Interrogation of genomes by molecular copy-number counting (MCC)

Angelika Daser[1,2,4], Madan Thangavelu[1,2,4], Richard Pannell[1], Alan Forster[1], Louise Sparrow[3], Grace Chung[1,2], Paul H Dear[1] & Terence H Rabbitts[1]

Human cancers and some congenital traits are characterized by cytogenetic aberrations including translocations, amplifications, duplications or deletions that can involve gain or loss of genetic material. We have developed a simple method to precisely delineate such regions with known or cryptic genomic alterations. Molecular copy-number counting (MCC) uses PCR to interrogate miniscule amounts of genomic DNA and allows progressive delineation of DNA content to within a few hundred base pairs of a genomic alteration. As an example, we have located the junctions of a recurrent nonreciprocal translocation between chromosomes 3 and 5 in human renal cell carcinoma, facilitating cloning of the breakpoint without recourse to genomic libraries. The analysis also revealed additional cryptic chromosomal changes close to the translocation junction. MCC is a fast and flexible method for characterizing a wide range of chromosomal aberrations.

With the completion of the human genome sequencing, methods for analyzing individual genomes are required to implement the objectives of personalized, preventative and predictive medicine. Genomes of cancer cells have many abnormalities (http://cgap.nci.nih.gov/Chromosomes/Mitelman)[1], and hereditary chromosomal abnormalities are associated with complex syndromes and disease predispositions. The proposed sequencing of complete cancer cell genomes (http://www.genome.gov/Pages/Research/SequenceMapsBAC/Workshops/GenomicComparisonsSummary.pdf) might be more effective if focused on regions in which aberrations are located, using methods of scanning genomes for such changes. Additionally, there are many other settings in which copy-number imbalances occur and methods to evaluate these differences are needed, such as for the detection of inherited developmental abnormalities and for analysis of copy-number variations, which are increasingly recognized as being a part of the spectrum of normal human variation.

The most widely recognized group of karyotypically abnormal cells in animals are neoplasms, which contain deletions, amplifications and translocations[1,2]. Whereas leukemias and sarcomas typically have reciprocal chromosomal translocations[2], epithelial tumors (which account for more than 90% of human cancers) are far more heterogeneous[1]. Major events in epithelial tumorigenesis are chromosomal gain or loss and unbalanced translocations. The unbalanced der(3)t(3;5) translocation t(3;5) in renal cell carcinoma associates specifically with nonpapillary renal cell carcinoma[3–5]. The nonreciprocal nature of these translocations has hindered attempts to define the precise der(3)t(3;5) breakpoints.

An approach to this problem could be based on the fact that nonreciprocal translocations result in a change in the copy number of genomic sequence. Several hybridization-based techniques are available to scan the copy number within genomes. These methods use genomic DNA[6] as a probe for arrays of bacterial or P1 artificial chromosomes (BACs or PACs)[7,8] or oligonucleotides[9,10] representing the human reference sequences or use representation genomic probes (ROMA)[11–13] or single nucleotide polymorphism (SNP) analyses[14,15]. Array-based methods, however, lack flexibility because a new array must be created for each set of targets to be examined. Additionally, these approaches are not always quantitative.

We now describe MCC, a new procedure based on PCR analysis of genomes, which does not involve hybridization of genomic DNA or the generation of probes from genomic samples, and does not require extensive optimization for each locus. MCC relies on the direct counting of target sequences in a series of genomic DNA samples at limiting dilution (**Fig. 1**). We have validated this method, by scanning the short arm of human chromosome 3 in renal cell carcinoma for copy-number changes, to locate the breakpoint of a nonreciprocal translocation to within 300 bp, allowing this region to be cloned by inverse PCR. This is the first time the breakpoint of a *de novo* nonreciprocal translocation in renal cell carcinoma has been cloned. The MCC method is fast and highly sensitive, can be applied to small quantities of low-molecular-weight DNA and is amenable to automation.

RESULTS

The MCC method

The essence of the MCC method is described in **Figure 1**. In cases in which the translocation occurs there is 1*n* copy number of the marker for the distal portion of the chromosome and 2*n*

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. [2]Present addresses: Institute of Human Genetics, Johannes Gutenberg University, Langenbeckstr. 1, D-55131 Mainz, Germany (A.D.), MRC Cancer Cell Unit, Hills Road, Cambridge, CB2 2XZ, UK (M.T.) and Department of Anatomical and Cellular Pathology, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong (G.C.). [3]Present address: Prime Health, L1, St Georges Tce, Perth 6000, Australia. [4]These authors contributed equally to this work. Correspondence should be addressed to M.T. (mt370@hutchison-mrc.cam.ac.uk).

RECEIVED 7 MARCH; ACCEPTED 17 APRIL; PUBLISHED ONLINE 23 MAY 2006; CORRECTED AFTER PRINT 8 JUNE 2006; DOI:10.1038/NMETH880

