC DNA ranged from 50 to 200 ng/ $\mu$ l.

### BAC DNA Amplification

Amplification of BAC DNA was performed using the GenomiPhi whole genome amplification kit (WGA kit, GE Healthcare, UK) according to the manufacturer's instructions.<sup>20–22</sup> In all, 10 ng of BAC DNA (1  $\mu$ l) was added to 9  $\mu$ l of the sample buffer before the mixture was incubated at 95°C for 3 min and then cooled on ice for 5 min. In all, 10  $\mu$ l of  $\phi$ 29 polymerase reaction mix (1  $\mu$ l of  $\phi$ 29 polymerase in 9  $\mu$ l of reaction buffer) was added, and the mixture was then incubated at 30°C for  $\geq$ 16 h. To inactivate the enzyme the mixture was heated for 10 min at 65°C. The DNA was precipitated and washed with 70% ethanol. Finally, the concentration of the amplified DNA was measured with a spectrophotometer, yielding approximately 5–7  $\mu$ g of DNA.

## Probe End Sequencing

Each sequencing reaction was performed from 1  $\mu\text{g}$  GenomPhi-amplified BAC DNA, 0.5  $\mu\text{l}$  of 20  $\mu\text{M}$  SP6 or T7 primers (0.66 mM), 6  $\mu\text{l}$  BigDye terminator version 1.1 (Applied Biosystems, Foster City, CA, USA) and ultrafiltered water to a final volume of 15  $\mu\text{l}$ . The sequencing reaction was carried out on a thermocycler according to the following conditions: 96°C for 4 min followed by 100 cycles at 96°C for 10 s, 50°C for 10 s and 60°C for 4 min. Samples were ethanol precipitated at room temperature (RT) (15  $\mu\text{l}$  sequencing reaction, 21  $\mu\text{l}$  ultrafiltered water, 64  $\mu\text{l}$  100% ethanol) for 15 min and then centrifuged at 14 000 r.p.m. for 30 min. Pellets were washed with 500  $\mu\text{l}$  of 70% ethanol, air dried, resuspended in 10  $\mu\text{l}$  of Hi-Di formamide (Applied Biosystems, Foster City, CA, USA) and sequenced on a Biosystems 3100 Genetic Analyser (Applied Biosystems). Data were analysed with the Applied Biosystems 3100 Genetic Analyser. Traces were screened by Sequencher (Gene Codes Corporation, Ann Arbor, MI, USA) and Autoassembler (Applied Biosystems) and raw sequence data were processed with Perl. The genomic locus of each BAC clone was determined according to the UCSC BAC End Pair algorithm. Alternatively, each BAC clone sequence was manually analysed on UCSC Golden Path genome browser (<http://genome.ucsc.edu/index.html>).

## Probe Labelling

To create probes, BAC DNA was labelled with either biotin or digoxigenin (DIG) using the Bioprime kit (Invitrogen Life Technologies, UK) according to the manufacturer's instructions with minor modifications. For biotin-labelled probes, a 44  $\mu\text{l}$  reaction mixture was prepared by adding 20  $\mu\text{l}$  of random primer to 300 ng of GenomPhi-amplified BAC DNA. For DIG labelling, 42.25  $\mu\text{l}$  of reaction mixture was prepared by adding 20  $\mu\text{l}$  of random primer to 300 ng of BAC DNA. For cases where more than one BAC was used to create the probe, the 300 ng of DNA was composed of a pool of equal amounts of each BAC used. The DNA mixture was then denatured at 95°C for 5 min and snap frozen on ice for 10 min. In all, 5  $\mu\text{l}$  of 10  $\times$  dNTP biotinylate mixture (Bioprime Kit, Invitrogen) and 1  $\mu\text{l}$  of Klenow DNA polymerase were then added and the mixture was incubated at 37°C for 3 h. For the DIG-labelling method, 1.75  $\mu\text{l}$  DIG-11-dUTP (pH 7.5) (Roche Applied Science, Mannheim, Germany), 5  $\mu\text{l}$  of 10  $\times$  dNTP mixture (1 mM dATP, 1 mM dGTP, 1 mM dCTP and 0.65 mM dTTP) and 1  $\mu\text{l}$  of Klenow DNA polymerase were then added and the mixture was incubated at 37°C for 3 h. Of stop buffer, 5  $\mu\text{l}$  was added and the probes were purified using Microspin G50 columns (Amersham Biosciences, UK), ethanol precipitated with 30  $\mu\text{g}$  human Cot-1 DNA (Invitrogen) and 10  $\mu\text{g}$  salmon sperm DNA (Invitrogen). The DNA pellets

were dried at 37°C for 15–30 min, resuspended in 30  $\mu\text{l}$  of Tris buffer (pH 7.5) and the DNA concentrations assessed by spectrophotometry. For 12 probes created, DNA concentrations ranged from 168 to 442 ng/ $\mu\text{l}$  (median = 304 ng/ $\mu\text{l}$ , mean = 293.2  $\pm$  76.3 ng/ $\mu\text{l}$ ).

After performing the optimization steps, where FISH was being carried out, probes were resuspended in 12  $\mu\text{l}$  of hybridization buffer I (50% formamide; 20% dextran sulphate; 2  $\times$  SC, 0.1 mM EDTA, pH 8.0; 0.2 Mm Tris-HCl, pH 7.6). For CISH, probes were resuspended in 80  $\mu\text{l}$  of hybridization buffer II (60% deionized formamide; 12% dextran sulphate; 2.4  $\times$  SSC, 0.14 mM EDTA, pH 8.0; 0.4 mg/ml of salmon sperm DNA).

