Minireview

Molecular Assays for the Diagnosis of Minimal Residual Head-and-Neck Cancer: Methods, Reliability, Pitfalls, and Solutions

Vida M. M. van Houten, Maarten P. Tabor, Michiel W. M. van den Brekel, Fedor Denkers, Richard G. A. Wishaupt, J. Alain Kummer, Gordon B. Snow, and Ruud H. Brakenhoff[1]

Section Tumor Biology, Department of Otolaryngology/Head and Neck Surgery [V. M. M. v. H., M. P. T., M. W. M. v. d. B., F. D., R. G. A. W., G. B. S., R. H. B.] and Department of Pathology [J. A. K.], University Hospital Vrije Universiteit, 1007 MB Amsterdam, the Netherlands

Abstract

The prognosis of cancer patients is determined by the radicalness of treatment: residual tumor cells will grow out and develop in manifest local recurrences, regional recurrences, and distant metastases. Classical diagnostic methods such as radiology and histopathology have limited sensitivities, and only by molecular techniques can minimal residual disease be detected. In tissue samples containing the normal tissue counterpart of a tumor, only tumor-specific markers can be exploited, whereas in other samples, tissue-specific markers can be used. At present, there are two main methodologies in use, one based on antigen-antibody interaction and the other based on amplified nucleic acids. The most commonly used nucleic acid markers are mutations or alterations in tumor DNA (tumor-specific markers) or differentially expressed mRNA (tissue-specific markers). Many reports and reviews have been published on the assessment of minimal residual disease by molecular markers, showing either positive or negative clinical correlations. One of the main reasons for these contradictory findings is the technical difficulty in finding the small numbers of tumor cells in the large number of normal cells, which necessitates sensitivities of the assays up to 1 tumor cell in 2×10^7 normal cells. These assays often are complex, demand considerable experience, and usually are laborious. In this review, we will address a number of the technical issues related to molecular assays for tumor cell detection that make use of nucleic acids as markers. Many difficulties in data interpretation are at least in part because of technical details that might have been solved by the incorporation of one or more appropriate controls. We hope that this review clarifies a number of these issues and help clinicians and investigators interested in this field to understand and weigh the contradictory findings in the published studies. This will help move the field forward and facilitate clinical implementation.

Introduction

The success rate of many cancer treatments is largely dependent on the often microscopically undetectable tumor cells (MRD)[2] that remain in the body. Classical diagnostic modalities such as histopathology and radiology are not sensitive enough to detect these small numbers of cells, but recent advances in molecular diagnostic methods based on tissue- or tumor-specific markers are filling this gap. Particularly in head-and-neck cancer, a more sensitive detection of MRD is of importance because residual tumor cells play a crucial role in the relatively high recurrence rates observed in these patients. Local recurrences occur in ~10–30% of the cases with advanced tumors, even with histopathologically tumor-free surgical margins (1). Moreover, although locoregional control has improved during the last decades, 5-year survival rates have not changed accordingly, which can be explained by a higher frequency of distant metastases (15–25% of the patients; Ref. 2). In a recent study, the long-term results of the treatment of head-and-neck cancer at the base of the tongue with surgery and radiotherapy were published with a follow-up period of 10 years, showing that the most common site of treatment failure was the development of distant metastases in 24% of the patients *versus* local recurrence in only 11% (3). As stated earlier, available diagnostic methods are not sensitive enough to detect minimal residual head-and-neck cancer, and molecular analysis might improve the staging of these patients. However, we and others have experienced several problems and pitfalls in developing and optimizing molecular methods for this purpose, as well as in implementing them in laboratory practice.