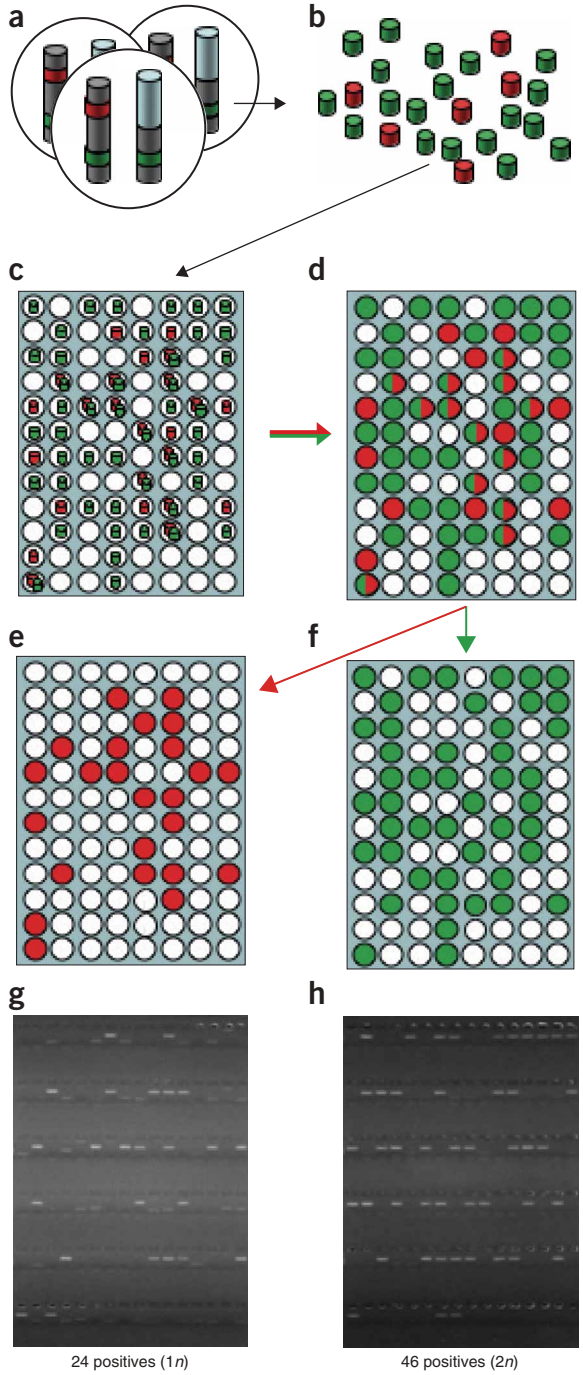

Figure 1 | Overview of the MCC method. (**a**) Cells carry a nonreciprocal translocation, and therefore one marker sequence (green) is present in twice as many copies as a second marker (red), which lies telomeric of the translocation breakpoint. (**b**) DNA is prepared from the target cells and diluted; for simplicity, only the markers are illustrated. (**c**) DNA is dispensed at less than one genome per aliquot into wells of a 96-well plate. (**d**) In an initial multiplex PCR all markers are amplified by a modest amount (using forward and reverse primers for both markers in one multiplex reaction); wells in which 'red' and 'green' markers are amplified are illustrated in this example. (**e**,**f**) The multiplex reaction products are split into replica plates, and a second phase of PCR is carried out with semi-nested primers specific for each marker (using forward-internal and reverse primers specific to each marker in the respective reaction). The distribution of the amplified 'red' (**e**) and 'green' (**f**) marker product in this example is illustrated. (**g**,**h**) Gel electrophoresis analysis of the semi-nested PCR products in MCC applied to kidney carcinoma. In this example, the 'red' marker is found in 24 of the wells (**g**) and the 'green' marker in 46 of the wells (**h**), corresponding to a twofold increase in copy number.

copy number of the marker proximal to the translocation. For MCC, genomic DNA is highly diluted as described for HAPPY mapping[16] (mapping based on the analysis of approximately haploid DNA samples using PCR), and aliquots containing less than one haploid genome of DNA are distributed to 88 wells of a 96-well microtiter plate, leaving 8 wells for negative PCR controls. The first round of PCR analysis is a multiplexed amplification step for each aliquot with all pooled outer primers in each well of the 96-well plate (**Fig. 1d**), so that all copies of any target sequence are amplified to some extent. Previous data have demonstrated that very high levels of multiplexing have been achieved using a similar protocol, indicating that this is a robust approach[17]. The second PCR round is a semi-nested PCR carried out separately for each individual marker, using as the template the PCR product transferred from the multiplex plate into replica plates (**Fig. 1e**,**f**). These second-round PCR products are separated on polyacrylamide gels (**Fig. 1g**,**h**) and scored. The proportion of aliquots that are positive for any marker reflects the relative copy number of that marker in the genome.

Examples of gel visualization of the PCR products in MCC applied to kidney carcinoma are shown in **Figure 1g**,**h** for markers distal or proximal to the nonreciprocal breakpoint. Second-round amplification with distal marker–specific primers yields a PCR product in 24 wells (**Fig. 1g**), whereas amplification with primers for the proximal marker yields a product in 46 wells (**Fig. 1h**), indicating approximately twofold copy-number difference between the markers. When carrying out an initial scan of a chromosomal region, widely spaced markers are used, and the distance between them can be varied according to need. About 70 markers at about 2-Mb spacing would be sufficient to collect information for the entire chromosome 3.

Locating a nonreciprocal translocation in kidney cancer

One motivation behind our development of the MCC method was to locate the breakpoints of the recurrent nonreciprocal chromosomal translocation t(3;5)(p;q) in renal cell carcinoma. Cell lines have been established from tumor material[18] and fluorescence *in situ* hybridization (FISH) analysis of the SK-RC-9 cell line (incompletely tetraploid) revealed the presence of a nonreciprocal t(3;5) chromosomal translocation (**Supplementary Fig. 1** online). We initially investigated the location of the t(3;5) chromosomal translocation breakpoint by FISH using BAC clones from an approximately 3-Mb region (73.2–75.8 Mb) of chromosome 3p (**Fig. 2**). The most proximal BAC clone located 'telomeric' of the t(3;5) breakpoint is RP11-24E1 (at 73256358–73419679 bp from the 3p telomere) (**Fig. 2a**). The BAC clone RP11-781E19 (at 74210885–74236089 bp from the 3p telomere) is located just proximal to the breakpoint as it hybridizes to both wild-type 3 and t(3;5) chromosomes (**Fig. 2b**). These data located the breakpoint in the SK-RC-9 to a region of at most 1 Mb of chromosome 3p13-p12.3 (**Fig. 2b** and **Supplementary Data** online).