## Probe FISH-Mapping

BAC clones were FISH-mapped to ensure that each probe specifically hybridized to only the expected genomic regions. Briefly, 12  $\mu\text{l}$  of the resuspended probe was denatured at 75°C for 10 min followed by preannealing at 37°C for 30 min. The denatured probe was applied to the centre of a 22  $\times$  50 mm coverslip, placed on a slide with a denatured preparation of metaphase spreads (Vysis Inc., Downers Grove, USA) and hybridized overnight at 37°C in a humidified chamber. After hybridization, slides were washed in 0.4  $\times$  SSC/0.3% Igepal CA-630 (Sigma-Aldrich, Poole, UK) for 2 min at 73°C, two washes in 2  $\times$  SSC/0.1% Igepal CA-630 for 5 min each at RT followed by one wash with water for 3 min and two washes in PBS for 3 min. 1  $\mu\text{l}$  of Avidin-cy3 (Sigma, UK) was diluted 1:200 in PBS/0.5% BSA and the mixture was added to the slide and incubated in the dark for 40 min at RT. The slides were dehydrated through a series of alcohols and mounted in an antifade medium (Vector laboratories, Peterborough, UK) containing 4,6-diamino-2-phenylindole (DAPI) as a counterstain. The probes were visualized using a Zeiss Axioskop microscope and a COHU camera, and analysed with Cytovision software (Applied Imaging International, Newcastle Upon Tyne, UK).

## Optimization of the CISH Protocol

To test the intensity of the CISH probe signal and to reduce the nonspecific background staining, a number of optimization steps were carried out.<sup>23</sup> Factors known to affect the ability of the probe to hybridize specifically to a single genomic region and to reduce nonspecific background staining and nonspecific signals were tested, including the deparaffinization (dewaxing) method, digestion time, probe concentration, hybridization buffer and posthybridization washes.<sup>23</sup> The effect of the number of overlapping BACs on the signal size was also investigated.

Suboptimal dewaxing is an important factor that may hinder probe hybridization. Most preparations

of paraffin melt at 58°C therefore dewaxing methods using xylene washes of different lengths (5, 10, 20, 30 and 40 min) at both RT, 42 and 58°C were analysed. Suboptimal proteolytic digestion (over or under digestion) with pepsin may lead to absence of signals, loss of morphological detail and/or reduced specificity of the probe (eg, multiple signals of varying sizes in normal cells, nonspecific spots juxtaposed to the nuclear envelope and increased background<sup>23</sup>). Therefore, we tested varying digestion times between 4 and 10 min. These two steps were optimized using commercially available probes for epidermal growth factor receptor (*EGFR* SpotLight *EGFR* amplification probe, Zymed) and *MYC* (*CMYC* amplification probe, Zymed).

The effect of different probe concentrations (1, 10, 20, 50, 100, 200 and 400 ng/μl) on the size and intensity of the probe signal and the level of background staining was analysed with probes mapping to 8p12 (composed of BACs RP11-350N15, RP11-148D21 and RP11-359P11) and 12q13 (RP11-66N19, RP11-277A02 and RP11-549D07). Both the hybridization conditions and the stringency of the posthybridization washes may also affect the background levels, the presence of nonspecific spots and the size and intensity of the actual signals.<sup>23</sup> First, the effect of using different hybridization buffers was evaluated. Three were tested: (i) hybridization buffer I (50% formamide; 20% dextran sulphate; 2 × SSC, 0.1 mM EDTA, pH 8.0; 0.2 mM Tris-HCl, pH 7.6), an in house buffer used for CGH analysis and FISH mapping; (ii) hybridization buffer II (60% deionized formamide; 12% dextran sulphate; 2.4 × SSC, 0.14 mM EDTA, pH 8.0; 0.4 mg/ml of salmon sperm DNA), which is similar to a commercially available hybridization mix reported to show optimal results for CISH on FFPE tissue sections<sup>23</sup> and (iii) the Sigma hybridization solution (H-7782; Sigma-Aldrich, Poole, UK). After selecting the optimal hybridization buffer, two different posthybridization washes were compared, a formamide wash (50% formamide in 2 × SSC at 45°C) and a nonformamide wash (0.5 SSC at RT). Subsequently, the size and intensity of the detected probe signal were also investigated using probes mapping to the same genomic region (8p12) containing one (RP11-350N15), two (RP11-350N15 and RP11-148D21) and three BACs (RP11-350N15, RP11-148D21 and RP11-359P11).

### Chromogenic *In Situ* Hybridization

Representative FFPETS (2–4 μm) were mounted on polylysine-coated slides and subjected to CISH analysis. Dewaxed sections were heated with CISH pretreatment buffer (SPOT-light tissue pretreatment kit, Zymed, South San Francisco, CA, USA) on a hotplate for 15 min at ≥98°C as described in the manufacturer's protocol. The treated sections were then washed with water before being digested with pepsin (Zymed) for up to 10 min, depending on the

tissue type and its fixation (see Results). Subsequently, sections were washed with water, dehydrated with graded ethanol, and air dried.

The hybridization process was carried out by applying the resuspended probe to the centre of a 22 × 50 mm<sup>2</sup> cover slip, placing the cover slip onto the tissue section and then sealing the edges of the cover slip with rubber cement. The slides were placed on a hot plate at 97°C for 5–10 min to denature the tissue sections and probes, before hybridizing overnight in a humidified chamber at 37°C. Coverslips were removed and slides were rinsed in 0.5 × SSC at RT and then washed in 0.5 × SSC for 5 min at 78°C. Endogenous peroxidase activity and nonspecific protein binding were blocked by incubation with 3% H<sub>2</sub>O<sub>2</sub> in methanol for 10 min and CAS-block (Zymed) for 10 min, respectively. HRP-streptavidin (Zymed) was applied to the tissue section for 30 min at RT and then washed in PBS/0.025% Tween 20. For DIG-labelled probes, mouse antiDIG antibody (anti-DIG) was added before incubation with HRP-streptavidin-goat (anti-mouse). Development of the immunohistochemical staining was completed using DAB (Zymed) as the chromogen. Tissue sections were lightly counterstained with haematoxylin and then mounted. The CISH results were analysed on a light microscope. Signals were counted at ×400 and ×630 final magnifications.

### CISH Probe Validation

To assess the specificity and sensitivity of our in house-produced probes, we compared our probe for cyclin D1 (*CCND1*) to a commercially available probe for the same gene (Zymed). Our CISH probe consisted of three overlapping BAC clones (RP11-681H17, RP11-699M19 and RP11-775I17) mapping to the genomic region encompassing the *CCND1* gene on chromosome 11q13.3. Results were scored blindly by two of the authors (MBKL and JSRF).

