

Chapter 2

Molecular Genetics Methods in Discovery of Chromosome Structure

Donna G. Albertson

Contents

| | | |
|---------|--|----|
| 2.1 | Introduction..... | 16 |
| 2.1.1 | Cancer Genome Alterations | 16 |
| 2.1.2 | Cytogenetic and Molecular Techniques | 18 |
| 2.1.2.1 | Fluorescent <i>In Situ</i> Hybridization (FISH)..... | 18 |
| 2.1.2.2 | Comparative Genomic Hybridization | 19 |
| 2.1.2.3 | Amplification-Based Methods for Genome Copy Number Measurement..... | 20 |
| 2.1.2.4 | Multiplex Ligation-Dependent Probe Amplification (MLPA)..... | 20 |
| 2.1.2.5 | nCounter®..... | 21 |
| 2.1.2.6 | Whole Genome Sequencing..... | 21 |
| 2.1.3 | Combining Technologies to Better Study Tumours | 22 |
| | References..... | 24 |

Abstract Somatically acquired tumour genome alterations underlie many of the changes in gene expression that promote tumour formation. These changes, ranging from single nucleotide changes to those involving parts of chromosomes or whole chromosomes, likely reflect the many different solutions taken by individual tumours to escape normal growth regulatory mechanisms. A variety of molecular and cytogenetic techniques, differing in resolution and capabilities for high throughput or single cell analysis, for example, have been used to investigate the altered state of tumour genomes. Some of these methods have become the mainstay of clinical cancer diagnosis and patient management.

Keywords Copy number • Chromosome aberrations • FISH • CGH • Whole genome sequencing

D.G. Albertson (✉)

Department of Oral and Maxillofacial Surgery, Bluestone Center for Clinical Research,
New York University College of Dentistry, New York, NY 10010, USA
e-mail: albertson@cc.ucsf.edu

2.1 Introduction

Human cancer genomes typically have somatically acquired genome alterations, spanning the range from single nucleotide changes to those involving parts of chromosomes or whole chromosomes. These aberrations underlie many of the changes in gene expression that promote tumour formation, such as increases or decreases in copy number of coding regions, inactivation or activation of genes by point mutations or gene disruption, and activation of genes by mutation and rearrangements that create fusion genes with new properties and gene expression patterns. A variety of molecular and cytogenetic techniques have been used to investigate the altered state of tumour genomes, some of which have become the mainstay of clinical cancer diagnosis and patient management.

2.1.1 *Cancer Genome Alterations*

Great variation in the numbers and types of chromosome level alterations present in human tumours has been reported, which is likely to reflect the many different solutions taken by individual tumours to escape normal growth regulatory mechanisms. Figure 2.1 illustrates chromosome level alterations that may occur in a tumour genome.

Some changes result in loss of heterozygosity (LOH), i.e., change in the normal equal contribution to the diploid genome from both the maternal and paternal chromosomes. Such alterations do not affect the number of copies of regions of the genome and are often referred to as “copy neutral” changes; however, gene expression may be altered by LOH and contribute to tumour formation. For example, at a locus heterozygous for a mutation in a tumour suppressor gene, somatic recombination could result in loss of the wild type allele and its replacement with the mutant copy, such that the cell would be homozygous for the mutation. Intra-chromosomal inversions, another class of copy neutral alterations might also alter gene expression as a result of a change in the gene’s neighbourhood or fusion with part of another gene, for example.

By contrast, chromosome level alterations often result in net gain or loss of whole chromosomes (aneuploidy) or parts of chromosomes (insertions, deletions, non-reciprocal translocations). Gene amplification, defined as a copy number increase of a restricted region of a chromosome arm, may also occur [1]. The analysis of amplified DNA in mammalian cell lines and tumours has revealed that it may be organized as extra chromosomal copies, called double minutes, in tandem arrays as head to tail or inverted repeats within a chromosome, often forming a cytologically visible homogeneously staining region (HSR) or distributed at multiple locations in the genome [1]. The unit of amplified DNA in some cases may involve sequences from two or more regions of the genome, indicating a more complex process of formation involving multiple chromosomes [2]. Regions of focal copy