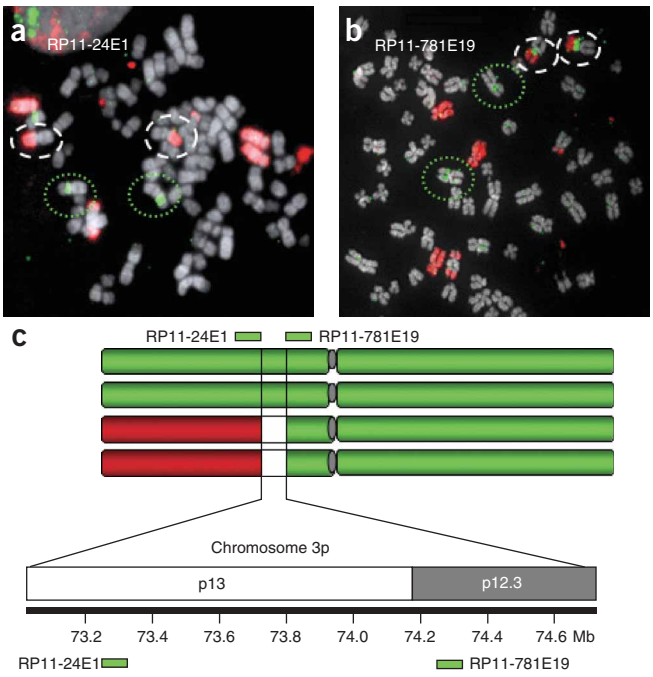

Figure 2 | *In situ* hybridization of BAC clones with SK-RC-9 chromosomes to localize the t(3;5) translocation breakpoint. (**a**,**b**) BAC clones defining an approximately 1-Mb region of human chromosome 3 short arm (clone RP11-24E1 located at 73256358–73419679 bp (**a**) and clone RP11-781E19, located at 74210885–74236089 bp (**b**)) were fluorescently labeled green for FISH analysis of metaphase spreads from SK-RC-9 chromosomes in combination with whole chromosome 5 paints (red fluorescence). BAC clone RP11-24E1 hybridizes to the short arms of two normal chromosomes 3 (**a**; green dotted ovals), but not to the chimeric translocated chromosome t(3;5) (**a**; white dashed ovals) indicating that this BAC is telomeric of the translocation t(3;5) breakpoint. BAC clone RP11-781E19 hybridizes to both normal chromosomes 3 (**b**; green dotted ovals) as well as to the chimeric chromosomes t(3;5) (**b**; white dashed ovals). Thus, this BAC maps centromeric of the t(3;5) breakpoint. These data show that the translocation breakpoint occurs between or within the chromosome region defined by these two BAC clones. (**c**) Schematic representation of normal chromosomes 3 and chimeric t(3;5) in SK-RC-9 showing the region of chromosome 3 that contains the putative t(3;5) translocation breakpoint as determined by FISH. Chromosomes 3 are depicted in green and chromosomes 5 in red. The region containing the breakpoint is represented in white and expanded below, comprising approximately the region of chromosome 3p13–p12.3.

We located the translocation breakpoint on chromosome 3 using multiple rounds of MCC at progressively higher resolution. We performed an initial round of MCC examining the copy number of 12 markers spaced at intervals of 0.2–0.5 Mb over about 3.8 Mb in the region of chromosome 3p13–p12.3 (**Fig. 3**). The results of the round 1 amplification revealed a twofold shift in relative copy number between markers located at 73760583 bp and 74333559 bp, defining the putative translocation breakpoint within a window of ~570 kb. We conducted a subsequent round of MCC using 12 markers at intervals of ~50 kb, further refining the putative breakpoint region to within ~40 kb. MCC rounds 3 and 4 further localized the copy-number shift to ~1–4 kb and to 300 bp, respectively. In addition, the fourth round of MCC revealed an apparent deletion on the 'centromeric' side of the putative translocation (see below).

We confirmed that the MCC data corresponded to genomic alterations using filter hybridizations. We hybridized a probe made from the region of chromosome 3p (**Fig. 4**; probe 3pA) corresponding to the region where the copy-number shift is observed, to filters carrying restriction-digested genomic DNA from the SK-RC-9 cells and, as a control, a different renal cell line SK-RC-12 whose t(3;5) breakpoint lies proximal to that of SK-RC-9 (A.D., G.C. and T.H.R., unpublished data). There are rearranged bands in two different restriction enzyme digests of SK-RC-9 (**Fig. 4a**), showing that the MCC method had identified a genuine abnormality in chromosome 3 in this cell line.

Cloning the t(3;5) translocation breakpoint

Round 4 of the MCC analysis of SK-RC-9 DNA showed a copy-number shift that could be the translocation between chromosome 3 and 5. We designed a pair of chromosome 3 primers for inverse PCR cloning of the DNA corresponding to the region in which we detected the abnormality (**Fig. 5a**). The sequence of the PCR product showed that it comprised the junction of the t(3;5) nonreciprocal chromosomal translocation (**Fig. 5b**) in which a region of chromosome 5q (the location is shown in **Fig. 5c**, coordinate 105386443 bp) had fused with a region of chromosome