### FISH Dual Labelling on FFPETS

Translocation and co-amplification are common features of cancer genomes. To map these chromosomal aberrations on FFPE tissue sections, a dual labelling FISH approach was performed. Two probe pairs were generated with our in house protocol: (i) fibroblast growth factor receptor 1 (*FGFR1*, Table 1), labelled with biotin and *CCND1* (Table 1), labelled with DIG and (ii) ETS variant gene 6 (*ETV6*, Table 1) labelled with biotin and neurotrophic tyrosine kinase receptor type 3 (*NTRK3*, Table 1) labelled with DIG. Each probe pair was precipitated together with 30 μg of human COT-1 and 10 μg of salmon sperm DNA. Slides were dewaxed, pretreated and left to hybridize overnight as stated in the CISH protocol. Tissue sections were rinsed in 0.5 × SSC at RT and then washed in 0.5 × SSC for 5 min at 78°C.

**Table 1** *In situ* hybridization probes: details on the cytogenetic band and bacterial artificial chromosomes (BACs)

Probe	Cytoband	BAC clones	BAC length (bp)	Probe size (bp)
<i>FGFR1</i>	8p12	RP11-350N15	200918	313937
		RP11-148D21	200374	
		RP11-359P11	175197	
<i>MYC</i>	8q24.21	RP11-440N18	181231	291231
		RP11-237F24	135280	
		CTD-2034C18	173138	
<i>CCND1</i>	11q13.3	RP11-699M19	147904	284428
		RP11-681H17	164791	
		RP11775I17	75099	
<i>ETV6</i>	12p13.2	RP11-434CI	210778	305585
		CTD-2335C03	103991	
		RP11-407P10	210778	
<i>NTRK3</i>	15q25.3	RP11-452K23	150370	198450
		RP11-730G13	40869	
		RP11-114I9	169686	
<i>CDK4/SAS</i>	12q14.1	RP11-66N19	150611	430617
		RP11-277A02	149338	
		RP11-549D07	174738	
<i>MDM2</i>	12q15	CTD-2538A02	185484	405652
		RP11-797C20	178739	
		RP11-675P21	167929	

BAC: bacterial artificial chromosome; bp: base pairs; *CDK4*: cyclin-dependent kinase 4; *CCND1*: Cyclin D1; *ETV6*: ETS translocation variant 6; *FGFR1*: fibroblast growth factor receptor 1; *MDM2*: double minute 2; *MYC*: Myc proto-oncogene; *NTRK3*: neurotrophic tyrosine kinase receptor type 3; *SAS*: sarcoma-amplified sequence.

Nonspecific protein binding was blocked by incubation with CAS-block (Zymed) for 10 min. The slides were then mounted with a mixture of 1  $\mu$ l of Avidin-cy3/1  $\mu$ l of anti-DIG-fluorescein diluted in 1:200 in PBS/0.5% BSA for 45 min in the dark at RT. The mixture was blotted off the slides and the slides were washed with PBS/0.025% Tween 20. The slides were counterstained with antifade medium containing DAPI, and images were collected sequentially in three channels (DAPI, FITC, Cy3) on a TCS SP2 confocal microscope (Leica, Milton Keynes, UK).

### Array Comparative Genomic Hybridization

For aCGH analysis, an invasive lobular carcinoma of the breast and a pleomorphic adenoma of the salivary gland were used. For each sample, two consecutive 5  $\mu$ m sections stained with nuclear fast red and microdissected with a sterile needle under a stereomicroscope (Olympus SZ61, Tokyo, Japan). Estimated purity of neoplastic cells in both samples was 90%. DNA was extracted in 200  $\mu$ l digestion buffer (100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 25 mM EDTA (pH 8.0), 0.5% SDS, 0.5 mg/ml Proteinase K (Invitrogen Life Technologies) and was purified by phenol/chloroform extraction in Phase Lock Gel Light tubes (Eppendorf, Hamburg, Germany). The purified DNA samples were ethanol precipitated and the DNA resuspended in TE buffer (pH 7.5). The DNA yield was assessed by spectro-

photometry and the quality (DNA fragment size range) by agarose gel electrophoresis.

Array CGH was performed using an array of 5623 BACs spaced at approximately 1 Mb intervals throughout the genome.<sup>4,9</sup> Array CGH was performed according to a previously described protocol.<sup>4,9</sup>

Slides were scanned using an Axon 4000B scanner (Axon Instruments, Burlingame, CA, USA) and images were analysed using Genepix Pro 4.1 software (Axon Instruments). The median localized background slide signal for each clone was subtracted and each clone Cy5/Cy3 ratio subjected to row Loess normalization.<sup>24</sup> BAC clone replicate spots were averaged, and clones with poor reproducibility between replicates (s.d. > 0.2) were excluded. In addition, clones with missing/poor values in > 70% samples were excluded, as were those with no mapping information (May 2004 build of the human genome sequence, hg17). The final quality-filtered data set comprised four hybridizations and 4606 clones. Data were smoothed using a local polynomial adaptive weights smoothing procedure for regression problems with additive errors.<sup>25</sup> A categorical analysis was applied to the BACs after classifying them as representing gain, loss, or no-change according to their smoothed log<sub>2</sub> ratio values. Log<sub>2</sub> ratio values < -0.09 were categorized as losses, those > 0.09 as gains, and those in between as unchanged. These threshold values were chosen to correspond to three s.d. of the normal ratios obtained from the filtered clones mapping to chromosomes 1–22, assessed in comparisons between DNA extracted from a pool of male and female blood donors (data not shown). All data transformation and statistical analysis was carried out in R 2.0.1 (<http://www.r-project.org/>) and BioConductor 1.5 (<http://www.bioconductor.org/>), making extensive use of modified versions of the packages, in particular aCGH, marray and aws.<sup>26</sup>

### Results

Optimization of the signal intensity and background are essential. Each signal should be strong enough to be visible at  $\times 400$  magnification and the background should be clear enough to avoid misinterpretation of nonspecific deposition of DAB as probe signals. Suboptimal conditions lead to (i) speckled background, (ii) inconspicuous signals and (iii) nonspecific spots juxtaposed to the intranuclear surface of the nuclear envelope, hindering the assessment of copy numbers. To improve the signal size and to reduce the background, we tested several parameters:

*Section dewaxing step:* The presence of discrete signals was found to be dependant on the dewaxing step. If sections were suboptimally dewaxed (< 3  $\times$  5–10 min in RT xylene), signals were significantly decreased or absent. Although signals were

obtained with the protocols recommended by one of the manufacturers (Zymed),<sup>27</sup> we observed that results were more consistent when slides were either dewaxed at 55°C for 15 min (3 × 5 min) or dewaxed at 42°C for 15–30 min before a second incubation for 5 min with xylene preheated to 58°C in a fume cupboard.