At present, it is widely accepted that cancer arises as a result of the accumulation of genetic alterations in oncogenes and tumor suppressor genes, followed by clonal evolution. Some of these alterations occur specifically in the genes that play a crucial role in the normal behavior of the cell, but often these changes appear in less crucial sequences and are, therefore, a mere reflection of the genetic instability of the tumors (4). Hence, tumor cells harbor specific clonal genetic changes that can be used as molecular markers for the detection of cancer cells in clinical samples. In addition to these tumor-specific genetic alterations, tissue-specific markers can also be ex-

Received 3/2/00; revised 7/20/00; accepted 7/20/00.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom requests for reprints should be addressed, at Section Tumor Biology, Department of Otolaryngology/Head and Neck Surgery, University Hospital Vrije Universiteit, P.O. Box 7057, 1007 MB Amsterdam, the Netherlands. Phone: 31-20-444-0953; Fax:31-20-444-3688; E-mail: rh.brakenhoff@azvu.nl.

[2] The abbreviations used are: MRD, minimal residual disease; RT-PCR, reverse transcription-PCR; HNSCC, squamous cell carcinoma of the head and neck; MASA, mutant-allele specific amplification; OLA, oligonucleotide ligation assay; LNM, lymph node metastasi(e)s; LOH, loss of heterozygosity; MSI, microsatellite instability; CEA, carcinoembryonic antigen.

ploited, but with two important considerations: (*a*) the marker should still be expressed homogeneously in the tumors derived from the tissue; and (*b*) the marker should preferably not be expressed in the clinical sample of interest. The choice of a particular marker or assay depends on the necessary sensitivity and specificity of the assay, the origin of the clinical sample, and the laboriousness of the assay. The introduction of *in vitro* nucleic acid amplification methods, most notably PCR, which amplifies minute amounts of DNA >1 million-fold, triggered the development of novel sensitive technologies for the detection of tumor cells by molecular markers. When directed toward genetic abnormalities or characteristics specific for tumors, PCR-based methods might be powerful tools to detect low numbers of tumor cells in the presence of an excess of normal cells (5, 6). Consequently, an improvement of the sensitivity of detecting residual or disseminated disease can be obtained. In the first part or this review, we will discuss molecular assays that use DNA markers for tumor cell detection. In the second part, we will discuss highly sensitive RT-PCR assays that use tissue-specifically expressed RNA markers.

DNA Markers

The first molecular progression model was described for colorectal cancer in which the accumulation of genetic alterations had been demonstrated (4). The transitional stages of this model, ranging from normal epithelium via adenoma to carcinoma *in situ* and metastases, are associated with mutations affecting oncogenes (*e.g., K-ras*) and tumor suppressor genes (*e.g., p53*). Moreover, it was shown that at these various stages, cancer cells often display genetic instability, which can be observed at the DNA level as amplifications, deletions, or alterations of DNA repeat sequences, known as microsatellites (7). A variant of this sequential DNA damage model was later described for HNSCC carcinogenesis (8). Because we are exploring MRD in head-and-neck cancer, we will focus on two established DNA markers: *p53* mutations and microsatellite alterations, which can be exploited as tumor markers in HNSCC but also in many other tumor types (9–15). *K-ras* mutations are relatively rare in HNSCC, which prohibits their use in molecular approaches. In other tumor types, *K-ras* mutations are often used as markers.

Tumor Cell Detection Using *p53* Mutations as a Marker

General Aspects. A genetic event that appears to be of significance in the progression of many tumor types, including HNSCC, is loss of the *p53* tumor suppressor gene, through either allelic deletion and/or mutation (16–18). *p53* alterations are an integral part of cancer progression in almost all types of human cancer, and the assumption that they precede the stage of invasive cancer favors their use as marker for cancer staging.