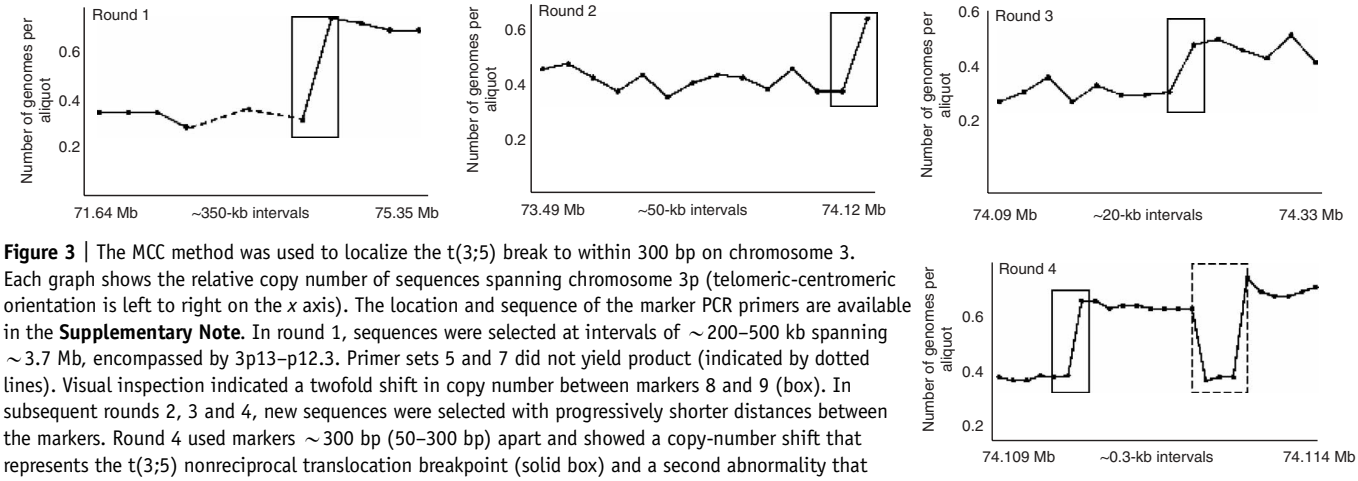

Figure 3 | The MCC method was used to localize the t(3;5) break to within 300 bp on chromosome 3. Each graph shows the relative copy number of sequences spanning chromosome 3p (telomeric-centromeric orientation is left to right on the *x* axis). The location and sequence of the marker PCR primers are available in the **Supplementary Note**. In round 1, sequences were selected at intervals of ~200–500 kb spanning ~3.7 Mb, encompassed by 3p13–p12.3. Primer sets 5 and 7 did not yield product (indicated by dotted lines). Visual inspection indicated a twofold shift in copy number between markers 8 and 9 (box). In subsequent rounds 2, 3 and 4, new sequences were selected with progressively shorter distances between the markers. Round 4 used markers ~300 bp (50–300 bp) apart and showed a copy-number shift that represents the t(3;5) nonreciprocal translocation breakpoint (solid box) and a second abnormality that represents a short deletion of about 700 bp in the region of chromosome 3, (dotted box) centromeric of the translocation junction.

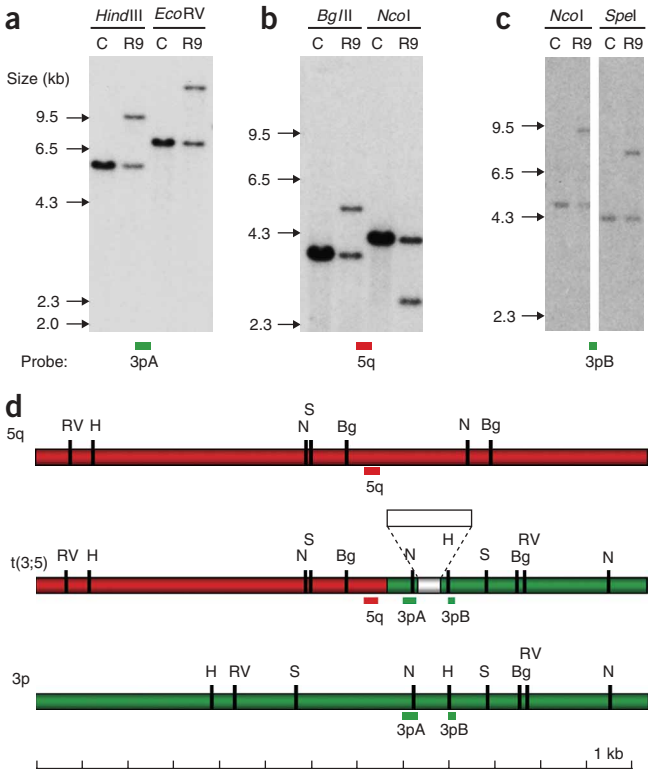

Figure 4 | Filter hybridization of SK-RC-9 DNA shows a rearranged segment and reveals an insertion accompanying a microdeletion. (**a–c**) Genomic DNA from SK-RC-9 or SK-RC-12 (a renal carcinoma with a der(3;5) proximal to that in SK-RC-9) was digested with the indicated restriction enzymes, fractionated and transferred to filters for hybridization to cloned PCR probes from either chromosome 3p (3pA or 3pB) or 5q. The chromosome 3p probes were designed based on the MCC data in **Figure 3**, and a repeat-free genomic sequence from this region was identified using the human genome sequence database. Similarly a probe from chromosome 5q was designed from the human genome sequence database after cloning and sequencing the t(3;5) junction. (**d**) Partial restriction maps of relevant regions of chromosomes 5q, t(3;5) and 3p showing the location of the hybridization probes on either side of the t(3;5) breakpoint (there were two probes for chromosome 3p, designated 3pA and 3pB). Indicated restriction site positions are approximate. The length of the insertion (gray) is not certain and it is possible that it contains additional restriction enzyme sites (as inferred from **a–c**). B, *Bgl*II; H, *Hind*III; N, *Nco*I; RV, *Eco*RV; S, *Spe*I; C, control genomic DNA from SK-RC-12; R9 = SK-RC-9 genomic DNA.