**Signal size and background optimization:** In order to attain strong signals with minimum background and following the guidelines published by Evans *et al*,<sup>23</sup> optimization of the following parameters was performed: proteolytic digestion time, probe concentration, hybridization buffers, posthybridization washes and number of BACs to make the probe.

**Proteolytic digestion:** We<sup>4,27</sup> and others<sup>15,18</sup> demonstrated that the tissue digestion solution from SPoT-Light Tissue Pretreatment Kit (Zymed) provides consistent results. For all commercially available probes tested, we observed that digestion for <4 min leads to total absence of signals in the majority of tissues, whereas digestion for >10 min results in poor morphological features and blurred signals. In our hands, 6 min at RT was the optimal condition for a number of different tissue types tested, including breast, brain and salivary gland, whereas 4 min at RT was the optimal condition for kidney sections. Our results suggest that proteolytic digestion has to be empirically optimized for each tissue type.

**Probe concentration:** Signals were weak and inconsistent when low concentrations (<20 ng/μl) of probe were used. A higher probe (>100 ng/μl) concentration leads to increased nonspecific binding and background. The best results (clear signals with negligible background) were obtained with a concentration between 50 and 100 ng/μl. Depending on the size of the section, a total volume ranging from 12 to 20 μl of resuspended probe was used.

**Hybridization buffer:** For CISH analysis, hybridization buffer I and the commercial hybridization solution (Sigma) invariably gave higher levels of background staining when compared to our hybridization buffer II. Therefore, the latter was used for further analysis.

**Posthybridization washes:** The level of background noise was assessed for both formamide- and nonformamide-containing washes. As reported by Evans *et al*,<sup>23</sup> no significant differences were seen between washes with and without formamide. Hence, a nonformamide (0.5 × SSC) wash was employed in the posthybridization washes.

**Number of BAC clones:** The signal size generated by the probe was directly proportional to the number of BAC clones used to construct the probe. Probes that consisted of one BAC clone, with an average size of 159 kb, showed small but discrete signals. When the number of overlapping BACs mapping to the same region was increased to 3, a final size ranging from 198 to 431 Kb (mean size = 319 kb, Table 1), signals were readily detected at × 400 magnification with little background noise.

Therefore, all of the CISH probes subsequently generated in house consisted of three overlapping BAC clones.

### CISH Probe Validation

The specificity and sensitivity of our probes for the *CCND1* gene were compared to a commercial probe from Zymed using tissue sections from 10 invasive ductal breast carcinomas, five of which were known to harbour *CCND1* amplification (Reis-Filho JS, unpublished results) and five with normal *CCND1* copy numbers (Figure 1). The agreement between the commercially available (Zymed) probe and our in house probe was perfect ( $\kappa=1$ ) (Figure 1). Subsequently, probes for *FGFR1*, cyclin-dependent kinase 4 (*CDK4*), *ETV6*, *NTRK3* and *MDM2*, genes were made (Table 1).