Mutated *p53* has been used successfully as a marker to detect tumor cells in various tissues such as tumor resection margins, lymph nodes, urine, and saliva (9, 10, 12). In recent years, numerous methods have been described for the detection of point mutations in clinical samples, among which are the plaque hybridization assay (9), MASA (19), OLA (20), POINTEXACCT (a modified oligonucleotide ligation assay; Ref. 21), "enriched" RFLP-PCR (22–25), and very recently, digital PCR (26). In a pilot study in 25 head-and-neck cancer patients, molecular diagnosis using *p53* mutations as the marker and the plaque hybridization assay as the measurement method appeared to be superior to histopathology in detecting HNSCC cells in surgical margins (12). This pilot study demonstrated the potential of these molecular approaches. Moreover, in this pilot study the data were shown to be clinically relevant: the plaque hybridization assay with *p53* mutations as the marker correctly predicted local recurrences in patients with positive surgical margins. Therefore, we initiated a clinical trial with 200 head-and-neck cancer patients for the assessment of residual tumor cells, using *p53* mutations as marker for the analysis of tumor cells in resection margins and *E48* gene expression to detect tumor cells in lymph nodes, blood, and bone marrow. In addition, we evaluated the suitability of *p53* mutations as marker and explored a number of DNA assays. As yet, the plaque hybridization assay is still the most reliable, quantitative, and robust method for tumor cell detection using *p53* mutations as marker. The plaque hybridization assay will be discussed in detail as well as a number of alternative techniques.

Mutation Identification. Sequencing of the *p53* gene should in theory be straightforward, but in practice it is rather complex. This statement underlies the large reported differences in the mutational frequencies of *p53*; a considerable number of these differences can be explained by the method used for sequencing. Most researchers use direct DNA cycle sequencing, either with fluorescent chromophores or radioactive labels. In addition, the *p53* GeneChip assay (Affymetrix, Santa Clara, CA; Ref. 18) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (27) are being used for (*p53*) sequencing. For allele discrimination in blood samples, these methods seem equivalent, but for sequencing DNA derived from tumor tissue, the method used becomes more critical. The number of tumor cells often is low in a tumor biopsy, and the DNA of poor quality, which might result in the missing of mutations or the introduction of artificial mutations. In general, both fluorescent and radioactive cycle sequencing are based on Sanger's dideoxynucleotide method. Because tumor samples are always contaminated by stroma, neoplastic areas in sections of the tumor need to be microdissected before DNA is isolated. Subsequently, the appropriate fragment, usually the domain encoded by exons 5–9, of the *p53* gene is amplified. A large fragment of 1.8 kb containing exons 5–9 or the separate exons can be amplified. We and others have tested both solid-phase fluorescent cycle sequencing and radioactive cycle sequencing for their effectiveness in detecting mutations in the *p53* gene (ABI 373 sequencing technology and dye-labeled deoxynucleotides; Refs. 28, 29). We missed approximately half of the mutations detected by radioactive cycle sequencing (data not shown). A comparable observation (30% missed) was described by Ahrendt *et al.* (18) in a very large study. Obviously, a large number of remarks could be made about the various sequencing strategies, which are beyond the scope of this review (for more details, see Ref. 18), but one particular observation is worth mentioning. Even the amplified template used might considerably influence the sequencing result (Fig. 1). When a small amplicon was used as template (a single exon), a mutation could not be demonstrated by radioactive cycle sequencing, whereas when a large amplicon was used as template (exons 5–9), a

