Review TRENDS in Genetics Vol.22 No.8 August 2006

Gene amplification in cancer

Donna G. Albertson

Cancer Research Institute and Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA 94143, USA

Gene amplification is a copy number increase of a restricted region of a chromosome arm. It is prevalent in some tumors and is associated with overexpression of the amplified gene(s). Amplified DNA can be organized as extrachromosomal elements, as repeated units at a single locus or scattered throughout the genome. Common chromosomal fragile sites, defects in DNA replication or telomere dysfunction might promote amplification. Some regions of amplification are complex, yet elements of the pattern are reproduced in different tumor types. A genetic basis for amplification is suggested by its relative frequency in some tumor subtypes, and its occurrence in 'early' preneoplastic lesions. Clinically, amplification has prognostic and diagnostic usefulness, and is a mechanism of acquired drug resistance.

Introduction

The development of tumors is associated with the acquisition of genetic and epigenetic alterations and the corresponding changes in gene expression that modify normal growth control and survival pathways. These changes can be brought about at the genomic level in a variety of ways, including altered karyotypes, point mutations and epigenetic mechanisms. Genomic DNA copy number aberrations are frequent in solid tumors and are expected to contribute to tumor evolution by copy number-induced alterations in gene expression. Cytogenetic and more recently molecular and array-based analytic methods (Box 1) have found great variation in the numbers and types of chromosome-level alterations present in human tumors – see the Mitelman Database of Chromosome Aberrations in Cancer (http://cgap.nci.nih.gov/Chromosomes/Mitelman) – which are likely to reflect the many different routes taken by individual tumors to escape normal protective mechanisms. Often they result in net gain or loss of whole chromosomes (aneuploidy) or parts of chromosomes (deletions, nonreciprocal translocations). Gene amplification, defined as a copy number increase of a restricted region of a chromosome arm, can also occur [1]. For example, Figure 1 shows DNA copy number profiles for three ductal invasive breast tumors as determined by array comparative genomic hybridization (array CGH). The tumors vary in the extent to which the genome has acquired low-level gains and losses of parts of chromosomes. In some tumors focal high-level amplifications are evident.

Corresponding author: Albertson, D.G. (albertson@cc.ucsf.edu).
Available online 19 June 2006

Gene amplification is a process that characterizes some normal developmental states, such as oogenesis in *Drosphila*, in which case it is strictly developmentally controlled [2]. However, studies carried out >20 years ago, in which bacteria, yeast, *Drosophila* or immortalized mammalian cells were challenged with a drug (Box 2), showed that drug-resistant cells were permissive for amplification because resistant cells with amplified target genes arose at frequencies of 10^{-7} to 10^{-4} [3]. Amplification was not detected in normal cells (frequency $<10^{-9}$), indicating that mechanisms exist in normal cells to prevent amplification [4,5]. Further experiments selecting for cells capable of growth in the presence of drugs in a variety of microbial and mammalian systems have extended these initial studies. In cultured mammalian cells, they show that gene amplification is likely to be initiated by a DNA double-strand break. Moreover, it can occur only in cells that are able to progress inappropriately through the cell cycle with this damaged DNA [6–10], whereas normal cells would arrest due to activation of cell-cycle checkpoints. Similarly, in mouse models, amplification can be promoted by double-strand breaks [11,12].

The analysis of amplified DNA in mammalian cell lines and tumors has revealed that it can be organized as extrachromosomal 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 various locations in the genome [1] (Figure 2). The unit of amplified DNA in some cases can involve sequences from two or more regions of the genome, indicating a complex process of formation involving multiple chromosomes [11–16]. The unit of amplified DNA in tumors can range in size from kilobases to tens of megabases. The organization of amplicons encompassing *MYCN*, *EGFR* and *ERBB2*, three genes frequently amplified in cancers, is summarized in Box 3.

Recent applications of genome-wide scanning techniques facilitate identification of regions of tumor genomes that are often amplified and thus likely to harbor genes of importance for tumor development. For cancer biologists and clinicians, gene amplification has utility for identifying novel oncogenes, and significance for diagnosis, patient prognosis, response to therapy and development of resistance. Here we review some areas of active research focusing on gene amplification.

Amplicons pinpoint candidate oncogenes or oncogenic pathways

A longstanding interest in mapping amplicons stems from the notion that amplification of a genomic region

 0168-9525/$ - see front matter doi:10.1016/j.tig.2006.06.007

Box 1. Molecular methods for mapping and measuring amplification

Comparative genomic hybridization (CGH)
CGH detects and maps DNA sequence copy number variation throughout the normal genome. In array CGH, relative copy number is measured at specific loci, which are represented by arrays of mapped clones or oligonucleotides. Fluorescently labeled test and reference DNAs are hybridized to the arrayed elements together with unlabeled Cot-1 DNA to suppress hybridization from repeated sequences. The resulting ratio of the fluorescence intensities of the two fluorochromes at each locus is approximately proportional to the ratio of the copy numbers of the corresponding DNA sequences in the test and reference genomes. The previous use of metaphase chromosomes as the hybridization target limited the resolution of CGH to 10–20 Mb, prohibiting resolution of closely spaced aberrations, and only allowing linkage of CGH results to genomic information with cytogenetic accuracy. In array CGH, the resolution is determined by the genomic spacing of the array elements [73,74].

Fluorescent *in situ* hybridization (FISH)
FISH refers to a variety of cytogenetic applications in which one or more nucleic acid probes are labeled with a fluorochrome-conjugated nucleotide or other hapten that can be detected by fluorescently labeled molecules. The labeled probe(s) is then hybridized to whole organisms, tissue sections, cells or subcellular constituents such as metaphase chromosomes, nuclei or extended chromatin fibers, and the site of the nucleic acid sequence is visualized by fluorescence microscopy [75,76].

Digital karyotyping
Genomic DNA copy number aberrations are mapped by quantifying short DNA fragments. The DNA fragments are generated by initially digesting the DNA with a restriction enzyme ('mapping enzyme') to generate ~10-kb pieces. Biotinylated linkers are added and then a second digest is carried out with an enzyme that recognizes a four-base site ('fragmenting enzyme'). The biotinylated DNA fragments are recovered and a linker with a *Mme*I site is added. *Mme*I is a type IIS restriction enzyme that cleaves DNA 21 bp from its recognition site and thus will generate 21-bp tags, derived from sequences adjacent to the fragmenting enzyme site nearest to the site cleaved initially by the mapping enzyme. The tags are ligated together, sequenced and computationally matched to the genome sequence and quantified [77].

Quantitative microsatellite analysis (QuMA)
QuMA measures relative DNA sequence copy number of a test locus relative to a pooled reference by using quantitative, real-time PCR amplification of loci carrying simple sequence repeats [78].