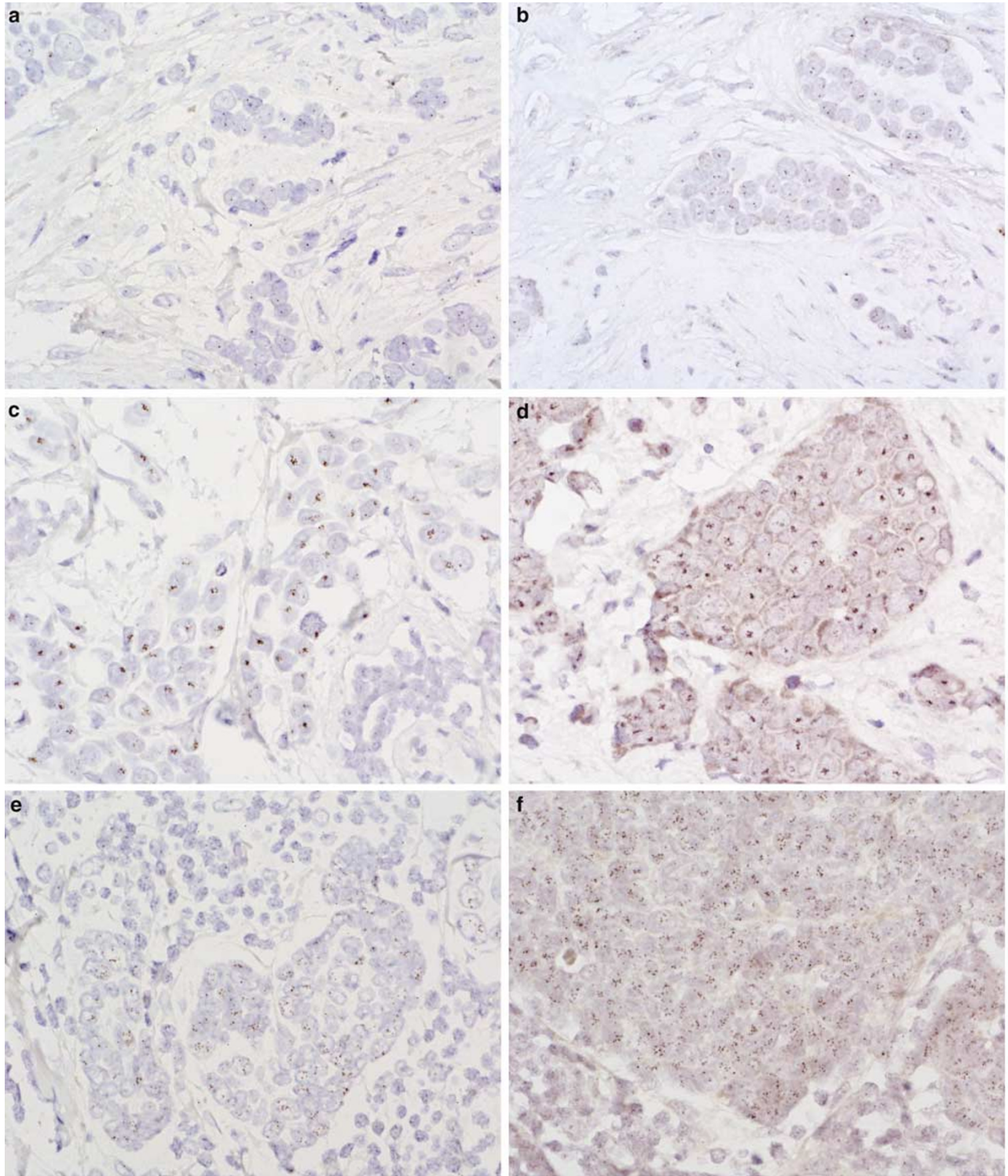
### Confirmation of Array CGH Results

The complete details of the genome profile of the cases of invasive lobular carcinoma and pleomorphic adenoma are described elsewhere<sup>4</sup> (and Reis-Filho *et al*, manuscript in preparation). The array CGH profile for the invasive lobular carcinoma showed high-level copy numbers in two regions, 8p12 (30 636.68–42 348.47 kb) and 11q13.3 (63 963.52–69 679.24 kb) (Figure 2). We have employed a simple dual FISH protocol using probes designed/mapping to the *FGFR1* (8p12, RP11-350N15, RP11-148D21 and RP11-359P11) and *CCND1* (11q13, RP11-681H17, RP11-699M19 and RP11-775I17) loci to validate this dual amplification on FFPETS. Images obtained with the confocal microscope demonstrated coamplification of *FGFR1* and *CCND1* (Table 1), corroborating the results of microarray CGH (Figure 2). The pleomorphic adenoma showed amplification of 12q13.3–q14.1 (55 277.47–62 196.196 kb) (Figure 3). Probes mapping to *CDK4/sarcoma-amplified sequence (CDK4/SAS)*, 56.4 Mb, RP11-66N19, RP11-277A02 and RP11-549D07) and mouse double minute homolog 2 (*MDM2*, 67.5 Mb, CTD-2538A02, RP11-797C20 and RP11-675P21) were applied to FFPE tissue sections, corroborating amplification of *CDK4/SAS* (Figure 3) and not the region distal to the amplicon (67.5 Mb).

### Detection of Balanced Chromosomal Translocations

Secretory carcinomas of the breast are reported to consistently harbour the balanced chromosomal translocation t(12;15)(p13;q25), involving the genes *ETV6* and *NTRK3*.<sup>28,29</sup> After generating probes for *ETV6* and *NTRK3* (Table 1)<sup>29</sup> labelled with biotin and DIG, respectively, as described above, we applied the probe pairs to FFPETS of a bona fide case of secretory carcinoma using our in house method for dual FISH analysis. Furthermore, these probes were hybridized to the same set of 10

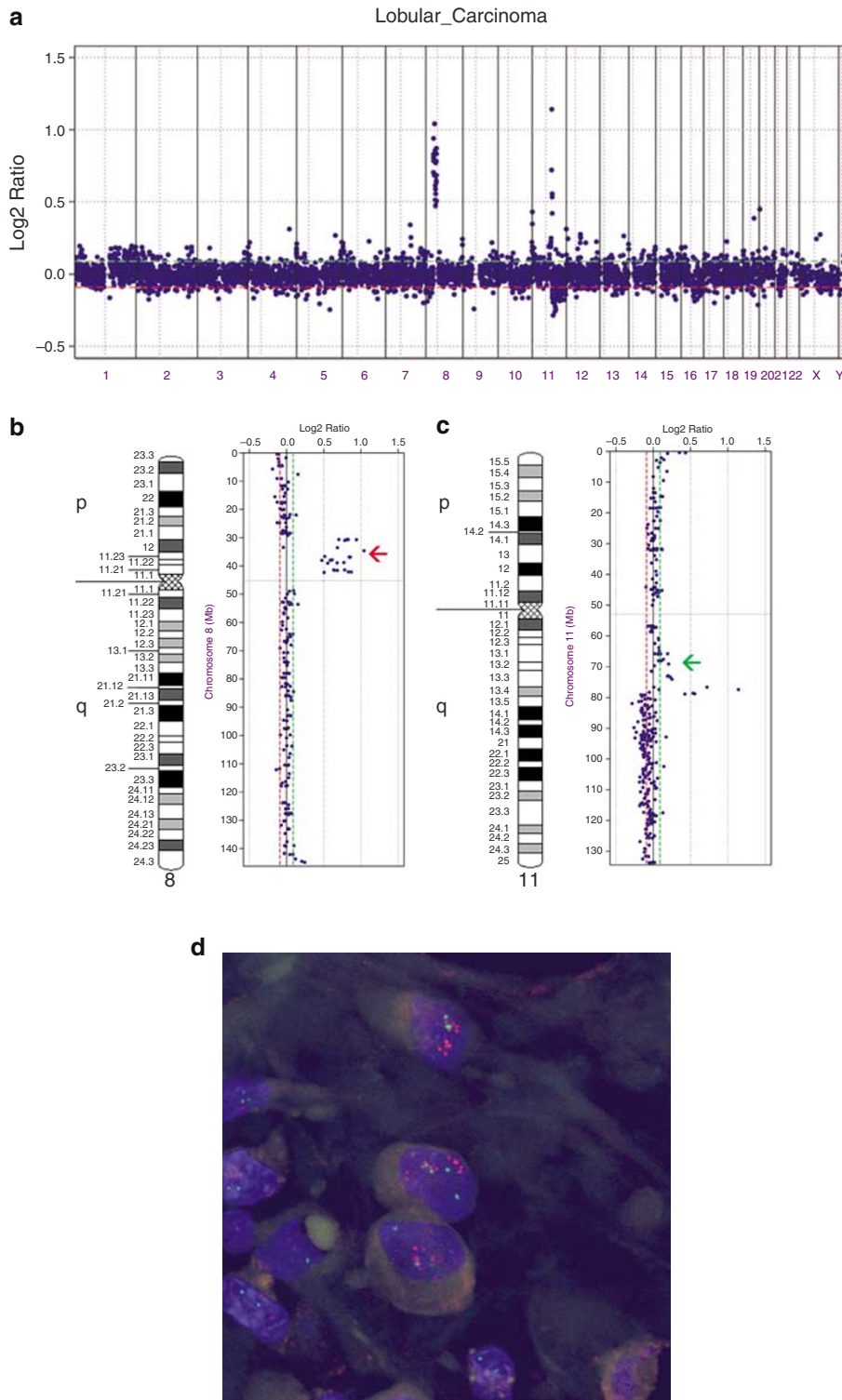




**Figure 1** Comparison between Zymed Spot-Light *CCND1* amplification probe (**a**, **c** and **e**) and a *CCND1* in house probe (**b**, **d** and **f**) in a series of invasive breast carcinomas. (**a** and **b**) No amplification; (**c** and **d**) amplification, large signal clusters; (**e** and **f**) amplification, > 5 individual signals per nucleus.