3p (**Fig. 5d**, coordinate 74111893 bp). The adenine at the junction may derive from either chromosome. The rearrangement of this chromosome 5 segment in DNA from SK-RC-9 cells was formally shown using filter hybridization studies (**Fig. 4b**; probe 5q).

MCC can detect cryptic chromosomal changes

While defining the translocation breakpoint by the MCC analysis, we observed an additional copy-number reduction over a region covering ~700 bp just centromeric of the breakpoint (**Fig. 3**). We verified this finding with two independent MCC experiments (**Supplementary Fig. 2** online). A possible explanation for this anomaly in SK-RC-9 DNA could be a small deletion on the t(3;5) chromosome, just centromeric of the translocation breakpoint. We sought to substantiate this by genomic PCR with primers flanking the region. Although we could amplify a fragment of the expected size (907 bp) from normal chromosome 3, there was no evidence of a smaller product that should have been amplified across the deleted segment of DNA (data not shown). The most likely possibility, consistent with this PCR result, is that a short deletion is accompanied by an insertion of sequences from another location. The filter hybridization analysis (**Fig. 4**) confirmed the presence of an insertion centromeric to the breakpoint. For SK-RC-9 DNA digested with *Bgl*II, we expected to detect a rearranged fragment from the der(3;5) chromosome of ~3 kb by hybridization with the 5q probe if the translocation was simply associated with a small deletion. Instead we observed a larger fragment of ~5 kb (**Fig. 4b**). Additionally, the observed sizes of *Nco*I and *Spe*I fragments detected using the 3pB probe (**Fig. 4c**) were both substantially larger than expected for a simple deletion. Taken together, the MCC, PCR and filter hybridization blot data suggest that the nonreciprocal translocation was accompanied by a combination of a short deletion (of ~700 bp) and a large insertion. The origin and exact size of the large insertion are not known. Owing to the complexity of this rearrangement, we have not directly confirmed the short deletion repeatedly observed by MCC by an independent method.

A chromosome 3 deletion in a kidney cancer

A further exemplification of the sensitivity of MCC was provided by the characterization of a cryptic deletion in the cell line SK-RC-12. We performed MCC using a panel of markers spanning chromosome 3 at 76.7–88.0 Mb, at intervals of about 0.35 Mb (**Fig. 6a**). We observed a copy-number shift between markers 22 and 23 (the analytical gels for PCR products of markers 21, 22, 24 and 25 are available in **Supplementary Fig. 3** online). Rounds 2 and 3 of MCC with more closely spaced markers (**Fig. 6a**) resolved the region to about 2 kb, and a final round of MCC localized the copy-number shift to within 400 bp (**Fig. 6a**). We confirmed the presence of a genomic alteration by filter hybridization using a 237-bp probe from chromosome 3 (**Supplementary Fig. 4** online).

We cloned the genomic region corresponding to the round-4 copy-number shift, using inverse PCR, revealing that the copy-number change resulted from a deletion of ~289 kb of chromosome 3p (**Fig. 6b**; 81.64 and 81.94 Mb). The sequences of the regions flanking the deletion point from normal chromosome 3 are compared to the fused chromosome 3 in SK-RC-12 (**Fig. 6b**). This discloses the identity of a 6-bp region on both ends of the deletion segment (**Fig. 6b**) suggesting that this microhomology may have been used in nonhomologous end-joining to repair the double-stranded breaks.

DISCUSSION

MCC has several advantages over other methods for locating alterations in copy number. The method offers effectively unlimited resolution as sequences can be examined from wide intervals down to a few hundred base pairs. Because MCC uses genome sequence information, it only requires a genome database and a series of PCR primers for its implementation. We anticipate that libraries of PCR primers, formatted for use in MCC, can be established to allow the rapid scanning of chromosomal regions or complete genomes. Additionally, the method should be applicable to assaying regions of genomic amplifications, as well as deletions.