BAC end sequencing
The use of bacterial artificial chromosome (BAC) end sequences to map and measure genomic aberrations has been termed ESP (end sequence profling). In this approach, a BAC library is constructed from a test genome, BAC end sequences are obtained and the end-sequence pairs are mapped onto the normal genome sequence. The density of BAC end-sequences provides information on copy number. Structural aberrations can also be recognized by BAC end pairs that map abnormally far apart on the normal sequence [79].

represents selection for expression of gene(s) that promote growth of the tumor. Thus, genes mapping within amplicons are candidate oncogenes. Genome-scanning techniques now enable researchers to identify recurrent regions of amplification, to define efficiently amplicon boundaries and to use the genome databases to identify candidate oncogenes in the region. The high-throughput nature of array-based genome scanning also allows rarer, perhaps small amplicons to be found that are particularly informative for identifying candidate oncogenes or pathways (or both) [17,18]. Nevertheless, proving that a gene within a region is the driver gene for that amplicon remains problematic, because several genes might map in the amplified region. In general, one narrows the field of candidates based on whether the candidate oncogene is expressed in the tumor, whether expression is enhanced when amplified, and perhaps also whether it is found to be overexpressed in tumors by other mechanisms when not amplified. The shape of the amplicon or variation in amplitude across the amplified region has also been used as a guide [19]. The rationale for this approach is based on the observation that under continued selection pressure amplicons can become narrower and more focused on the gene under selection. Thus, the driver gene for amplification is expected to map at the region of maximal amplitude. Subsequent studies to investigate the function of the candidate oncogene in promoting tumor formation are then required to confirm the candidate. The design of these confirmatory studies will be dictated by the nature of the probable function of the gene in cancer.

Models for initiation of amplification

Bearing in mind that gene amplification is likely to be initiated by a DNA double-strand break, in cells lacking robust checkpoints [6–10], processes and genome sequences likely to promote breakage and subsequent genome rearrangements consistent with observed amplicon structures have been proposed. Much attention has been focused on errors in DNA replication, telomere dysfunction and contributions from specific genome sequences.

DNA replication

A role for errors in DNA replication promoting gene amplification was proposed by Alexander Varshavsky [20] >20 years ago and subsequently elaborated on by Robert Schimke and colleagues [21]. Although these initial proposals, based on extra rounds of replication because of misfiring of replication origins, appear to be incorrect, modification of models invoking replication of extrachromosomal DNA [22], possibly formed by collapse of replication bubbles with or without subsequent reintegration into a chromosome, have been proposed to explain the formation of *MYCN* [23], *EGFR* [24] and *ERBB2* [25] amplicons. Double-strand DNA breaks are frequent in replicating cells and can be generated by collapse of replication forks unable to progress as a result of encountering lesions in the DNA, for example; hence they provide an opportunity to initiate the amplification process.

Telomere dysfunction

Telomeres function to prevent the loss of DNA sequences that occurs as a consequence of the incomplete replication of linear DNA at chromosome ends. They are replicated by telomerase, which is inactive in most somatic cells, and thus at each generation the ends of the DNA shorten by 40–50 bp. When telomeres shorten to a critical length,

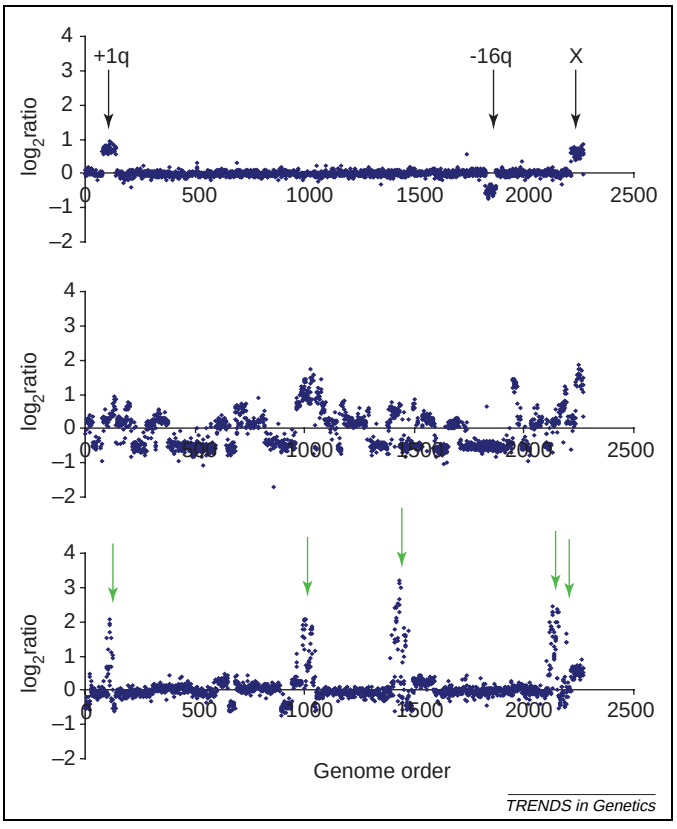

Figure 1. Copy number profiles of breast tumors. Normalized copy number ratios of genomic DNA from breast tumors compared with normal male reference DNA. The normalized log2 ratio is plotted for each clone sorted by chromosome and ordered according to genome position from the p arm to the q arm (clones are numbered sequentially on the *x*-axis). The tumors vary widely in the number of low-level copy number changes. The tumor profile in the top panel shows low-level copy number changes in the tumor involving gain of chromosome 1q (+1q) and loss of chromosome 16q (−16q). The increased copy number of the X chromosome reflects the difference in copy number of the X chromosome between females and males. The tumor in the middle panel shows low-level copy number changes involving almost all chromosomes. The tumor profile in the bottom panel shows low-level gains and losses, and amplification on chromosomes 1, 8, 11, 20 and 21 (green arrows).

cellular senescence is induced and normal cells cease proliferating. In most tumors, however, cells have acquired the capability to maintain telomere length through reactivation of telomerase, or by using a recombination-based mechanism, alternate lengthening of telomeres (ALT). However, before adapting to telomere dysfunction, the tumor cells could have continued proliferating with such short telomeres that the ends of the chromatids were susceptible to fusion and subsequent abnormal cell division [26,27]. A role for telomere dysfunction in promoting gene amplification is supported by the observation that the genomes of tumors arising in mice lacking the RNA component of telomerase ($Terc^{-/-}$ mice) contain amplifications and deletions commonly associated with human tumors. Such rearrangements are rare in most mouse tumor models, presumably because of the longer murine telomeres and also the efficient tumor promotion afforded by many of the oncogenes and tumor suppressors modeled in mice, which might obviate the need for genomic instability in tumor development in these mouse models [28,29]. Telomere attrition is observed in human tissues in preneoplastic stages, supporting a role for this mechanism in generating genomic aberrations in development of human tumors with high levels of chromosomal rearrangement [30–32].