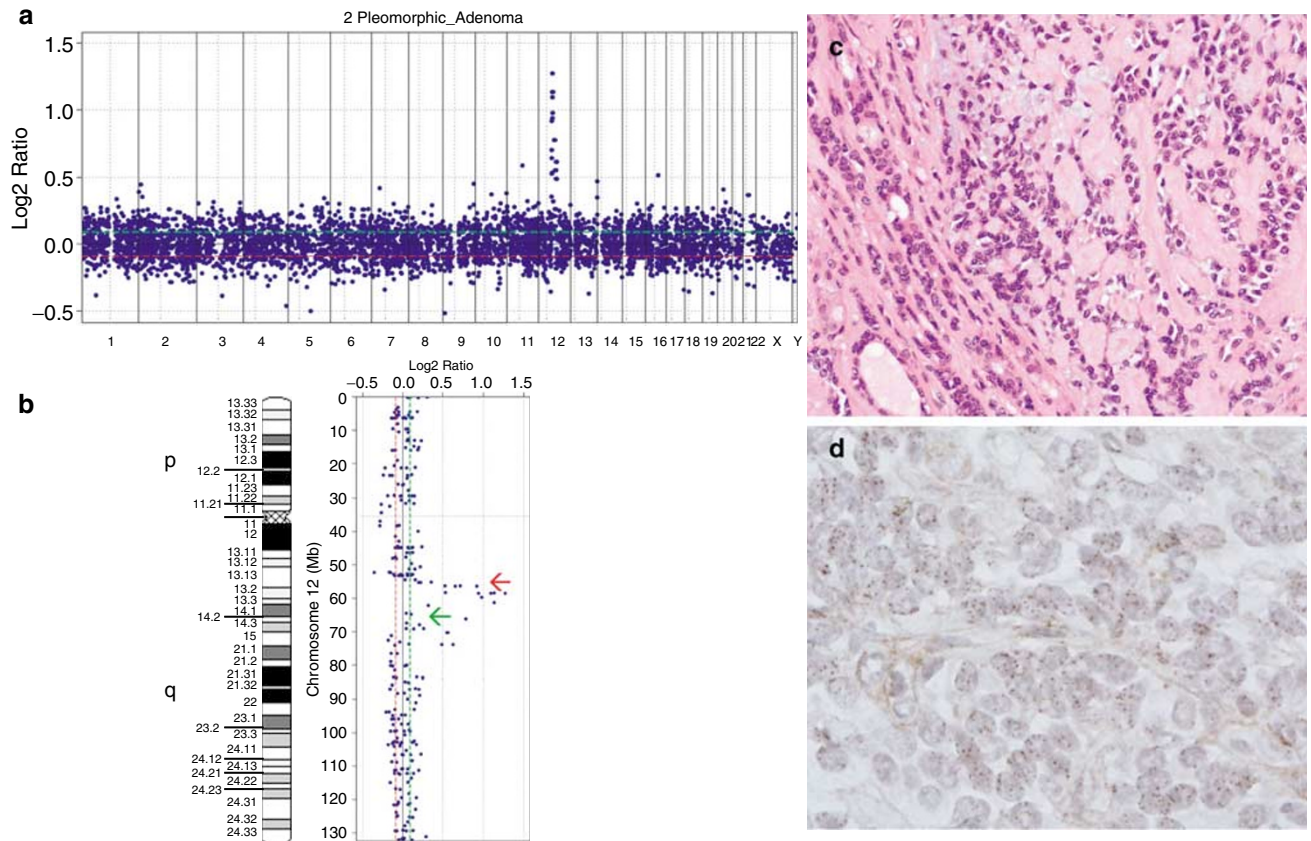
samples of invasive ductal carcinomas of no special type that were used for the assessment of our in house *CCND1* probe. Juxtaposed signals were identified in the majority of neoplastic cells of the

secretory carcinoma (Figure 4), whereas adjacent normal terminal duct-lobular units, endothelial cells, lymphocytes and fibroblasts, as well as all cases of invasive ductal carcinomas of no special



**Figure 2** Invasive lobular carcinoma: **(a)** Genome plot. Individual BAC clones are plotted according to genomic location along the X-axis. Log<sub>2</sub> ratios for each clone are plotted in blue on the Y-axis. Horizontal dashed lines correspond to log<sub>2</sub> ratios of 0.09 (green) and -0.09 (red). Note high-level gains of 8p and 11q, and losses on 11q and 16q. **(b)** Ideogram of chromosome 8 and chromosome plot of microarray CGH results. The red arrow highlights the region target by the probe (*FGFR1* locus). **(c)** Ideogram of chromosome 11 and chromosome plot of microarray CGH results. The red arrow highlights the region target by the probe (*CCND1* locus), which shows a low-level copy number gain. **(b** and **c)** Individual BAC clones are plotted according to genomic location along the Y-axis. Log<sub>2</sub> ratios for each clone are plotted in blue on the X-axis. Horizontal dashed lines correspond to log<sub>2</sub> ratios of 0.09 (green) and -0.09 (red). **(d)** Confocal micrograph demonstrating copy number gains of *FGFR1* (red) and *CCND1* (green).