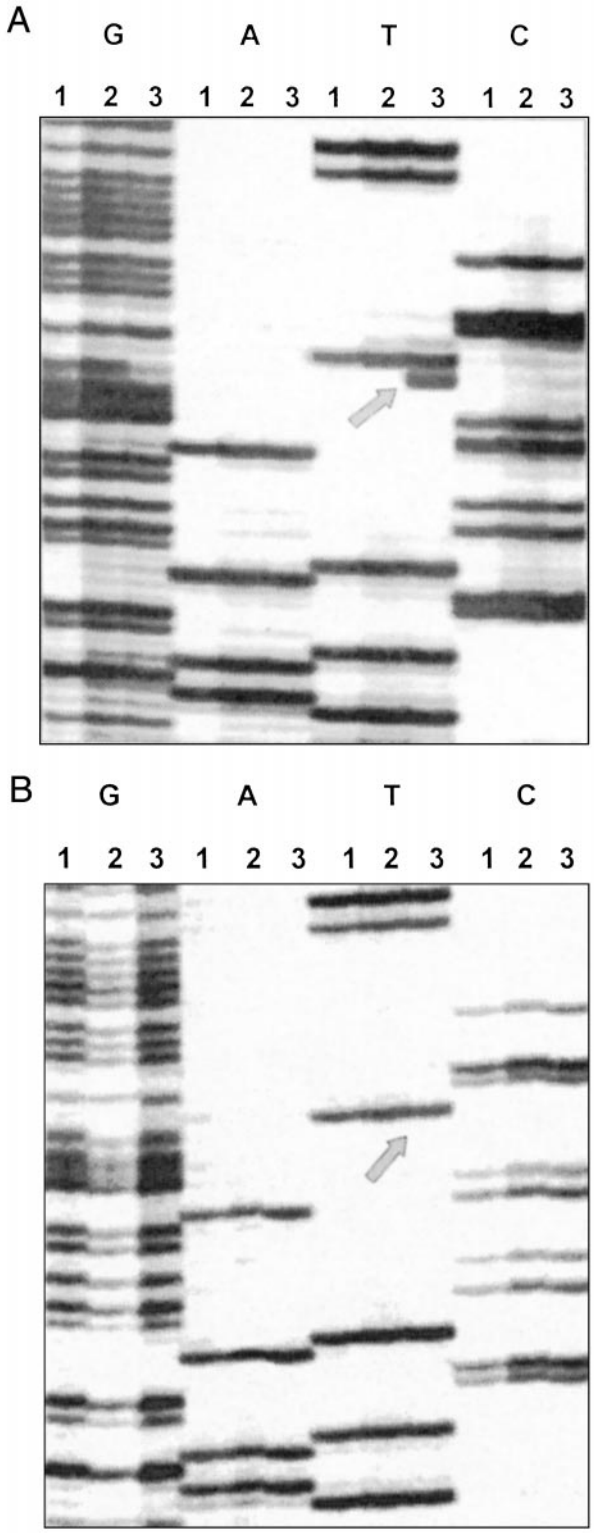


Fig. 1 Sequence analysis for *p53* mutations using different templates. Exons 5–9 (*A*) or exon 6 (*B*) were amplified on DNA from tumor 97-53 as template. The PCR products were purified and sequenced with sense primers. The experiment was repeated once. Both templates were mutation-positive in the plaque assay. Carryover contamination was excluded by appropriate negative controls. The visible mutated base when sequenced on the exon 5–9 template is indicated by an *arrow*.

mutation could clearly be visualized. The experiment was repeated several times to exclude carryover contamination or other experimental confounders, but the result was always the same. The mutation was correctly identified on the large template, as was confirmed by the plaque hybridization assay (see below).

In summary, none of the available methods for mutation sequencing in tumor DNA is 100% reliable, and it is therefore important to be critical about sequencing data published. In fact, to prevent ambiguous results, each mutation should be checked formally by an independent method such as differential hybridization or OLA. When extremely high or low mutational frequencies are reported, the data should be interpreted with caution, in particular when independent confirmation was not included.

Suitability of *p53* Mutations as a Molecular Marker. There are a number of criteria determining whether a molecular marker is reliable (30): (*a*) the molecular marker should be specific for tumor cells such that it correctly distinguishes between normal cells and tumor cells; (*b*) to qualify as a clonal marker, a genetic alteration should precede or occur at the stage of invasive cancer and be preserved during tumor progression and metastasis; and (*c*) the marker has to be broadly applicable, *i.e.*, the marker must be present in a large part of the study population. It has been firmly established by numerous studies that mutations in the *p53* gene meets the first criterion, including HNSCC. The position in the genetic progression model of HNSCC supports its suitability as molecular tumor marker (8, 17, 31).