Amplification as a result of telomere dysfunction in addition to breakage at fragile sites (discussed below) is thought to occur through breakage–fusion–bridge (BFB) events as originally proposed in maize [33]. In this process, a chromosomal break or dysfunctional telomere promotes fusion of chromosome ends, resulting in formation of a dicentric chromosome. At anaphase, breakage occurs, which, depending on the position of the break, generates a chromosome with an inverted duplication of terminal sequences. Because this chromosome also has broken ends the process can repeat in the next and subsequent divisions until the chromosomal ends are stabilized by addition of telomeric sequences, possibly by translocation to another chromosome. This model results in chromosomal amplicons often at the site of the gene under selection, in which the units are arranged as inverted repeats. Loss of the DNA sequences distal to the gene under selection, or their translocation to another chromosome, is also expected to occur. Evidence of BFB

Box 2. Studies of gene amplification *in vitro*

Amplification has been studied most often in mammalian cells by challenging cells with *N*-phosphonacetyl-L-aspartate (PALA) or methotrexate and quantifying the rate or frequency with which resistant colonies arise. Typically, a small number of cells or cells expanded from a single cell are challenged with drug to minimize contributions from pre-existing drug-resistant variants in the cell culture. A Luria–Delbrück analysis can be used to distinguish between resistance resulting from pre-existing resistant variants and adaptation to the drug challenge [80]. Colonies are also expanded and subjected to copy number analyses (Box 1).

In rodent cells, PALA resistance is almost always achieved by amplification of the target gene, carbamoyl-synthetase 2 aspartate transcarbamylase and dihydro-orotase (*Cad*), in tandem arrays on the chromosome that carries the normal copy of the gene. In human cells, however, the particular mechanism seems to depend on the organization of chromosome 2, which carries the *CAD* locus in the cell line being studied. Amplification of *CAD* occurs in cell lines with rearranged copies of chromosome 2. In human cells with few or no rearrangements involving chromosome 2, mechanisms other than amplification are used, including increases in aspartate transcarbamylase activity without altered copy number of *CAD*, or gain of an extra copy of the gene resulting from duplication of chromosome 2 or isochromosome 2p formation [81].

Methotrexate is an analog of folic acid. It kills cells by inhibiting dihydrofolate reductase (DHFR), which catalyzes reduction of dihydrofolate to tetrahydrofolate used in glycine, purine and thymidylate syntheses. The initial demonstration of gene amplification in mammalian cells employed selection for methotrexate resistance and subsequent analysis of *DHFR* copy number in resistant cells by nucleic acid reassociation [82]. Resistance to methotrexate is achieved by amplification of *DHFR*, but also by enhanced activity of *DHFR*, mutation to decrease affinity for methotrexate and alterations in transport of the drug into and out of the cell. Thus, when assessing whether amplification is enhanced in a particular genetic background following challenge with PALA or methotrexate, it is necessary to show that copy number of the target locus is increased; simply determining the frequency or rate of resistant colonies in response to drug challenge is not sufficient.