**Figure 3** Pleomorphic adenoma of the salivary gland: (a) Genome plot. Individual BAC clones are plotted according to genomic location along the X-axis. Log<sub>2</sub> and aws-smoothed ratios for each clone are plotted in grey and blue, respectively, on the Y-axis. Horizontal dashed lines correspond to log<sub>2</sub> ratios of 0.09 (green) and -0.09 (red). Note high-level gains on 12q. (b) Ideogram of chromosome 12 and chromosome plot of microarray CGH results. Individual BAC clones are plotted according to genomic location along the Y-axis. Log<sub>2</sub> and aws-smoothed ratios for each clone are plotted in grey and blue, respectively, on the X-axis. Horizontal dashed lines correspond to log<sub>2</sub> ratios of 0.09 (green) and -0.09 (red). The red arrow highlights *CDK4/SAS* locus, whereas the green arrow highlights the *MDM2* locus. (c) Photomicrograph of a representative area of the pleomorphic adenoma (H&E × 200). (d) Chromogenic *in situ* hybridization with probes targeting *CDK4/SAS*.

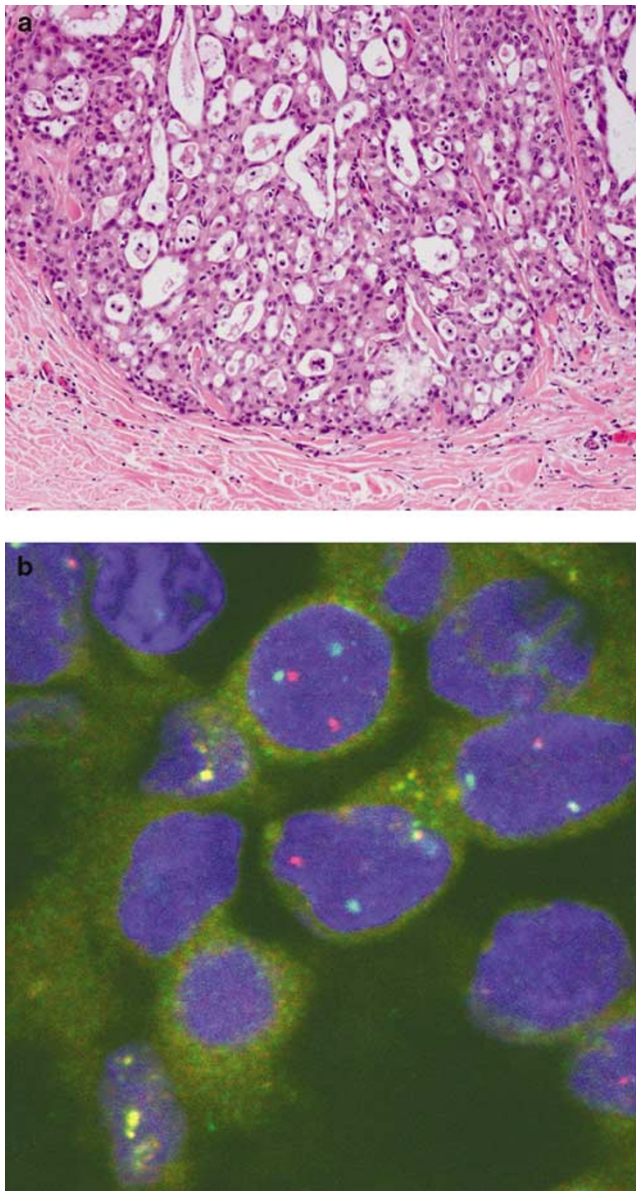
type showed the normal pattern that is no juxtaposed signals.

## Discussion

Fluorescent and CISH techniques have become part of the diagnostic armamentarium of a number of pathology laboratories.<sup>4,6,7,9,13–18,23,29</sup> These techniques are currently being used to identify aneusomy, amplification and balanced chromosomal translocations in human tumours.<sup>4,6,7,9,13–18,23,29</sup> FISH and CISH probes for some key oncogenes, including *EGFR*, *c-MYC*, *HER2*, *MYCN* and *CCND1* are currently commercially available, but by no means does this list cover all important genes amplified and/or rearranged in cancer.

The requirement for simple and reliable validation of genome-wide molecular genetic techniques prompted us to devise a protocol that could be applied to FFPETS. Our protocol generates probes which can be used for both chromogenic and

fluorescence experiments. As the interpretation of CISH is performed on a bright field microscope, this technique allows a direct correlation between copy number changes and morphological features, as well as permitting a quick and reliable assessment of specific chromosomal abnormalities in tissue microarrays. Previous studies have either used direct labelling of BAC probes by nick translation or amplification and labelling by DOP-PCR. In the present study, we employed  $\phi$ 29 amplified BACs as the source of the probes. There are several lines of evidence to suggest that  $\phi$ 29 has a higher fidelity than DOP-PCR and that  $\phi$ 29 amplified products are more representative of the original sequence than DOP-PCR amplified products.<sup>20,22</sup> Other advantages of this method include: (i) direct end sequencing of amplified products and (ii) direct FISH mapping of labelled probes, ensuring the high specificity of the probes for the genomic region of interest. Furthermore, our probes provide signals of similar size and minimal background levels when compared to those obtained with commercially available probes.



**Figure 4** Secretory carcinoma of the breast: (a) H&E; (b) Fluorescent *in situ* hybridization with probes mapping to *ETV6* (green) and *NTRK3* (red). Note the presence of juxtaposed signals confirming the presence of the t(12;15) (p13;q25) chromosomal translocation.

Finally, this protocol can be readily employed to validate aCGH experiments, as probes can be generated from the BACs present in the CGH platform. An alternative approach to generate ISH probes would be to amplify and biotin/DIG label BACs using  $\phi 29$ , without a random primer labelling step. However,  $\phi 29$  (GenomiPhi) does not perform well in incorporating nucleotide analogues, such as biotin or DIG-labelled dNTPs. In fact, we could not obtain biotin-, DIG- or Cy5-labelled products with such approach (data not shown).

In the era of high throughput molecular genetics, one can envisage that the results generated by aCGH should be tested in an independent, large cohort of

tumours, ideally in a high throughput fashion, using for example tissue microarrays. Although FISH can be applied to FFPE tissues, a correlation with morphological features is often very difficult. Moreover, when FISH is applied to tissue microarrays, the correct orientation of cores in a tissue microarray slide under a fluorescent microscope is time-consuming and, sometimes, misleading. For these particular practical purposes, CISH analysis performed on a bright field microscope is more suitable and less labour intensive.<sup>15,19</sup>

In conclusion, this quick protocol generates probes that are highly specific and sensitive. In addition, the amplified products and labelled probes can be end sequenced and FISH mapped, ensuring their specificity. As these probes work reproducibly on FFPE tissue sections and show negligible background, they are a step further to 'unlocking' pathology files for molecular genetic analysis. We envisage that this approach will be instrumental for validation of aCGH studies<sup>9</sup> and high throughput copy number change analysis on tissue microarrays.