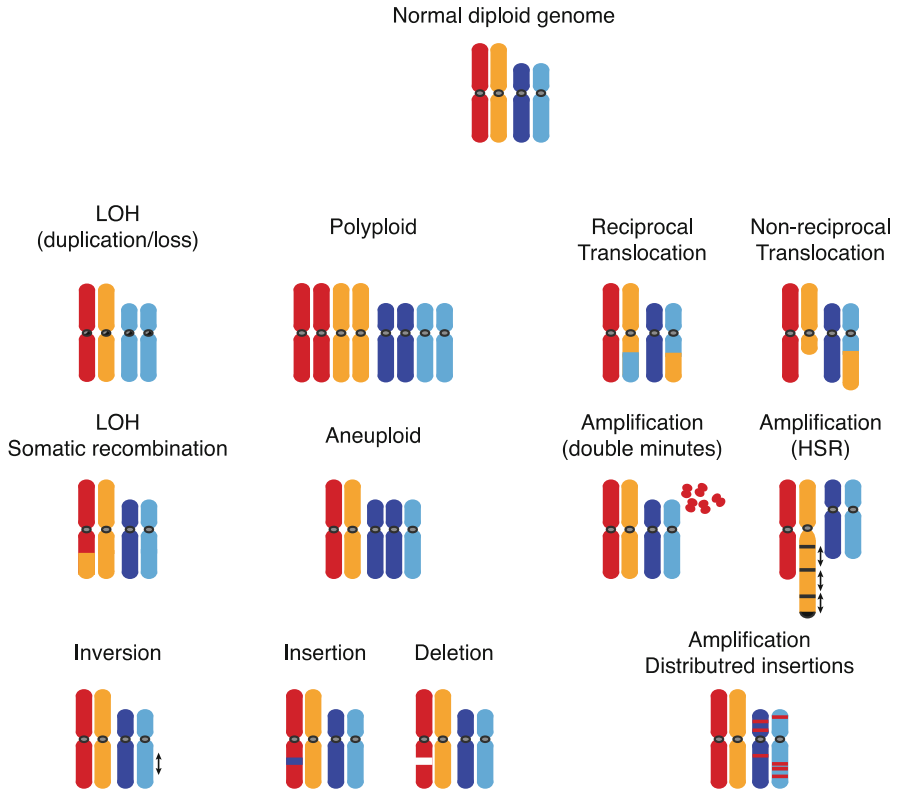


Fig. 2.1 Schematic illustration of mechanisms by which chromosomal aberrations arise. A two chromosome diploid genome with a large (*red and orange*) and smaller chromosome (*blue and light blue*) pair is depicted at the top of the figure. The maternal and paternal chromosomes of the pairs are distinguished by shades of *red/orange* and *blue/light blue*

number alterations, such as small deletions and amplicons focus attention on genes in these regions as potential tumour suppressors and oncogenes, respectively. Regions of amplification pinpoint genes whose elevated expression is likely to be beneficial to the tumour. As amplicons are unstable [3], it is likely that there is ongoing selection for their retention and the elevated expression of gene(s) in the region.

Cytogenetic and molecular methods have been applied to study the organization of frequently amplified oncogenes such as *MYCN*, *EGFR* and *ERBB2* [4–6]. Recently, application of allele specific copy number and high throughput sequencing technologies (see below) has provided fine resolution maps of amplicons and chromosome alterations in tumours. The term “chromothripsis” (derived from Greek, *chromo* for chromosome and *thripsis*, shattering to pieces) was used to describe the complex rearrangements involving multiple breakpoints and copy number alterations seen in 2–3 % of cancers [7]. It was imagined that the rearrangements occurred in a single cataclysmic event, which involved shattering of a

chromosome and its reassembly, rather than stepwise accumulation over time, the existing view of genome evolution in cancer. Although the term gained popularity and has even been considered a mechanism, its usage was subsequently appropriately criticized based on mathematical modelling and existing knowledge of cancer genomes and nuclear organization [8, 9]. Indeed, the complex amplicons that might be formed via a breakage-fusion-bridge process [10], in which repeated cycles of fusion of broken ends of chromosomes lead to failure to segregate properly at mitosis with subsequent breakage during anaphase, were considered examples of chromothripsis. Such a process, however, requires multiple rounds of cell division, generates unstable chromosomes [3], and results in heterogeneity in a population of cells, consistent with the observed copy number profiles characteristic of chromothripsis [9]. In summary, therefore, there appears to be little need for this terminology, the observed complex rearrangements being adequately explained by known genomic instability mechanisms.