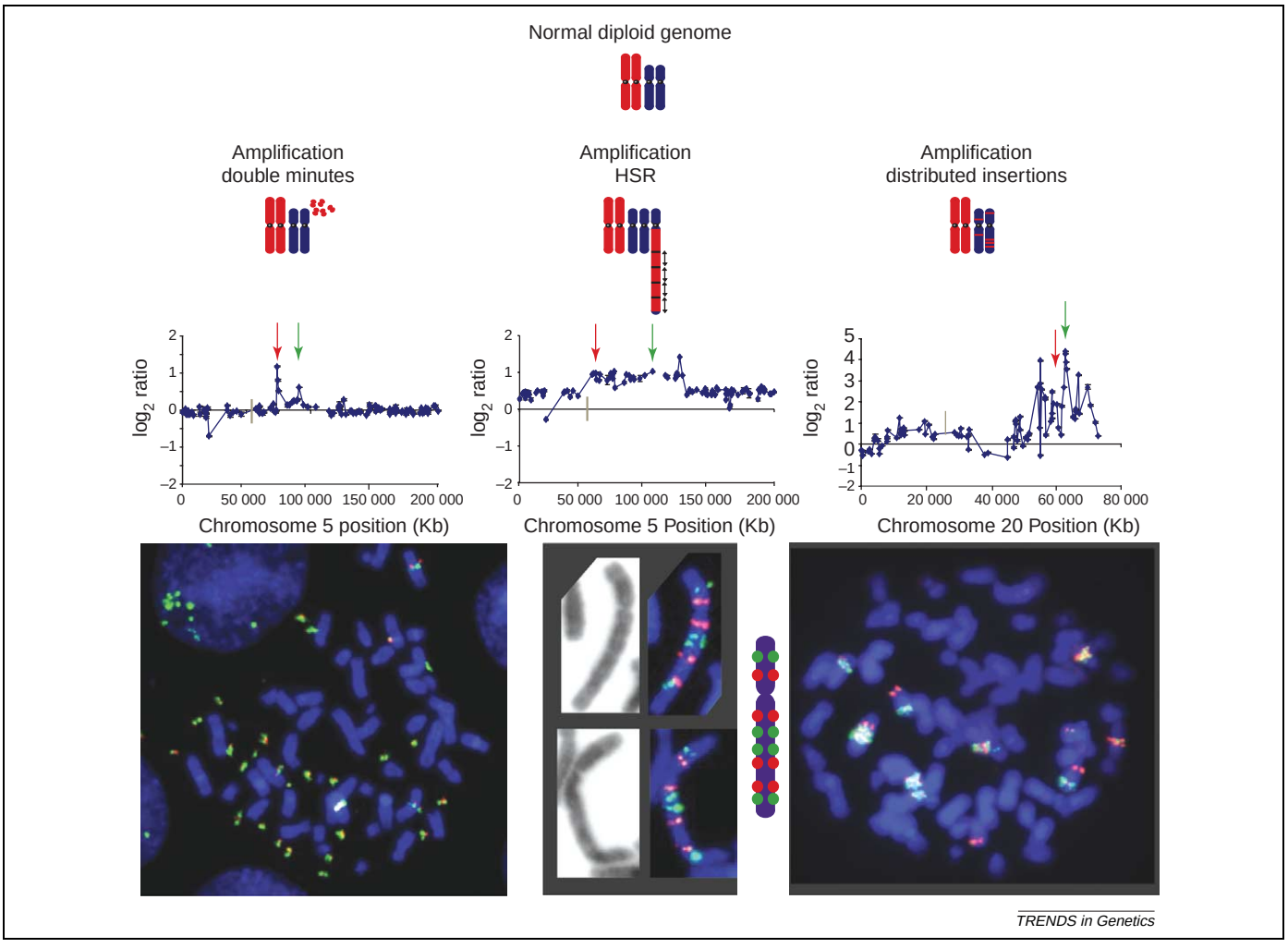

Figure 2. Cytogenetics of gene amplification. Amplified DNA can be present in various forms including double minutes, amplified regions on a chromosome, which can be cytogenetically visible as a HSR, or distributed across the genome. A two-chromosome genome is depicted at the top of the figure and possible amplification structures are drawn below. Examples of array CGH copy number profiles (plotted as the normalized $\log_2$ ratio) for each type of amplicon structure are shown with corresponding FISH pictures of the cells using BAC clones from the region of the amplicon indicated by the red and green arrows. Many red and green signals can be seen in the double minutes in a methotrexate-resistant human cell line (lower left panel). The amplified DNA in another methotrexate-resistant cell line is present as inverted repeats in the chromosomes [41] (lower center panel). Images of the chromosomes stained with the DNA-binding compound DAPI (4′-6-diamidino-2-phenylindole) are shown on the left. The inverted repeat organization is shown by the alternating green, red, red, green pattern of FISH signals repeated from top to bottom along the chromosomes in the composite FISH images (middle) and in the diagram on the right hand side. Amplified DNA from a region of chromosome 20q is distributed throughout the genome of MCF7 breast cancer cells [83] (bottom right panel). The fluorescence micrographs are courtesy of Antoine Snijders and Wen Lin Kuo.

Box 3. Organization of common amplicons in tumors

Amplification of *MYCN* has been reported variably to occur in ~20% of neuroblastomas. It is associated with more aggressive disease and was one of the first tumor amplicons to be studied in detail. The *MYCN* amplicons are visible as HSRs on chromosomes other than at the *MYCN* locus at chromosome 2p24. Amplification levels are typically 50- to 100-fold, and the amplified DNA ranges from 0.1 to >1 Mb. Core kilobase-sized units are organized as head-to-tail direct tandem repeats [23].

Amplification of *EGFR* occurs in 40% of gliomas. Deletions and point mutations in *EGFR* are found in these tumors and it is likely that amplification precedes mutation. Recent detailed studies revealed that the amplified DNA is present on double minute chromosomes, which ranged in size from 0.7 to 2.1 Mb and in level of amplification from 8- to 63-fold [24]. Within each tumor, amplicons were similar, leading to the conclusion that initially an extrachromosomal circular molecule encompassing *EGFR* was generated by joining the ends of an excised DNA fragment. The breakpoints of the amplified DNA in seven tumors were spread over a ~1 Mb region around *EGFR* and occurred in single-copy sequences or in interspersed repetitive elements, including Alus, a long interspersed nucleotide element (LINE) and a medium reiteration frequency repeat (MER), and in a segmental duplication. Microhomology was found in the sequences at the ends of the amplified DNA, implicating microhomology-based NHEJ in the process, rather than homologous recombination, which requires longer stretches of homology. As in the case of *MYCN*, *EGFR* was not deleted from its endogenous location on 7p.

Amplification of *ERBB2* at 17q21 occurs in breast, ovarian and gastric cancers, with amplification levels of 20-fold in some breast cancers, for example. Structural analysis of *ERBB2* amplicons in cell lines of these tumor types and in gastric tumors revealed a head-to-tail organization with uniform amplicon junctions; microhomology-mediated end joining was also evident [25].

cycles generating amplicons has been obtained by cytogenetic analyses of HSRs in tumor cell lines and in model systems with amplifications following challenge with drugs [34].

Fragile sites

The role of genomic context in promoting amplification was initially investigated >20 years ago in studies that showed that different regions of the genome were more prone to amplify than others, although the basis for this difference was unknown [35]. The contribution of the genome sequence to the propensity to amplify is still of considerable interest. The genomic sequences receiving the most attention are common chromosomal fragile sites, which are chromosomal regions that are late replicating and prone to breakage under conditions of replication stress. There are ~100 known fragile sites in the human genome, as defined by treating cells with agents that interfere with DNA replication. The sites are further distinguished depending on their sensitivity to the drugs used to induce their expression, such as aphidicolin. As a group, fragile sites are heterogeneous. Those that have been molecularly characterized seem to extend over broad regions 0.3–9-Mb long. The DNA sequences can display particular characteristics of flexibility, for example, and are associated with a high frequency of recombinogenic events, including colocation at cytogenetic resolution with sites of recurrent chromosome aberrations associated with various cancers [36]. Thus, they are attractive candidates for regions prone to breakage that could initiate the amplification process.

If one considers the boundaries of the amplified region as markers of original DNA strand breaks initiating the amplification, evidence for the involvement of fragile sites in this process can be found, because the boundaries of several amplicons generated in hamster cells have been mapped to common chromosomal fragile sites [7,8,37]. In human cell lines there are fewer examples supporting a role for fragile sites [38,39], but there is evidence that they participate in amplification in human tumors, because the boundaries of *MET* amplicons in esophageal adenocarcinoma map within the fragile site FRA7G [40]. However, in a human cell culture system in which amplicons encompassing *DHFR* were induced in response to methotrexate challenge, no evidence of recurrent amplicon boundaries was found using array CGH, suggesting that breakage at a limited number of specific genomic sites did not contribute to the process [41]. Because the frequency of expression of fragile sites varies, it could be that the aphidicolin-sensitive fragile sites FRA5D, FRA5F and FRA5C, which map distal to *DHFR* on 5q, are infrequently expressed in this experimental system and thus less likely to contribute to the amplification process. In addition, it could be necessary for the gene under selection to be in close proximity to fragile sites, similar to the *MET* amplicon. Currently, most fragile sites have been mapped only to cytogenetic resolution. Thus, definitive information regarding their role in setting other amplicon boundaries awaits molecular mapping of additional fragile sites. Other candidate classes of sequences that might be involved in generating amplicons include repetitive elements such as Alu sequences, which could promote amplification through inappropriate homologous recombination-based mechanisms; or shorter sequences promoting microhomology-mediated joining, which could be more prevalent (Box 3). Application of improved high-throughput sequence-based analysis methods for characterizing the organization of amplified DNA and sequences at amplicon boundaries might help to shed light on this question in human tumors.

Comparative genomics of amplicons

Recent array CGH studies of oral squamous cell carcinoma (SCC) [18], melanoma [42] and breast adenocarcinoma [84] provide the opportunity to investigate how amplicons compare in tumors from different tissues. These three tumors amplify a region at 11q13, including *CCND1*. Array CGH copy number profiles of the region show at least eight amplicon cores centered on previously identified oncogenes *RIN1* (amplicon A), *CCND1* (amplicon C), *WNT11/EMSY* (amplicon E), *GARP* (amplicon F), *PAK1* (amplicon G) and three additional regions for which candidate genes have not been identified (Figure 3). The complexity of the amplicon raises the question of whether the cores are a by-product of the amplification process or the result of selection. The amplicons and amplicon boundaries encompassing *CCND1* and *PAK1* are common to all three tumor types, but amplification of the region encompassing the amplicon cores D, E and F, which are present in breast adenocarcinoma and melanoma, is rare in oral SCC as noted previously [43].