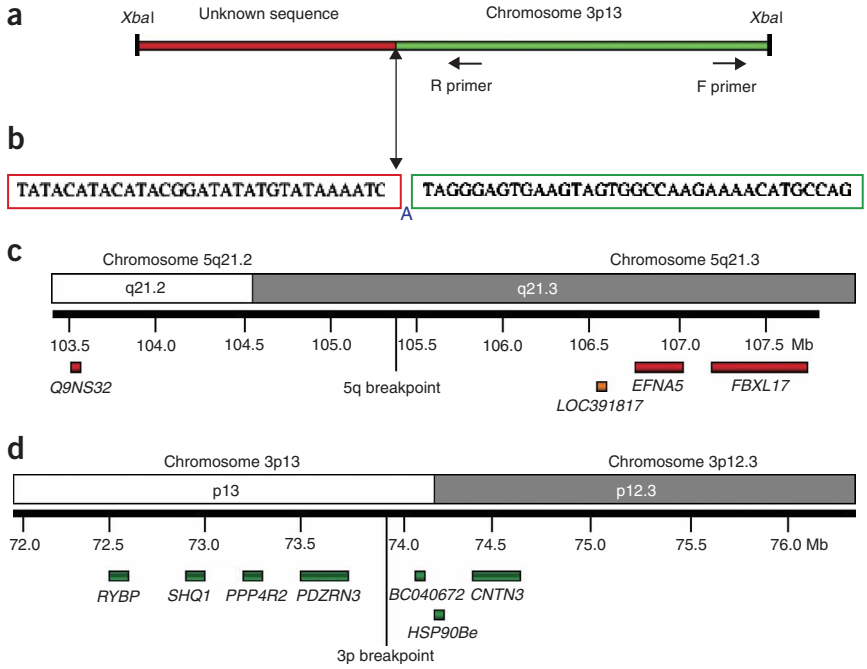

Figure 5 | The sequence and chromosomal location of the t(3;5) nonreciprocal translocation junction. The t(3;5) translocation junction in the SK-RC-9 was located to within 300 bp by MCC. The human genome sequence facilitated identification of restriction sites around this translocation, and inverse PCR was used to clone the junction. SK-RC-9 DNA was digested with *Xba*I, self-ligated at high dilution to obtain intramolecular circles and a PCR product obtained by amplification with the primers F and R (**a**; see **Supplementary Data** for primer sequences). Sequencing of the PCR product confirmed that the junction of the t(3;5) nonreciprocal translocation abutted 5q (**b**; sequence boxed in red) and 3p (**b**; sequence boxed in green) sequences. The location of the translocation breakpoints was determined in chromosome 5q and 3p using the human Ensembl database (NCBI release 35) and indicated in **c** (5q) and **d** (3p). In each of these panels, the top line indicates the relevant chromosome bands and distances (in Mb), and below are shown various known or putative genes. The position of Ensembl genes is shown for both chromosome 5 (**c**) and chromosome 3 (**d**). A Genscan-predicted transcript (LOC391817) is shown at 106.58 Mb on chromosome 5 and two mRNAs located on chromosome 3, BC040672 (RIKEN cDNA) and *HSP90AA1* at 74.10 and 74.2, respectively. The t(3;5) translocation of SK-RC-9 does not split genes on either chromosome 3 or chromosome 5.

Unlike array-based technologies for copy-number determination, MCC does not require whole-genome amplification or any hybridization step. This obviates any problems that might arise from biased amplification, incomplete suppression of repeat sequences within the probe or cross-hybridization, as can occur when using short oligonucleotide arrays or through amplification of *Escherichia coli* DNA contaminating BACs or PACs for array comparative genomic hybridization[8]. Accurate copy-number quantitation by MCC depends upon successful amplification of all of the copies of a locus in the panel of aliquots. In our experience (this work, and unpublished observations) most PCR primer sets either work well or not at all. The latter are obvious and can be discarded from the analysis, as was done for markers 5 and 7 (**Fig. 3**). Nevertheless, a single marker of apparently low copy number must be viewed with caution. MCC is essentially a digital approach that simplifies interpretation of results, whereas the micro-array approaches are quantitative and often require complex algorithms for interpretation[19]. Although MCC is easily applicable to manual operation, it also lends itself well to automation.

MCC requires minuscule amounts of genomic DNA, being applicable to only hundreds of cells, and the DNA does not have to be of high molecular weight. This suggests that MCC will allow hitherto impractical studies, such as the detailed analysis of pre-neoplastic biopsy material from patients, the retrospective analysis of archival tumor samples, or the exploration of genomic variability across different small regions of a tumor. MCC should also simplify the analysis of hereditary chromosomal abnormalities that affect copy number or are a part of the normal spectrum of human variation. The application of MCC described here used cell lines in which the imbalances are constant from cell to cell. The application of MCC to constitutional copy-number differences will be similar because all the DNA will be identical. This will not necessarily be the case for biopsies of disease-based material. Tumor samples may be a particular issue as resections will comprise cancer cells, stromal cells and inflammatory cells. Nevertheless, the small amounts of material required for MCC and the sensitivity of the approach should allow copy-number anomalies to be detected even against some background of normal DNA.

The specific breakpoint that we have cloned represents the first example of cloning a *de novo* nonreciprocal chromosomal translocation. Kidney cancer has a very poor prognosis[20] and tumors arising in the proximal tubule (nonpapillary kidney cancer) often have a nonreciprocal chromosomal translocation t(3;5). The breakpoints on chromosome 3 cluster to three different regions of the short arm[5] and the one in SK-RC-9 cells locates at the most distant cluster (3p13). The breaks do not involve cleavage within any known or putative genes (**Fig. 5**). The use of MCC to analyze breakpoints of other nonreciprocal translocations will shed more light on the mechanism and consequences of these translocations. In this study, the new MCC technology has also been used to determine the presence of cryptic modifications on chromosome 3. As the chromosomes involved in interchromosomal translocations are inherently unstable at the time of double-strand breakage, the incidence of additional changes may not seem too surprising but still can be functionally important.

METHODS

Molecular Copy-number Counting (MCC). We cultured SK-RC-9 and SK-RC-12 cell lines[18] in Dulbecco's modified Eagle's medium with 10% fetal calf serum. We prepared genomic DNA from SK-RC-9 and SK-RC-12 cells using the Qiagen DNeasy Tissue kit. We diluted DNA with distilled water to approximately 10 genomes/µl (about 30 pg/µl), and stored it at –70 °C in aliquots.

For each round of the MCC assay, we used three PCR primers (forward and reverse, plus a nested forward-internal primer) for each locus (that is, genomic marker) to be assayed. In the first phase, we combined multiple PCR primer pairs (forward and