2.1.2 Cytogenetic and Molecular Techniques

2.1.2.1 Fluorescent *In Situ* Hybridization (FISH)

A variety of cytogenetic applications use FISH to detect changes in the copy number of loci, a change in the organization of the loci on a chromosome (e.g., inversions, deletions duplications, amplifications) and between chromosomes (e.g., translocations, amplifications). The method uses one or more nucleic acid probes labelled with a fluorochrome conjugated nucleotide or other hapten, such as biotin, that can be detected by fluorescently labelled molecules such as avidin or hapten specific antibodies. The labelled probe(s) are hybridized to whole organisms, tissue sections, cells or subcellular constituents such as metaphase chromosomes, nuclei or extended chromatin fibres and the site of the nucleic acid sequence visualized by fluorescence microscopy [11, 12]. Single locus FISH probes are currently in routine use in clinical laboratories, for example, to assess amplification of *ERBB2* in tumours as a guide to therapeutic decisions [13] and to detect aneuploid cells in urine as a non-invasive alternative to cystoscopy to monitor bladder cancer patients for disease recurrence [14]. Recurrent translocations characteristic of certain cancers are also used to identify cancers. Probes that flank the breakpoints are labelled in different colours and following hybridization, the presence of the translocation is readily observed by separation of the two normally overlapping coloured signals in the cancer cell nuclei [15, 16].

Analysis of tumour karyotypes using FISH to differentially label whole chromosomes or parts of chromosomes (painting probes) can provide higher resolution information on chromosome rearrangements than is possible by standard G-bands by Trypsin using Giemsa (GTG) metaphase chromosome analysis, especially for the common situation in which it is not possible to prepare well banded metaphases from the tumour. Two of the first described painting probe approaches, Spectral

karyotyping (SKY) and M-FISH developed for this purpose used chromosome specific probe libraries differentially labelled with four to seven different fluorophores [17, 18]. Following hybridization and imaging, sequences from the 24 human chromosomes can be distinguished based on the spectroscopic properties of the probes and localized on metaphase spreads prepared from the tumour. Aneuploidies and the composition of abnormal marker chromosomes can be revealed using whole chromosome paints, but within chromosome structural aberrations, such as inversions, deletions, insertions, and duplications cannot be detected. Variations on labelling with whole chromosome paints include, for example, addition of region specific probes obtained by chromosome microdissection or locus specific probes [19]. These alternatives can provide higher resolution information on specific genome regions or types of aberrations.

2.1.2.2 Comparative Genomic Hybridization

Described in 1992, comparative genomic hybridization (CGH) provided the first efficient approach to scan the entire genome for variations in DNA copy number [20]. In the original implementation of CGH, total genomic DNAs isolated from test and reference cell populations were differentially labelled and hybridized to metaphase chromosomes. The binding of sequences at different genomic locations was measured relative to the physical position on the chromosomes. Subsequently, as the human genome mapping and sequencing projects progressed, chromosomes were replaced by DNA microarrays containing elements, initially bacterial artificial chromosome (BAC) [21] or cDNA [22] clones spanning the genome, which had been mapped directly to the physical map of the genome or genome sequence. With either representation of the genome, copy number is determined from the relative hybridization intensity of the test and reference signals at a given genomic location and is proportional to the relative copy number of those sequences in the test and reference genomes. Typically, the reference sample has a normal genome, so that increases and decreases in ratio directly indicate DNA copy number variation in the genome of the test sample. Data are typically normalized so that the modal ratio for the genome is set to some standard value, typically 1.0 on a linear scale or 0.0 on a logarithmic scale. With the completion of the human genome sequence, arrays comprised of short oligonucleotides containing single nucleotide polymorphisms (SNP arrays) became commercially available that allowed information on allele specific copy number to be obtained [23, 24], thereby providing data on both copy number and LOH. At present, whole genome sequencing is replacing microarray-based methods for measuring copy number (see below).