The similarity of the 11q13 amplicon among these tumor types is particularly interesting, because they represent three different types of tumors arising in tissues with different developmental origins. Moreover, it is notable that whereas mutations in *TP53* are highly correlated with amplification of *CCND1* in oral SCC (Fisher exact test, $P=0.009$) [18], *TP53* mutations in breast cancer are not associated with *CCND1* amplification (or other recurrent amplicons), although other copy number alterations are significantly associated with *TP53* mutations [84]. The etiology of these tumor types is also likely to be different. In melanoma, *CCND1* amplification is most prevalent in acral and mucosal melanoma subtypes [42]. Factors promoting these melanomas are unknown, but because these occur on sun-protected skin they are not associated with UV exposure. By contrast, smoking is a known risk factor for oral SCC [44], and arguably could also increase risk for breast cancer [45]. Chromosomal aberrations indicative of breakage have been reported to be more prevalent in lymphocytes of smokers compared with nonsmokers [44,46], thus smoking might enhance breakage at sensitive sites, as discussed above, creating cells permissive for amplification and setting amplicon boundaries. Amplicon boundaries might also be determined by selection for enhanced expression of genes in the amplified regions, or selection against copy number-induced changes in expression of genes mapping outside the amplicon cores. The absence of amplicons D–F in oral SCC, and the demonstration that genes in these different amplicons have oncogenic or transforming potential (or both)

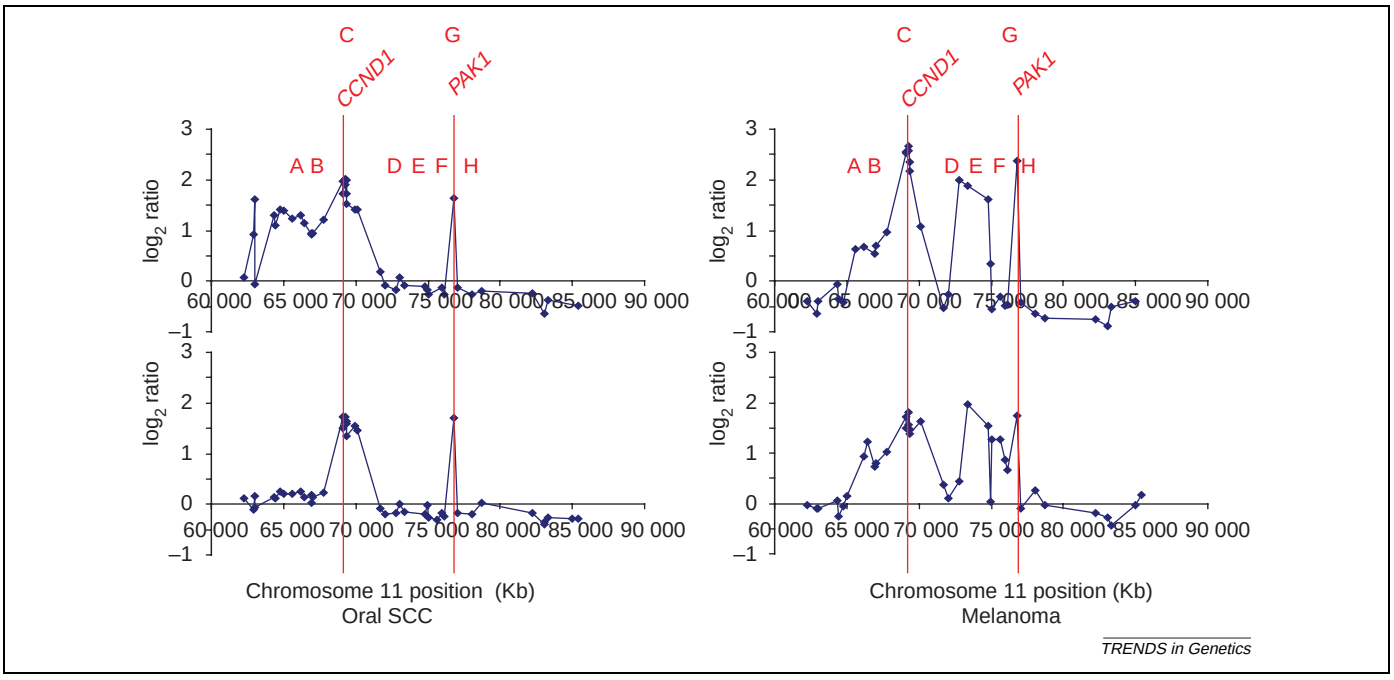

Figure 3. Copy number profiles of *CCND1* amplification in oral squamous cell carcinoma (SCC) and melanoma. Two copy number profiles of oral SCC (left) and melanoma (right) are shown in a ~3-Mb region encompassing *CCND1* and *PAK1*. The amplicon peaks at *CCND1* (C) and *PAK1* (G) are clearly visible in both tumor types, but additional amplicon cores are present in the melanoma profiles. The approximate positions of amplicon cores A–G are indicated.

indicate that the complexity of the amplicon is not simply a by-product of the amplification process. It remains to be seen whether the same oncogenes or pathways (or both) mapping in these amplicon cores are functioning similarly in these different tissues or if different oncogenes are driving amplification in the different tissue types.

Are there genes that promote amplification (an amplifier phenotype)?

One expects that the genetic alterations seen in tumors reflect underlying failures in mechanisms that maintain genetic stability, in addition to selection for changes that provide a growth advantage. This expectation is most clearly seen for hereditary and sporadic tumors with defects in mismatch repair, in which mutations at the nucleotide level drive tumorigenesis and there are few chromosomal level alterations [41,47]. Similarly, hereditary *BRCA1* mutant tumors display a high level of chromosomal alterations, which are distinct from the spectrum of alterations seen in *BRCA2* mutant tumors [48]. Neither hereditary *BRCA1* breast tumors, nor mismatch repair-defective tumors have a high frequency of amplifications.