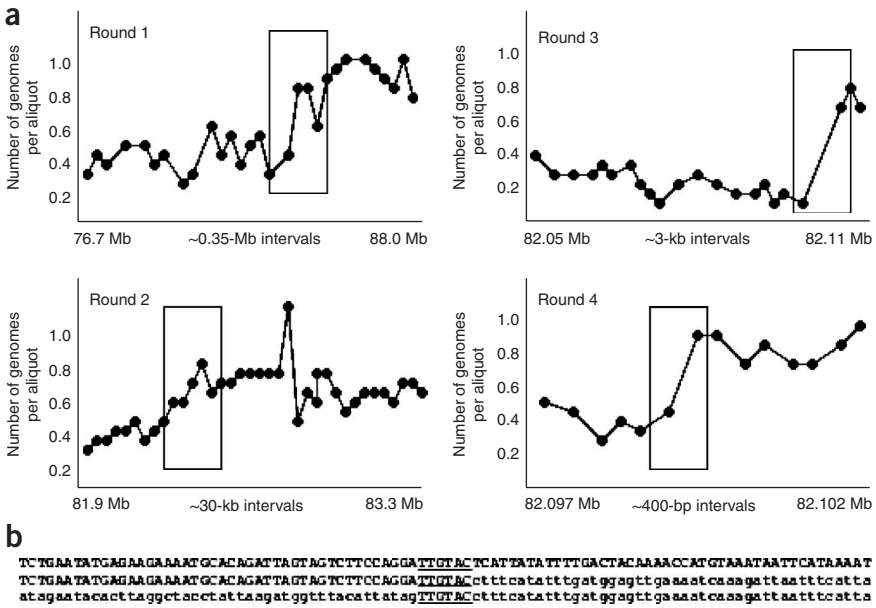

Figure 6 | MCC mapping of a chromosome 3 deletion in the SK-RC-12. (**a**) For MCC round 1 mapping, a panel of 35 markers was used to screen a genomic region of ~9 Mb with marker intervals around 0.25 Mb apart. A copy-number shift was detected between markers 22 and 23. Distances between markers do not reflect genomic distances but are numeric ones. The boxed zones represent shifts from low copy to high copy. Based on round-1 data, a second set of primers were designed spaced at an average of 30 kb apart, and copy-number variation mapped to between markers 9 and 13. Subsequently, two more rounds of MCC were conducted using markers at about 3 kb (round 3) and 400 bp apart (round 4). A copy-number shift occurred between markers 6 and 7 in round 4 defining a location of the shift within 800 bp of chromosome 3. The numbers on the *x* axes delineate the chromosomal regions examined in each round of MCC. (**b**) The genomic region of SK-RC-12 DNA with this copy-number shift was cloned after inverse PCR and revealed a cryptic deletion of chromosome 3 in SK-RC-12 cells. The top sequence (upper-case letters) spans the telomeric end of the chromosome 3p deletion, the bottom sequence (lower-case letters) spans the centromeric end of the chromosome 3p deletion, and the middle sequence is the fusion found at the junction of the deletion in SK-RC-12 DNA, located at 81.64 and 81.94 Mb. Underlined are the six nucleotides of sequence identity on both ends of the deletion segment.

reverse) in a single multiplex reaction. We used the products of this reaction as the templates for the second-phase PCRs, each of which used the forward-internal and reverse primers for each single sequence (see **Supplementary Data**). We selected primers using simple criteria (similar to design for conventional uses of PCR) after masking repetitive elements from the human genomic sequence (Ensembl database, NCBI release 35; http://www.ensembl.org) using RepeatMasker (http://www.repeatmasker.org). Typically, primer length was 18–20 bp, with a melting temperature (T_m) of 52–60 °C (based on the calculation $T_m = 2 \times (A + T) + 4 \times (G + C)$). Design requires at least two guanine or cytosine bases at the 3′ end and at least one at the 5′ end. No runs of any single base longer than 4 bases were allowed. Internal amplimer length was designed to be 80–150 bp, and the position of the external primer no more than 150 bp upstream of forward-internal primer.

DNA concentration is crucial in MCC analysis, but it need not be too precise as accuracy is obtained with average DNA content of ~0.2–0.6 haploid genomes per aliquot. We determined the starting concentration of the genomic DNA preparations using ultraviolet spectrophotometry, and based on these measurements, we made a series of dilutions (conveniently, 6), which we expected to give between 0.25 and 8 genomes of DNA per sample. We assayed 16 aliquots of each dilution (using MCC in 96-well plates, as described above) for each of four markers, corresponding to DNA segments believed to be present at one copy per haploid genome. We used the proportion of aliquots (wells) that scored positive (averaged across all four markers) to calculate the actual DNA concentration in each dilution. In turn, we used these data to determine the exact degree of dilution required for MCC analysis. When the working concentration for MCC has been determined, the DNA dilution may be used for all MCC steps. Each new preparation of DNA requires independent titration.

For MCC, we prepared a master mix containing the forward and reverse PCR primers for all sequences to be assayed (0.15 µM of each oligo), 1× Gold PCR buffer (Perkin-Elmer), 2 mM $MgCl_2$, 200 µM each dNTP and 0.1 u/µl *Taq* Gold DNA polymerase (Perkin-Elmer) and about 0.03 genomes/µl (0.09 pg/µl) of genomic DNA. We dispensed 10 µl of this mix into each of 88 wells of a 96-well plate, and the remaining 8 wells (negative controls) received 10 µl of a similar mix lacking DNA. We overlaid all samples with 20 µl of mineral oil. Thermocycling was carried out with hot start at 93 °C for 9 min, followed by 25 cycles of 20 s at 94 °C, 30 s at 52 °C and 1 min at 72 °C. We diluted each PCR reaction to 500 µl with water, and used 5-µl samples as template in each second-phase (marker-specific) semi-nested PCR (1.5 mM $MgCl_2$, 1 µM of the relevant forward-internal and reverse primers, other concentrations as before and thermocycling at 93 °C for 9 min, followed by 33 cycles of 20 s at 94 °C, 30 s at 52 °C and 1 min at 72 °C. After the semi-nested PCR, we added 8 µl of 2× loading buffer (15% (wt/vol) Ficoll 400, 0.1 mg/ml bromophenol blue, 4× Sybr Green, 1× TBE) to each well, and analyzed amplification products by electrophoresis for 10 min at 10 V/cm in precast 108-well horizontal 6% polyacrylamide gels (MIRAGE gels; Genetix) scoring the presence or absence of PCR product in each sample. In later experiments, we scored the results by melting-curve analysis using an ABI 7900HT with the manufacturer's SDS software.