An alternative comparative genomic hybridization platform uses molecular inversion probes (MIPs). Available as the Oncoscan™ FFPE Assay from Affymetrix, Inc., the technology can be used to detect selected cancer relevant single nucleotide mutations and measures copy number and LOH with 300 kb resolution from small amounts of DNA extracted from frozen or FFPE material [25]. The technology uses padlock probes [26]. The probes are designed such that the two ends of the probes

hybridize to ~40 bp regions in the genome leaving a single nucleotide gap. The gap is filled (allowing SNP detection) and the ends of the probes ligated to generate circular probes. Exonuclease digestion is used to remove other nucleic acid sequences and the probes are hybridized to an array via a specific tag sequence included in each probe. Copy number is then determined relative to a normal reference, ideally a patient matched normal sample. Advantages of the technology include requirement for small amounts of sample DNA, compatibility with degraded DNA extracted from FFPE samples, and simultaneous copy number, LOH and SNP detection. A version of the technology is the first chromosomal microarray to receive FDA approval for postnatal testing for germline chromosomal copy number alterations associated with developmental delay, intellectual disability, congenital anomalies, and/or dysmorphic features.

2.1.2.3 Amplification-Based Methods for Genome Copy Number Measurement

Real time quantitative polymerase chain reaction (PCR) has been used to measure copy number at specific loci in the genome relative to a reference locus. An advantage of this approach is the rapid turnaround time and possibility of automating the assay for hundreds of samples. The choice of reference locus for studies of cancer genomes, however, can be problematic. The copy number of the reference locus may not be known, it may not be single copy in tumours and the copy number may vary amongst tumours in a cohort under study. To address this problem, a multicopy reference has been introduced (Qiagen, Inc.). A reference sequence present in >20 copies per diploid genome and distributed across the genome is relatively insensitive to changes in copy number that affect a single locus or a few of the loci, and gain or loss of one or a few copies will result in only a small change in the measured C_T value for the reference.

2.1.2.4 Multiplex Ligation-Dependent Probe Amplification (MLPA)

The MLPA method measures copy number in a multiplex polymerase chain reaction [27]. Locus specific probes recognizing adjacent regions in the genome are annealed, ligated and amplified using universal primer sequences. Probe sets are designed such that ~50 amplification products can be distinguished and quantified following separation by capillary electrophoresis. Comparison of a test sample to a reference sample provides information on copy number. Probes can also be designed to interrogate SNPs and methylation status [28]. While MLPA offers advantages in terms of cost, turnaround time and capability to use degraded FFPE DNA, it is limited to the simultaneous analysis of ~50 loci. Performance of MLPA is also sensitive to the choice of reference DNA, ideally it should be normal DNA from the same individual extracted in the same manner as the test sample [28].

2.1.2.5 nCounter®

The nCounter® system from NanoString® captures and directly counts individual molecules without the need for amplification and can be used to determine copy number at defined loci in the genome [29]. The system uses a 35–50 base pair capture probe complementary to a nucleic acid sequence of interest and a second 35–50 base pair reporter probe complementary to a second region of the nucleic acid sequence of interest. The reporter probe carries a coloured barcode consisting of a DNA sequence annealed with complementary *in vitro* transcribed RNA sequences each labelled with a fluorophore. Multiplex hybridization of the region specific probes takes place in solution. Following hybridization, excess probe is washed away and the hybridized complexes are oriented and extended on a capture surface by application of an electric field. The linear order of the fluorophores in the barcodes of single molecules are then imaged and counted to determine the copy number of each locus. The nCounter® system allows simultaneous interrogation of several hundred loci and is suitable for use with DNA obtained from fresh frozen or FFPE samples.

2.1.2.6 Whole Genome Sequencing

Next generation sequencing (NGS) or high throughput whole genome sequencing technology offers the opportunity to sequence millions of reads in a cost effective manner. Four general methods are used to identify copy number alterations using NGS, including assembly-based methods, depth of coverage or read depth methods, paired-end or read-pair and split-read methods [30, 31]. Assembly-based methods, which reconstruct a genome *de novo* are best suited to studies of small genomes and have not been widely applied in human genome studies. The other three methods rely on aligning sequence reads to a previously established reference genome for the organism.

Depth of coverage methods (DOC) use short single or paired end reads and determine copy number based on number of reads that fall within a bin of defined size, e.g., 15 kb. There is an underlying assumption of uniform sequence coverage of the genome; however, the variation of counts amongst bins is affected by the DNA copy number variations, the Poisson statistics of counting reads, and by biases of the analytical process that have substantial dependences on such factors as the GC content and mappability of sequences in the bins. Coverage is reduced in regions of the genome with high or low GC content and in repetitive regions in which reads cannot be mapped unambiguously. Algorithms to correct for these biases have been developed. Alternatively, comparisons to sequencing data from appropriate reference genomes have been used to normalize data from test samples. Algorithms incorporating information on SNP heterozygosity have also been used to call both copy number and loss of heterozygosity. The capability of DOC methods to use short single end reads offers an advantage when working with archival FFPE tumour specimens from which DNA is likely to be fragmented. A further benefit for tumour

genome analysis is the use of reference free DOC methods, since matched normal reference DNA may not be available [32].