The question arises then as to whether there is evidence for a genetic basis for amplification in tumors. Two observations lend support to a possible genetic basis. First, amplification can occur early, because it has been observed in the dysplastic or benign, but potentially preneoplastic, tissue adjacent to carcinoma [49,50] and in histologically normal melanocytes adjacent to *in situ* melanoma [51], indicating that fields of amplified cells exist. Thus, in these cases, amplification is not a 'late' event in tumorigenesis, so that it is not simply an adventitious feature of highly rearranged genomes of advanced tumors. Moreover, in breast tumors, amplifications occurred in tumors with relatively few additional types of chromosomal copy number aberrations, and in those with many [84]. Second, the presence of multiple regions of amplification in some tumors lends support to a possible genetic basis. For example, one of the three breast tumor subtypes defined on the basis of copy number aberrations is distinguished by the presence of amplifications (Figure 1) [84]. Co-amplification of the more commonly amplified genes in breast cancer has been reported to occur at a greater frequency than expected from the frequency of amplification of the individual genes [52]. Furthermore, the genomic alterations in almost all mucosal and acral melanomas include amplifications (85% and 89% of tumors, respectively), but they are rare in other melanoma types [42]. The genes that are amplified in mucosal and acral melanoma are frequently overexpressed without amplification in the more common form of melanoma, superficial spreading melanoma, indicating that the propensity to amplify is the mechanism of genetic instability that drives the evolution of acral and mucosal melanomas. By contrast, prostate tumor genomes are characterized by fewer DNA copy number changes and many of the aberrations detected by array CGH are deletions [53], suggesting that a different type of genome instability is active in this tumor type.

Studies in hamster cell lines and yeast have provided evidence that cells differ in their permissiveness or propensity to amplify, originally referred to as an 'amplificator' phenotype [54]. The underlying genetic changes in four hamster cell lines with enhanced amplification have not been identified, whereas cells with mutations in *ERCC6* do show the amplificator, or amplifier, phenotype [55]. Similarly, deficiency for the DNA-dependent protein kinase catalytic subunit (*Prkdc*,

DNA-PKcs) enhances amplification in hamster cells permissive for amplification [56]. Overexpression of several genes, including *Myc*, *Hras1*, *Rrm2*, *Fgf2*, *CCND1* and mutant *JUN* also enhances the frequency of amplification, whereas wild-type *TP53* is necessary but not sufficient to maintain cells in a nonpermissive state with respect to amplification [57–60]. Mutations in *RAD3* (human homolog, *XPD*), *MSH2* and *RAD52* have been shown to affect the rate of amplification in a yeast system [61], but it is unknown whether they influence the frequency of amplification in human cells.

Deficiencies in nonhomologous end joining (NHEJ) DNA repair and *Trp53* function together have been implicated in amplicon formation in murine tumors. In this particular case, the amplification process seems to be initiated by programmed double-strand DNA breaks in lymphocytes during *V(D)J* recombination, which occurs in the course of T- and B-cell development. The doubly deficient NHEJ/*Trp53* animals are susceptible to pro-B lymphomas, characterized by amplicons involving translocations that place the IgH locus adjacent to the *Myc* or *Mycn* locus [62]. These amplicons resemble the complex amplicons comprising multiple genomic regions that are found in human tumors; however, such amplicons are not usually found in human lymphomas.

Numerous genes have been shown to have chromosomal instability phenotypes when knocked out, knocked down or overexpressed in human cells; however, in most cases analyses of their instability phenotypes have not rigorously determined whether gene amplification is enhanced. Thus, although there is evidence suggesting that specific genetic defects in tumor cells might underlie the amplifier phenotype, currently we can only speculate that altered expression of genes that participate in processes such as mitosis, DNA replication, repair, cell-cycle checkpoints and telomere maintenance are likely candidates. This is an area for further research.

Clinical considerations

Amplification in general and in specific amplicons has importance for both prognosis and targeted therapies. For example, amplifications of *MYC*, *ERBB2*, *CCND1*, *EGFR* or *MDM2* have been reported to be individually significantly associated with high grade in breast cancer [52], whereas decreasing survival is associated with number of amplifications [52,63]. Specific amplicons, such as *MYCN* in neuroblastoma [23], and *MYC* and *ERBB2* in breast cancer, also have prognostic significance [52].

Amplified genes are also targets of specific therapies, because the tumor can become 'addicted' to their enhanced expression [64]. The reliance of tumors, compared with normal tissues, on a continued high level of expression of the oncogene provides a differential sensitivity to therapies targeting the oncogene. The use of specific antibodies against ERBB2 (Herceptin) to treat breast cancer patients with amplified *ERBB2* is an example, which also illustrates the diagnostic utility of the amplification itself, because measuring *ERBB2* copy number by fluorescent *in situ* hybridization (FISH) has been shown to have the greatest predictive power for response to Herceptin [65]. In the clinic, amplification is observed to be one of the underlying causes of resistance to therapy. Examples include amplification of *DHFR* in response to methotrexate [66], *TYMS* in response to 5-fluorouracil [67] and *BCR-ABL* in response to Imatinib mesylate [68]. However, recent studies of the response of lung cancer patients to the small molecule tyrosine kinase inhibitor, gefitinib, which most effectively targets *EGFR* with gain-of-function mutations in the kinase domain, indicate that patients with *EGFR* amplification have improved survival [69]. These patients had greater response rates and longer times to disease progression, although they did not experience longer overall survival, leading to the suggestion that *EGFR* amplification might be associated both with survival and resistance to gefitinib. Resistance to gefitinib has been associated with a specific second mutation in exon 20 of the kinase domain. It has been suggested that improved outcome with amplified *EGFR* might be explained by assuming that a minimal level of expression of the mutant alleles is required for resistance. Thus, acquisition of more alleles with the resistance mutation would be required to overcome greater dilution by the increased number of *EGFR* copies in cells with amplification [70]. Amplification of mutant *Ras* genes to overcome tumor suppressive effects of the wild-type gene in mouse models is an example of this type of phenomenon [71,72].

Concluding remarks

Gene amplification in human cancer provides a means of overexpression of oncogenes. The amplified DNA serves both as a diagnostic marker and an indicator that the amplified region of the genome harbors candidate oncogenes likely to have a role in tumor development. Thus amplicons are currently both diagnostic and research tools. Observations support a genetic basis for propensity to amplify in tumors, but the genes or processes promoting this type of genetic instability in tumors have not been identified. Because amplification is often associated with poor prognosis and is a mechanism of resistance to therapies, it will be important to identify the genes or pathways (or both) that promote amplification in tumors so that they might be targeted as part of combination therapy to prevent evolution of resistance to drugs designed to arrest or eradicate the tumor.

Acknowledgements

Work in the author's laboratory was supported by National Institutes of Health grants CA84118, CA90421, CA94407 and CA101359.