In the case of melting-curve analysis, we modified the PCR mixture for the second-phase PCRs to contain 4 mM $MgCl_2$ and 0.5× SYBR Green I (Cambrex). All second-phase PCR reactions were set up robotically in multiple 384-well microtiter plates (each plate containing the 96 reactions for each of four markers). Note that the second, semi-nested PCR could be carried out in a second 96-well plate, using a multichannel pipette for transfers, rather than a robotic system.

Statistical analysis of MCC data. If DNA molecules are distributed randomly among a series of aliquots then, from the number of aliquots scoring positive for any given sequence, the concentration

of that sequence (expressed in copies per aliquot) can be determined from the Poisson equation (see **Supplementary Note**). If two or more sequences are analyzed in this way on the same set of aliquots of genomic DNA, then their relative concentrations, and hence their relative abundance in the genomic DNA, can be calculated. For instance, if two DNA markers (A and B) were scored on the same set of 88 aliquots, and if the numbers of aliquots scoring positive for each marker were 34 and 56 respectively, then the average concentrations of the two sequences can be calculated (see **Supplementary Note**) as 0.49 and 1.0 copies per aliquot, respectively. Hence, if sequence A is known to be present at n copies per genome, it may be inferred that sequence B is present at $2n$ copies per genome.

Additional methods. Descriptions of FISH, filter hybridization and inverse PCR cloning are available in **Supplementary Methods** online.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council. G.C. was partly supported by the Kay Kendall Leukemia Fund. We thank L. Old for generously supplying the renal cell line SK-RC-9 and N. Copeland for providing an inverse PCR protocol.

AUTHOR CONTRIBUTIONS

A.D., M.T., R.P., A.F. and G.C. conducted the experimental procedures; A.D., M.T., P.H.D. and T.H.R. devised the project; A.D., P.H.D. and T.H.R. wrote the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Albertson, D.G., Collins, C., McCormick, F. & Gray, J.W. Chromosome aberrations in solid tumors. *Nat. Genet.* **34**, 369–376 (2003).
2. Rabbitts, T.H. Chromosomal translocations in human cancer. *Nature* **372**, 143–149 (1994).
3. Kovacs, G., Szucs, S., De Riese, W. & Baumgartel, H. Specific chromosome aberration in human renal cell carcinoma. *Int. J. Cancer* **40**, 171–178 (1987).
4. Kovacs, G. & Brusa, P. Recurrent genomic rearrangements are not at the fragile sites on chromosomes 3 and 5 in human renal cell carcinomas. *Hum. Genet.* **80**, 99–101 (1988).
5. Kovacs, G. *et al.* The Heidelberg classification of renal cell tumours. *J. Pathol.* **183**, 131–133 (1997).
6. Pinkel, D. & Albertson, D.G. Comparative genomic hybridization. *Annu. Rev. Genomics Hum. Genet.* **6**, 331–354 (2005).
7. Fiegler, H. *et al.* DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosom. Cancer* **36**, 361–374 (2003).
8. Menten, B. *et al.* arrayCGHbase: an analysis platform for comparative genomic hybridization microarrays. *BMC Bioinformatics* **6**, 124 (2005).
9. Brennan, C. *et al.* High-resolution global profiling of genomic alterations with long oligonucleotide microarray. *Cancer Res.* **64**, 4744–4748 (2004).
10. Carvalho, B., Ouwerkerk, E., Meijer, G.A. & Ylstra, B. High resolution microarray comparative genomic hybridisation analysis using spotted oligonucleotides. *J. Clin. Pathol.* **57**, 644–646 (2004).
11. Lucito, R. *et al.* Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res.* **13**, 2291–2305 (2003).
12. Sebat, J. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528 (2004).
13. Jobanputra, V. *et al.* Application of ROMA (representational oligonucleotide microarray analysis) to patients with cytogenetic rearrangements. *Genet. Med.* **7**, 111–118 (2005).
14. Zhou, W. *et al.* Counting alleles reveals a connection between chromosome 18q loss and vascular invasion. *Nat. Biotechnol.* **19**, 78–81 (2001).
15. Zhao, X. *et al.* Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. *Cancer Res.* **65**, 5561–5570 (2005).
16. Dear, P.H. & Cook, P.R. Happy mapping: a proposal for linkage mapping the human genome. *Nucleic Acids Res.* **17**, 6795–6807 (1989).
17. Eichinger, L. *et al.* The genome of the social amoeba Dictyostelium discoideum. *Nature* **435**, 43–57 (2005).
18. Ebert, T., Bander, N.H., Finstad, C.L., Ramsawak, R.D. & Old, L.J. Establishment and characterisation of human renal cancer and normal kidney cell lines. *Cancer Res.* **50**, 5531–5536 (1990).
19. Daruwala, R.S. *et al.* A versatile statistical analysis algorithm to detect genome copy number variation. *Proc. Natl. Acad. Sci. USA* **101**, 16292–16297 (2004).
20. Cohen, H.T. & McGovern, F.J. Renal-cell carcinoma. *N. Engl. J. Med.* **353**, 2477–2490 (2005).

Corrigendum: Interrogation of genomes by molecular copy-number counting (MCC)

Angelika Daser, Madan Thangavelu, Richard Pannell, Alan Forster, Louise Sparrow, Grace Chung, Paul H Dear & Terence H Rabbitts
***Nat. Methods* 3, 447–453 (2006); published online 23 May 2006; corrected after print 8 June 2006.**

Correspondence should be addressed to M. Thangavelu (mt370@hutchison-mrc.cam.ac.uk) instead of T. H. Rabbitts. *The error has been corrected in the PDF version of the article.*