The suitability of *p53* mutations as a clonal marker (usually determined by mutation screening of lymph node metastases) is being debated, and literature on this issue yields conflicting data. In a number of studies, the clonal stability of *p53* mutations in HNSCC tumor progression was confirmed (32, 33), whereas in other studies, it could not be demonstrated (34–36); reported values ranged from 100% concordance (32) to a mere 25% concordance (36). We compared known *p53* mutations of 14 HNSCC tumors, detected by direct radioactive cycle sequencing and confirmed by plaque assay, with the *p53* mutations of 18 matched LNM. In all cases, concordant mutations were found between the tumor and the LNM. In addition, DNA of LNM of four HNSCC patients whose primary tumors lacked a *p53* mutation were analyzed, and no additional mutations were found in these lymph nodes.[3] These findings support the idea that *p53* is closely associated with tumor progression: *p53* mutations develop before LNM and are maintained during clonal outgrowth.

Intriguingly, a recent report by Offner *et al.* (37) questioned the value of mutated *p53* as a target for diagnosis and treatment; they found that *p53* gene mutations are not required for early dissemination of various tumors. The authors compared the *p53* status of primary and cultured micrometastatic tumor cells obtained from bone marrow aspirates of 46 patients with epithelial tumors and concluded that disseminated tumor cells do not necessarily harbor *p53* mutations. However, it

[3] Manuscript in preparation.

Table 1 Mutational frequency of the coding exons of the *p53* gene in HNSCC

Exons 5–9 were sequenced for mutations by cycle sequencing using a ^{33}P-labeled primer. Tumor DNA that did not contain a mutation in exons 5–9 was subjected to sequencing of the other exons. Not all exons could be successfully amplified in all cases.

	Exon									
	2	3	4	5	6	7	8	9	10	11
Mutation	0/49	1/49	2/50	6/55	6/55	6/55	8/55	4/55	0/42	0/51
Frequency (%)		2	4	11	11	11	15	6		

should be noted that these conclusions were based on 10 SV40-transformed cell lines derived from disseminated carcinoma cells and not on manifest metastases. Moreover, no data on HNSCC were available, and the various tumor types might differ considerably. Larger studies combining data from disseminated cells and manifest metastases might elucidate this issue further.

At present, the major limitation of *p53* mutation-based detection of rare tumor cells in HNSCC patients is that *p53* mutations are present in only 50–60% of the head-and-neck cancers. Its applicability is therefore limited. Recently Kropveld *et al.* (38) described a mutational frequency of almost 100% in HNSCC. However, the same authors reported a mere 25% concordance between mutations in the primary tumor and corresponding LNM (36), which as shown by us and others is not correct. It cannot be excluded that their data suffered from sequencing artifacts. To increase the frequency of mutations in our patient group, we analyzed exons 2, 3, 4, 10, and 11 of the *p53* gene in 21 HNSCC tumors that had no mutations in exons 5–9 by radioactive cycle sequencing and found two additional mutations in exon 4 and one in exon 3, leading to a an additional mutational frequency of 6% in the total study population. These mutations were confirmed by plaque assay (Table 1).

A second drawback of the use of *p53* mutations as a marker is the heterogeneity of mutations, which necessitates sequencing of all individual tumors but also makes it very difficult to find simple and sensitive assays for tumor cell detection (see below). On the other hand, the heterogeneity reduces technical artifacts such as carryover contamination.

In summary, *p53* mutations appear to be suitable tumor-specific markers that appear to occur before the formation of LNM. The major drawbacks seems the frequency, which is ~50% in HNSCC, and the heterogeneity of the mutations, which makes it more difficult to design simple assays for tumor cell detection.