References

1 Albertson, D.G. *et al*. (2003) Chromosome aberrations in solid tumors. *Nat. Genet.* 34, 369–376
2 Tower, J. (2004) Developmental gene amplification and origin regulation. *Annu. Rev. Genet.* 38, 273–304
3 Stark, G.R. and Wahl, G.M. (1984) Gene amplification. *Annu. Rev. Biochem.* 53, 447–491
4 Tlsty, T.D. (1990) Normal diploid human and rodent cells lack a detectable frequency of gene amplification. *Proc. Natl. Acad. Sci. U. S. A.* 87, 3132–3136
5 Wright, J.A. *et al*. (1990) DNA amplification is rare in normal human cells. *Proc. Natl. Acad. Sci. U. S. A.* 87, 1791–1795

6 Chernova, O.B. *et al*. (1998) MYC abrogates p53-mediated cell cycle arrest in N-(phosphonacetyl)-L-aspartate-treated cells, permitting CAD gene amplification. *Mol. Cell. Biol.* 18, 536–545
7 Coquelle, A. *et al*. (1997) Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* 89, 215–225
8 Kuo, M.T. *et al*. (1994) Chromosome breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug-resistant CHO cells. *Mol. Cell. Biol.* 14, 5202–5211
9 Paulson, T.G. *et al*. (1998) Gene amplification in a p53-deficient cell line requires cell cycle progression under conditions that generate DNA breakage. *Mol. Cell. Biol.* 18, 3089–3100
10 Pipiras, E. *et al*. (1998) Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. *EMBO J.* 17, 325–333
11 Difilippantonio, M.J. *et al*. (2002) Evidence for replicative repair of DNA double-strand breaks leading to oncogenic translocation and gene amplification. *J. Exp. Med.* 196, 469–480
12 Zhu, C. *et al*. (2002) Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* 109, 811–821
13 Bautista, S. and Theillet, C. (1998) CCND1 and FGFR1 coamplification results in the colocalization of 11q13 and 8p12 sequences in breast tumor nuclei. *Genes Chromosomes Cancer* 22, 268–277
14 Guan, X.Y. *et al*. (1994) Identification of cryptic sites of DNA sequence amplification in human breast cancer by chromosome microdissection. *Nat. Genet.* 8, 155–161
15 Lafage, M. *et al*. (1992) Fusion and amplification of two originally non-syntenic chromosomal regions in a mammary carcinoma cell line. *Genes Chromosomes Cancer* 5, 40–49
16 Muleris, M. *et al*. (1995) Identification of amplified DNA sequences in breast cancer and their organization within homogeneously staining regions. *Genes Chromosomes Cancer* 14, 155–163
17 Brown, L.A. *et al*. (2006) Amplification of EMSY, a novel oncogene on 11q13, in high grade ovarian surface epithelial carcinomas. *Gynecol. Oncol.* 100, 246–270
18 Snijders, A.M. *et al*. (2005) Rare amplicons implicate frequent deregulation of cell fate specification pathways in oral squamous cell carcinoma. *Oncogene* 24, 4232–4242
19 Albertson, D.G. *et al*. (2000) Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. *Nat. Genet.* 25, 144–146
20 Varshavsky, A. (1981) On the possibility of metabolic control of replicon 'misfiring': relationship to emergence of malignant phenotypes in mammalian cell lineages. *Proc. Natl. Acad. Sci. U. S. A.* 78, 3673–3677
21 Schimke, R.T. *et al*. (1986) Overreplication and recombination of DNA in higher eukaryotes: potential consequences and biological implications. *Proc. Natl. Acad. Sci. U. S. A.* 83, 2157–2161
22 Windle, B.E. and Wahl, G.M. (1992) Molecular dissection of mammalian gene amplification: new mechanistic insights revealed by analyses of very early events. *Mutat. Res.* 276, 199–224
23 Savelyeva, L. and Schwab, M. (2001) Amplification of oncogenes revisited: from expression profiling to clinical application. *Cancer Lett.* 167, 115–123
24 Vogt, N. *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
25 Kuwahara, Y. *et al*. (2004) Alternative mechanisms of gene amplification in human cancers. *Genes Chromosomes Cancer* 41, 125–132
26 Londono-Vallejo, J.A. (2004) Telomere length heterogeneity and chromosome instability. *Cancer Lett.* 212, 135–144
27 Sabatier, L. *et al*. (2005) The loss of a single telomere can result in instability of multiple chromosomes in a human tumor cell line. *Mol. Cancer Res.* 3, 139–150
28 O'Hagan, R.C. *et al*. (2002) Telomere dysfunction provokes regional amplification and deletion in cancer genomes. *Cancer Cell* 2, 149–155
29 Chang, S. *et al*. (2001) Modeling chromosomal instability and epithelial carcinogenesis in the telomerase-deficient mouse. *Semin. Cancer Biol.* 11, 227–239
30 Maser, R.S. and DePinho, R.A. (2002) Connecting chromosomes, crisis, and cancer. *Science* 297, 565–569
31 Gisselsson, D. *et al*. (2005) Telomere-mediated mitotic disturbances in immortalized ovarian epithelial cells reproduce chromosomal losses and breakpoints from ovarian carcinoma. *Genes Chromosomes Cancer* 42, 22–33
32 Gisselsson, D. and Hoglund, M. (2005) Connecting mitotic instability and chromosome aberrations in cancer – can telomeres bridge the gap? *Semin. Cancer Biol.* 15, 13–23
33 McClintock, B. (1942) The fusion of broken ends of chromosomes following nuclear fusion. *Proc. Natl. Acad. Sci. U. S. A.* 28, 458–463
34 Murnane, J.P. and Sabatier, L. (2004) Chromosome rearrangements resulting from telomere dysfunction and their role in cancer. *Bioessays* 26, 1164–1174
35 Wahl, G.M. *et al*. (1984) Effect of chromosomal position on amplification of transfected genes in animal cells. *Nature* 307, 516–520
36 Glover, T.W. *et al*. (2005) Mechanisms of common fragile site instability. *Hum. Mol. Genet.* 14 Spec No. 2, R197–205
37 Coquelle, A. *et al*. (1998) A new role for hypoxia in tumor progression: induction of fragile site triggering genomic rearrangements and formation of complex DMs and HSRs. *Mol. Cell* 2, 259–265
38 Ciullo, M. *et al*. (2002) Initiation of the breakage-fusion-bridge mechanism through common fragile site activation in human breast cancer cells: the model of PIP gene duplication from a break at FRA7I. *Hum. Mol. Genet.* 11, 2887–2894
39 Hellman, A. *et al*. (2002) A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell* 1, 89–97
40 Miller, C.T. *et al*. (2006) Genomic amplification of MET with boundaries within fragile site FRA7G and upregulation of MET pathways in esophageal adenocarcinoma. *Oncogene* 25, 409–418
41 Snijders, A.M. *et al*. (2003) Shaping of tumor and drug-resistant genomes by instability and selection. *Oncogene* 22, 4370–4379
42 Curtin, J.A. *et al*. (2005) Distinct sets of genetic alterations in melanoma. *New Engl. J. Med.* 353, 2135–2147
43 Huang, X. *et al*. (2002) High-resolution mapping of the 11q13 amplicon and identification of a gene, TAOS1, that is amplified and overexpressed in oral cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11369–11374
44 Johnson, N. (2001) Tobacco use and oral cancer: a global perspective. *J. Dent. Educ.* 65, 328–339
45 Terry, P.D. and Rohan, T.E. (2002) Cigarette smoking and the risk of breast cancer in women: a review of the literature. *Cancer Epidemiol. Biomarkers Prev.* 11, 953–971
46 Kao-Shan, C.S. *et al*. (1987) Increased fragile sites and sister chromatid exchanges in bone marrow and peripheral blood of young cigarette smokers. *Cancer Res.* 47, 6278–6282
47 Rajagopalan, H. and Lengauer, C. (2004) Aneuploidy and cancer. *Nature* 432, 338–341
48 van Beers, E.H. *et al*. (2005) Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. *Cancer Res.* 65, 822–827
49 Mottolese, M. *et al*. (2005) Phenotypic changes of p53, HER2, and FAS system in multiple normal tissues surrounding breast cancer. *J. Cell. Physiol.* 204, 106–112
50 Roh, H.J. *et al*. (2000) Visualization of the timing of gene amplification during multistep head and neck tumorigenesis. *Cancer Res.* 60, 6496–6502
51 Bastian, B.C. *et al*. (2000) Gene amplifications characterize acral melanoma and permit the detection of occult tumor cells in the surrounding skin. *Cancer Res.* 60, 1968–1973
52 Al-Kuraya, K. *et al*. (2004) Prognostic relevance of gene amplifications and coamplifications in breast cancer. *Cancer Res.* 64, 8534–8540
53 Paris, P.L. *et al*. (2004) Whole genome scanning identifies genotypes associated with recurrence and metastasis in prostate tumors. *Hum. Mol. Genet.* 13, 1303–1313
54 Giulotto, E. *et al*. (1987) Hamster cells with increased rates of DNA amplification, a new phenotype. *Cell* 48, 837–845
55 Mondello, C. *et al*. (1995) Gene amplification in Chinese hamster DNA repair deficient mutants. *Mutat. Res.* 346, 61–67
56 Mondello, C. *et al*. (2001) Increased gene amplification in immortal rodent cells deficient for the DNA-dependent protein kinase catalytic subunit. *Cancer Res.* 61, 4520–4525
57 Smith, K.A. *et al*. (1995) Regulation and mechanisms of gene amplification. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 347, 49–56