Paired end and split read methods require paired sequencing reads. Deletions and insertions are detected when the paired reads align to the reference genome at distances greater than or less than expected, respectively based on the length of the fragments being sequenced. Paired-end sequencing can also detect inversions and translocations depending on the manner in which the paired ends map to the reference genome. Breakpoints in the genome can be quite accurately mapped by analysis of split-reads in the case that one of the paired end reads maps to the sequence and the other read, which fails to align is considered to span the breakpoint of a genome rearrangement.

A number of algorithms have been described for detecting copy number alterations from NGS data. They address the general workflow of first inferring copy number profiles from the raw sequence, segmenting the profiles and calling aberrations. A comparison of algorithms revealed differences in sensitivity and specificity for different sizes and types of genome alterations [33]. Further refinements in algorithms for detecting tumour genome copy number and structure are expected to better address the technical biases inherent in the current sequencing methodology, as well as incorporating improved knowledge of human genome variation to identify germline copy number variants that could be misinterpreted as tumour genome alterations [34].

2.1.3 Combining Technologies to Better Study Tumours

Cytogenetic and molecular methods for detecting and measuring tumour genome alterations vary in resolution, utility for detecting previously unknown aberrations and sensitivity to admixed normal cells or tumour heterogeneity. The combined use of more than one technique can provide greater insight into alterations in the genomes of the tumour cells. An example is shown in Fig. 2.2, in which array CGH and FISH were used to study an oral squamous cell carcinoma (SCC) primary and recurrence. Analysis of the primary and recurrence by array CGH revealed low level gains and losses in the genomes, as well as amplification of *CCND1* on chromosome 11. By contrast, amplification of *EGFR* on chromosome 7 was observed only in the primary (Fig. 2.2a). Using FISH probes for *EGFR* and *CCND1*, amplification of both regions was apparent in the primary, but *EGFR* was only modestly elevated in the recurrence (Fig. 2.2b), consistent with the array CGH copy number analysis. Enumeration of FISH counts at five different regions in the primary, however, revealed that the tumour was heterogeneous with respect to amplification of *EGFR* with one of five regions having only modestly elevated copy number of *EGFR* similar to the recurrence (Fig. 2.2c). These observations suggest that the recurrence possibly originated from (residual) cells from this region lacking *EGFR* amplification.

In the above example, specific chromosome alterations were selected for analysis based on the genome-wide copy number information provided by array CGH. By