Tumor Cell Detection by the Plaque Hybridization Assay. During the last decades, numerous elegant methods of tumor cell detection have been developed. It is not possible to review all of the various assays in detail and to indicate the pros and cons extensively. Important considerations of a suitable assay include (*a*) sensitivity and specificity; (*b*) robustness, in particular related to the heterogeneity of mutations in the *p53* gene; (*c*) reliability; and (*d*) labor-intensiveness. Depending on the (experimental) question, quantitative aspects might also play a crucial role. For example, it cannot be excluded that in MRD monitoring the risk for recurrence or metastasis is related to the number of tumor cells in the clinical sample. Only after well-performed clinical studies using quantitative assays can this question be addressed; therefore, in the interim, quantitative assays are used.

Because *K-ras* mutations involve only a few codons, there are numerous optimized, quantitative assays available to detect tumor cells using these mutations as marker. However, for *p53* mutations, the mutational spectrum is very heterogeneous, which makes it difficult to find assays that use this marker that are sensitive, specific, robust, and quantitative. For these reasons, we have chosen the plaque hybridization assay as the gold standard. Our method has been modified slightly from the technique reported by Sidransky *et al.* (9). In short, based on the *p53* gene sequence of the tumor DNA, a mutant-specific and a corresponding wild-type-specific oligonucleotide are selected. As a rule of thumb, 17-mer oligonucleotides are selected centrally across the mutation on the target strand, an important condition being that they do not contain a "dC" nucleotide at the 5′ end because these are very difficult to label by polynucleotide kinase (39). In theory, there might be locations that are more favorable thermodynamically for the mutated base, resulting in a larger difference in melting temperature between the two oligonucleotides, but in practice the position around the center usually fulfills the requirements of discrimination. Using DNA from the specimen of interest as template, the *p53* gene (exons 5–9 or specific exons) is amplified by PCR and cloned into lambda phages. These are infected into host bacteria, and the clones (plaques) are transferred to nylon membranes. The plaques are analyzed by differential hybridization with the radioactively labeled tumor-specific and wild type-specific oligonucleotides as probes. Obviously, appropriate positive (primary tumor DNA) and negative (wild-type DNA) controls are included. The hybridization solution initially described was rather complex (9), but 6× SSC, including 0.1% salmon sperm DNA (included only during prehybridization) and 5× Denhardt's also work well (40). Hybridization takes place for 18 h at 11°C below the melting temperature of (one of) the probes, followed by three initial wash steps with 6× SSC at the hybridization temperature and one final stringent wash step at 1–5°C below the melting temperature.

Because each plaque contains identical phages with only wild-type or mutant DNA strands, the number of plaques hybridizing with the mutant oligonucleotide divided by the number of plaques hybridizing with the wild-type oligonucleotide is a reliable measure of the tumor cell DNA load in the original sample. At this point, we have added the confirmation of mutant-positive plaques to identify false-positive signals. When the number of hybridizing plaques is low (between 1 and 5), the identity is confirmed by classical rescreening; the positive plaque is stabbed from the agar, replated, and rescreened.

In summary, the plaque assay is quantitative, well-controlled, highly reproducible, and very robust. Notwithstanding the heterogeneity of *p53* mutations, the assay has never failed in our hands although some particular mutations are more difficult to discriminate from wild type than others. The major drawback is the laboriousness of the assay, and it is suitable only in the experimental setting. Implementation of this assay in a regular clinical setting will not be possible.

Alternative Methods for Tumor Cell Detection Using Point Mutations as Marker. As mentioned before, the plaque assay is among numerous methods that are based on point-mutation detection in (clinical) samples, some of which have only very recently been developed and would be more rapid and less laborious alternatives, including MASA, OLA, POINT-EXACCT, rolling-circle amplification, denaturing-HPLC, RFLP-PCR, restriction endonuclease-mediated selection-PCR, PCR-single strand conformation polymorphism, and digital PCR. A number of these methods have been explored for the analysis of heterozygotes based on point mutations, but they lack sensitivity and specificity for use in MRD detection. In particular, techniques based on the different behaviors of mismatched duplex DNA molecules in various (electrophoretic) separation systems such as denaturing-HPLC, PCR-single strand conformation polymorphism, and others are usually too insensitive to detect mutant DNA strands in a large excess of wild-type DNA strands.