## References

- 1 Reis Filho JS, Simpson PT, Gale T, *et al*. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. *Pathol Res Pract* 2005;201: 713–725.
- 2 Pinkel D, Albertson DG. Array comparative genomic hybridization and its applications in cancer. *Nat Genet* 2005;37(Suppl):S11–S17.
- 3 Simpson PT, Reis-Filho JS, Gale T, *et al*. Molecular evolution of breast cancer. *J Pathol* 2005;205:248–254.
- 4 Di Palma S, Lambros MBK, Savage K, *et al*. Oncocytic metaplasia in pleomorphic adenoma: molecular evidence in support of an origin in neoplastic cells. *J Clin Pathol* 2005 in press.
- 5 Hwang ES, DeVries S, Chew KL, *et al*. Patterns of chromosomal alterations in breast ductal carcinoma *in situ*. *Clin Cancer Res* 2004;10:5160–5167.
- 6 Devries S, Nyante S, Korkola J, *et al*. Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. *J Mol Diagn* 2005;7:65–71.
- 7 Shelley Hwang E, Nyante SJ, Yi Chen Y, *et al*. Clonality of lobular carcinoma *in situ* and synchronous invasive lobular carcinoma. *Cancer* 2004;100:2562–2572.
- 8 Wa CV, DeVries S, Chen YY, *et al*. Clinical application of array-based comparative genomic hybridization to define the relationship between multiple synchronous tumors. *Mod Pathol* 2005;18:591–597.
- 9 Reis-Filho JS, Simpson PT, Jones C, *et al*. Pleomorphic lobular carcinoma of the breast: role of comprehensive molecular pathology in characterization of an entity. *J Pathol* 2005;207:1–13.
- 10 Heim S, Teixeira MR, Dietrich CU, *et al*. Cytogenetic polyclonality in tumors of the breast. *Cancer Genet Cytogenet* 1997;95:16–19.
- 11 Pandis N, Teixeira MR, Adeyinka A, *et al*. Cytogenetic comparison of primary tumors and lymph node metastases in breast cancer patients. *Genes Chromosomes Cancer* 1998;22:122–129.

- 12 Going JJ, Abd El-Monem HM, Craft JA. Clonal origins of human breast cancer. *J Pathol* 2001;194:406–412.
- 13 Prentice LM, Shadeo A, Lestou VS, *et al*. NRG1 gene rearrangements in clinical breast cancer: identification of an adjacent novel amplicon associated with poor prognosis. *Oncogene* 2005;24:7285–7289.
- 14 Al-Kuraya K, Schraml P, Torhorst J, *et al*. Prognostic relevance of gene amplifications and coamplifications in breast cancer. *Cancer Res* 2004;64:8534–8540.
- 15 Bhargava R, Gerald WL, Li AR, *et al*. EGFR gene amplification in breast cancer: correlation with epidermal growth factor receptor mRNA and protein expression and HER-2 status and absence of EGFR-activating mutations. *Mod Pathol* 2005;18:1027–1033.
- 16 Bhargava R, Oppenheimer O, Gerald W, *et al*. Identification of MYCN gene amplification in neuroblastoma using chromogenic *in situ* hybridization (CISH): an alternative and practical method. *Diagn Mol Pathol* 2005;14:72–76.
- 17 Arnould L, Denoux Y, MacGrogan G, *et al*. Agreement between chromogenic *in situ* hybridization (CISH) and FISH in the determination of HER2 status in breast cancer. *Br J Cancer* 2003;88:1587–1591.
- 18 Isola J, Tanner M, Forsyth A, *et al*. Interlaboratory comparison of HER-2 oncogene amplification as detected by chromogenic and fluorescence *in situ* hybridization. *Clin Cancer Res* 2004;10:4793–4798.
- 19 Loring P, Cummins R, O'Grady A, *et al*. HER2 positivity in breast carcinoma: a comparison of chromogenic *in situ* hybridization with fluorescence *in situ* hybridization in tissue microarrays, with targeted evaluation of intratumoral heterogeneity by *in situ* hybridization. *Appl Immunohistochem Mol Morphol* 2005;13:194–200.
- 20 Rook MS, Delach SM, Deyneko G, *et al*. Whole genome amplification of DNA from laser capture-microdissected tissue for high-throughput single nucleotide polymorphism and short tandem repeat genotyping. *Am J Pathol* 2004;164:23–33.
- 21 Lage JM, Leamon JH, Pejovic T, *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.
- 22 Sato M, Ohtsuka M, Ohmi Y. Usefulness of repeated GenomiPhi, a phi29 DNA polymerase-based rolling circle amplification kit, for generation of large amounts of plasmid DNA. *Biomol Eng* 2005;22:129–132.
- 23 Evans MF, Aliesky HA, Cooper K. Optimization of biotinyl-tyramide-based *in situ* hybridization for sensitive background-free applications on formalin-fixed, paraffin-embedded tissue specimens. *BMC Clin Pathol* 2003;3:2.
- 24 Yang YH, Dudoit S, Luu P, *et al*. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucl Acids Res* 2002;30:e15–e24.
- 25 Hupe P, Stransky N, Thiery JP, *et al*. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. *Bioinformatics* 2004;20:3413–3422.
- 26 Paris PL, Andaya A, Fridlyand J, *et al*. Whole genome scanning identifies genotypes associated with recurrence and metastasis in prostate tumors. *Hum Mol Genet* 2004;13:1303–1313.
- 27 Reis-Filho JS, Milanezi F, Carvalho S, *et al*. Metaplastic breast carcinomas show EGFR, but not HER2, gene amplification and overexpression: immunohistochemical and chromogenic *in situ* hybridization analysis. *Breast Cancer Res* 2005;7:R1028–R1035.
- 28 Tognon C, Knezevich SR, Huntsman D, *et al*. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer Cell* 2002;2:367–376.
- 29 Makretsov N, He M, Hayes M, *et al*. A fluorescence *in situ* hybridization study of ETV6-NTRK3 fusion gene in secretory breast carcinoma. *Genes Chromosomes Cancer* 2004;40:152–157.