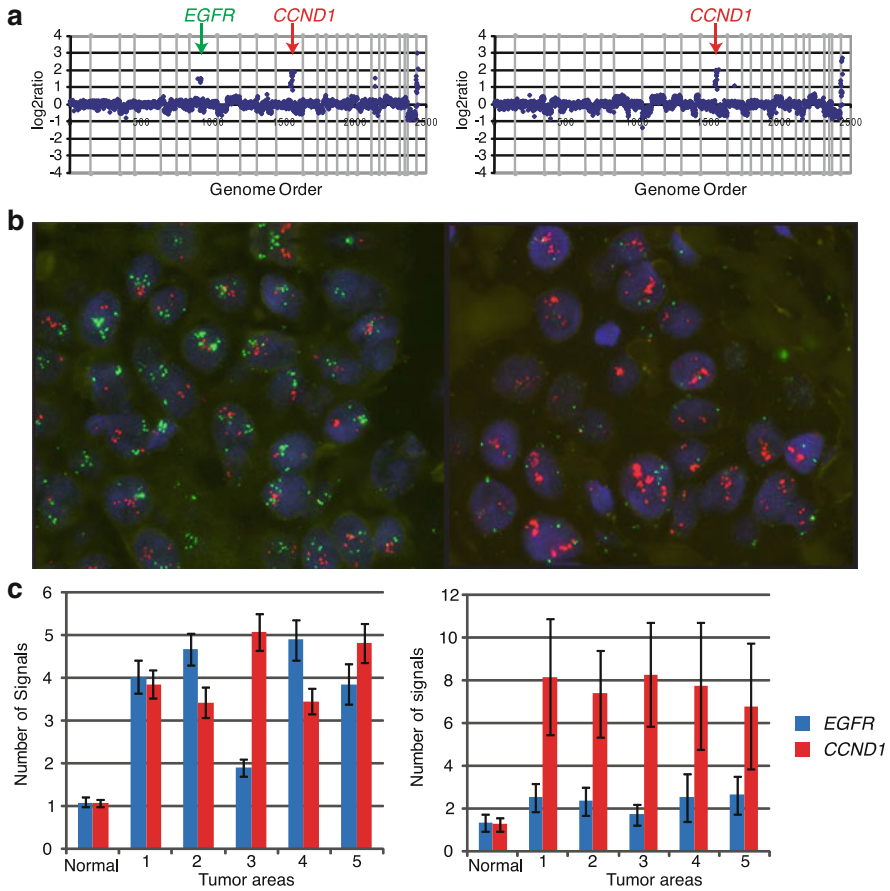


Fig. 2.2 Combining FISH and array CGH reveals tumour heterogeneity. (a) Copy number profiles of an oral SCC primary (left) and recurrence (right). The normalized \log_2 ratio is plotted at each locus sorted by chromosome and ordered according to genome position from the p-arm to the q-arm. Amplifications of *EGFR* on chromosome 7 and *CCND1* on chromosome 11 are present in the primary, but only amplification of *CCND1* in the recurrence. (b) Hybridization of FISH probes for *EGFR* (green) and *CCND1* (red) to tissue sections from the primary (left) and recurrence (right). The large clusters of green signals indicative of *EGFR* amplification are absent from the recurrence consistent with the array CGH profiles. (c) Enumeration of FISH signals from five regions of the sections from the primary (left) and (recurrence) right. While four of the five analyzed regions showed elevated counts for both *EGFR* and *CCND1* in the primary (4–5 times the number of counts for nuclei from non-tumour tissue), one region (region 3) was found in which *EGFR* copy number was only twice normal levels. Note, that due to truncation of nuclei by sectioning, fewer than a diploid number of FISH signals are routinely observed in normal tissue. In the recurrence, amplification of *CCND1* is observed in all regions with only modestly increased copy number of *EGFR* compared to normal tissue

contrast, a similar study using next generation sequencing and microdissection of tumour regions could provide genome-wide information on tumour heterogeneity, albeit with much greater computational effort. Nevertheless, there still appears to be an important role for cytogenetic techniques in the analysis of tumours. While the focus here has been on methods to study tumour genomes, techniques such as FISH are compatible with simultaneous analysis of expressed proteins by immunofluorescence [35] and spatial information on intra-tumour genome alterations and cellular phenotypes can be informative with respect to tumour evolution, for example [31].

The variety of cytogenetic and molecular technologies available for measurement of tumour genome alterations provides researchers and clinicians with many choices. Assessment of the differing capabilities, advantages and weaknesses of the technologies should allow selection of the platform best suited to particular applications, including considerations of cost, throughput, sensitivity and resolution.

References

1. Albertson DG, Collins C, McCormick F, Gray JW (2003) Chromosome aberrations in solid tumours. *Nat Genet* 34:369–376
2. Guan XY, Xu J, Anzick SL, Zhang H, Trent JM et al (1996) Hybrid selection of transcribed sequences from microdissected DNA: isolation of genes within amplified region at 20q11-q13.2 in breast cancer. *Cancer Res* 56:3446–3450
3. Snijders AM, Hermesen MA, Baughman J, Buffart TE, Huey B et al (2008) Acquired genomic aberrations associated with methotrexate resistance vary with background genomic instability. *Gene Chromosome Cancer* 47:71–83
4. Savelyeva L, Sagulenko E, Schmitt JG, Schwab M (2006) The neurobeachin gene spans the common fragile site FRA13A. *Hum Genet* 118:551–558
5. Vogt N, Lefevre SH, Apiou F, Dutrillaux AM, Cor A et al (2004) Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. *Proc Natl Acad Sci U S A* 101:11368–11373
6. Windle BE, Wahl GM (1992) Molecular dissection of mammalian gene amplification: new mechanistic insights revealed by analyses of very early events. *Mutat Res* 276:199–224
7. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR et al (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144:27–40
8. Righolt C, Mai S (2012) Shattered and stitched chromosomes-chromothripsis and chromoanasythesis-manifestations of a new chromosome crisis? *Gene Chromosome Cancer* 51:975–981
9. Sorzano CO, Pascual-Montano A, Sanchez de Diego A, Martinez AC, van Wely KH (2013) Chromothripsis: breakage-fusion-bridge over and over again. *Cell Cycle* 12:2016–2023
10. McClintock B (1942) The fusion of broken ends of chromosomes following nuclear fusion. *Proc Natl Acad Sci U S A* 28:458–463
11. Birchall PS, Fishpool RM, Albertson DG (1995) Expression patterns of predicted genes from the *C. elegans* genome sequence visualized by FISH in whole organisms. *Nat Genet* 11:314–320
12. Speicher MR, Carter NP (2005) The new cytogenetics: blurring the boundaries with molecular biology. *Nat Rev Genet* 6:782–792
13. Carlson RW, Moench SJ, Hammond ME, Perez EA, Burstein HJ et al (2006) HER2 testing in breast cancer: NCCN Task Force report and recommendations. *J Natl Compr Cancer Netw JNCCN* 4(Suppl 3):S1–S22; quiz S23–24