58 Dini, G. *et al*. (2002) Overexpression of the 18 kDa and 22/24 kDa FGF-2 isoforms results in differential drug resistance and amplification potential. *J. Cell. Physiol.* 193, 64–72
59 Huang, A. *et al*. (1997) Ribonucleotide reductase R2 gene expression and changes in drug sensitivity and genome stability. *Cancer Res.* 57, 4876–4881
60 Zhou, P. *et al*. (1996) Overexpression of cyclin D1 enhances gene amplification. *Cancer Res.* 56, 36–39
61 Peterson, C. *et al*. (2000) Mutations in RAD3, MSH2, and RAD52 affect the rate of gene amplification in the yeast *Saccharomyces cerevisiae*. *Environ. Mol. Mutagen.* 36, 325–334
62 Bassing, C.H. and Alt, F.W. (2004) The cellular response to general and programmed DNA double strand breaks. *DNA Repair (Amst.)* 3, 781–796
63 Blegen, H. *et al*. (2003) DNA amplifications and aneuploidy, high proliferative activity and impaired cell cycle control characterize breast carcinomas with poor prognosis. *Anal. Cell. Pathol.* 25, 103–114
64 Weinstein, I.B. (2002) Cancer. Addiction to oncogenes – the Achilles heal of cancer. *Science* 297, 63–64
65 Mass, R.D. *et al*. (2005) Evaluation of clinical outcomes according to HER2 detection by fluorescence *in situ* hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin. Breast Cancer* 6, 240–246
66 Gorlick, R. *et al*. (1996) Intrinsic and acquired resistance to methotrexate in acute leukemia. *New Engl. J. Med.* 335, 1041–1048
67 Wang, T.L. *et al*. (2004) Digital karyotyping identifies thymidylate synthase amplification as a mechanism of resistance to 5-fluorouracil in metastatic colorectal cancer patients. *Proc. Natl. Acad. Sci. U. S. A.* 101, 3089–3094
68 Gorre, M.E. *et al*. (2001) Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* 293, 876–880
69 Giaccone, G. and Rodriguez, J.A. (2005) EGFR inhibitors: what have we learned from the treatment of lung cancer? *Nat. Clin. Pract. Oncol.* 2, 554–561
70 Kaye, F.J. (2005) A curious link between epidermal growth factor receptor amplification and survival: effect of 'allele dilution' on gefitinib sensitivity? *J. Natl. Cancer Inst.* 97, 621–623
71 Nagase, H. *et al*. (2003) Allele-specific Hras mutations and genetic alterations at tumor susceptibility loci in skin carcinomas from interspecific hybrid mice. *Cancer Res.* 63, 4849–4853
72 Zhang, Z. *et al*. (2001) Wildtype Kras2 can inhibit lung carcinogenesis in mice. *Nat. Genet.* 29, 25–33
73 Pinkel, D. and Albertson, D.G. (2005) Comparative genomic hybridization. *Annu. Rev. Genomics Hum. Genet.* 6, 331–354
74 Pinkel, D. and Albertson, D.G. (2005) Array comparative genomic hybridization and its applications in cancer. *Nat. Genet.* 37(Suppl.), S11–S17
75 Birchall, P.S. *et al*. (1995) Expression patterns of predicted genes from the *C. elegans* genome sequence visualized by FISH in whole organisms. *Nat. Genet.* 11, 314–320
76 Speicher, M.R. and Carter, N.P. (2005) The new cytogenetics: blurring the boundaries with molecular biology. *Nat. Rev. Genet.* 6, 782–792
77 Wang, T.L. *et al*. (2002) Digital karyotyping. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16156–16161
78 Ginzinger, D.G. *et al*. (2000) Measurement of DNA copy number at microsatellite loci using quantitative PCR analysis. *Cancer Res.* 60, 5405–5409
79 Volik, S. *et al*. (2003) End-sequence profiling: sequence-based analysis of aberrant genomes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7696–7701
80 Tlsty, T.D. *et al*. (1989) Differences in the rates of gene amplification in nontumorigenic and tumorigenic cell lines as measured by Luria–Delbruck fluctuation analysis. *Proc. Natl. Acad. Sci. U. S. A.* 86, 9441–9445
81 Smith, K.A. *et al*. (1997) Multiple mechanisms of N-phosphonacetyl-L-aspartate resistance in human cell lines: carbamyl-P synthetase/aspartate transcarbamylase/dihydro-orotase gene amplification is frequent only when chromosome 2 is rearranged. *Proc. Natl. Acad. Sci. U. S. A.* 94, 1816–1821
82 Alt, F.W. *et al*. (1978) Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured murine cells. *J. Biol. Chem.* 253, 1357–1370
83 Pinkel, D. *et al*. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.* 20, 207–211
84 Fridlyand, J. *et al*. (2006) Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer* 6, 96