14. Gofrit ON, Zorn KC, Silvestre J, Shalhav AL, Zagaja GP et al (2008) The predictive value of multi-targeted fluorescent in-situ hybridization in patients with history of bladder cancer. *Urol Oncol* 26:246–249
15. Tanas MR, Goldblum JR (2009) Fluorescence in situ hybridization in the diagnosis of soft tissue neoplasms: a review. *Adv Anat Pathol* 16:383–391
16. van der Burg M, Poulsen TS, Hunger SP, Beverloo HB, Smit EM et al (2004) Split-signal FISH for detection of chromosome aberrations in acute lymphoblastic leukemia. *Leuk Off J Leuk Soc Am Leuk Res Fund UK* 18:895–908
17. Schrock E, du Manoir S, Veldman T, Schoell B, Wienberg J et al (1996) Multicolor spectral karyotyping of human chromosomes. *Science* 273:494–497
18. Speicher MR, Gwyn Ballard S, Ward DC (1996) Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 12:368–375
19. Liehr T, Starke H, Weise A, Lehrer H, Claussen U (2004) Multicolor FISH probe sets and their applications. *Histol Histopathol* 19:229–237
20. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW et al (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumours. *Science* 258:818–821
21. Snijders AM, Nowak N, Segreaves R, Blackwood S, Brown N et al (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 29:263–264
22. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A et al (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23:41–46
23. Huang J, Wei W, Zhang J, Liu G, Bignell GR et al (2004) Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics* 1:287–299
24. Bignell GR, Huang J, Greshock J, Watt S, Butler A et al (2004) High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res* 14:287–295
25. Wang Y, Cottman M, Schiffman JD (2012) Molecular inversion probes: a novel microarray technology and its application in cancer research. *Cancer Genet* 205:341–355
26. Nilsson M, Malmgren H, Samiotaki M, Kwiatkowski M, Chowdhary BP et al (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science* 265:2085–2088
27. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F et al (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 30:e57
28. Homig-Holzel C, Savola S (2012) Multiplex ligation-dependent probe amplification (MLPA) in tumour diagnostics and prognostics. *Diagn Mol Pathol Am J Surg Pathol B* 21:189–206
29. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N et al (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* 26:317–325
30. Teo SM, Pawitan Y, Ku CS, Chia KS, Salim A (2012) Statistical challenges associated with detecting copy number variations with next-generation sequencing. *Bioinformatics* 28:2711–2718
31. Zhao M, Wang Q, Jia P, Zhao Z (2013) Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinform* 14(Suppl 11):S1
32. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J et al (2012) Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28:423–425
33. Alkodsi A, Louhimo R, Hautaniemi S (2014) Comparative analysis of methods for identifying somatic copy number alterations from deep sequencing data. *Brief Bioinform* 16:242–254
34. Janevski A, Varadan V, Kamalakaran S, Banerjee N, Dimitrova N (2012) Effective normalization for copy number variation detection from whole genome sequencing. *BMC Genomics* 13(Suppl 6):S16
35. Chuang PT, Albertson DG, Meyer BJ (1994) DPY-27: a chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through association with the X chromosome. *Cell* 79:459–474