The MASA PCR-method was first outlined by Takeda *et al.* (19) for the detection of (known) *K-ras* mutations in sputum of lung cancer patients. The method is based on efficient amplification when perfectly matched primers are used and inefficient amplification when mismatched primers are used. In various reports, a sensitivity of at least 1:1,000 has been reported, but in our hands this technique did not exceed 1:10 when we used particular *p53* mutations as marker. Although there are various ways to improve the sensitivity of the assay (41), the major drawback is that a particular assay for a particular mutation necessitates considerable optimization. Moreover, the sensitivity of this assay is related to the particular mutation, a very unwanted situation. Finally, the assay is not easy to quantify, although real-time PCR approaches might solve that issue.

The OLA was initially described by Landegren *et al.* (20). The principle of this test is based on the specificity of DNA ligases: ligation when a perfect match is present and no ligation with mismatches. We tested a more extensive variant of this method, POINT-EXACCT, which was described by Somers *et al.* (21) for the detection of (known) *K-ras* mutations. To date, for the detection of *p53* mutations we have encountered a number of difficulties in reaching optimal sensitivity without high backgrounds. However, this method seems promising, and we are still working on it, but again the heterogeneity of *p53* mutations might seriously hamper exploitation.

Very promising approaches are the "enriched" PCR methods such as RFLP-PCR or restriction endonuclease-mediated selection-PCR. This approach was originally described by Kahn *et al.* (22) and Levi *et al.* (23) for K-ras mutation detection and was modified for use with *p53* mutations (24). The method makes use of slightly adapted primers for amplification (usually in two steps). The result is that the wild-type sequence generates a cleavage site for a specific endonuclease. When alternative PCR steps and endonuclease digestion steps are used, the mutant strands are enriched and can be visualized on an electrophoretic gel. Drawbacks of this technique include (*a*) the dependence on a mutation in a "near" restriction site, limiting the use for heterogeneous mutations such as in the *p53* gene; (*b*) difficulties in quantification; and (*c*) a very high risk for carryover contamination (see also below). The latter problem was in part solved by the use of thermostable endonucleases, which are active during the PCR reaction (24, 25). On the other hand, these assays are very sensitive and rapid, and suitable for implementation in the clinical setting.

Most recently, a novel method, named digital PCR, has been introduced as an approach to identifying predefined mutations in a minor fraction of a cell population (26). The strategy involves the isolation of single molecules by limiting dilution and separate amplification of the individual template molecules so that the resultant PCR product is completely wild type or mutant. In fact, this strategy is comparable to the cloning of amplified fragments in phages used in the plaque assay, with the major difference that "cloning" is performed by limiting dilution before amplification. Although in theory a sensitivity of 1:1000 could be reached, digital PCR seems mainly promising for a limited number of mutations (such as for *K-ras*) because for each mutation the molecular probes have to be optimized. Moreover, for a required MRD detection level of 1:5,000, at least 10,000 separate PCR reactions have to be carried out, which requires extensive automation and makes the assays expensive. To date, an "ultimate" assay that can exploit a large variable number of point mutations as marker and is less laborious than the plaque assay has not been described. However, when large prognostic studies might indicate the benefit of molecular staging, improved techniques need to be developed because the plaque assay cannot be implemented in the clinical setting.

Problems and Pitfalls Using Point Mutations as Marker

Sample Contamination. Assays using point mutations as marker need stringent control measures to prevent false-positive signals. Contamination of patient material might take place at different time points during sampling and technical processing. In our study, for example, tumor cell or tumor cell DNA contamination can occur (*a*) in the operating theater, (*b*) during histopathological processing of the samples, and (*c*) as a result of DNA amplimer contamination in the PCR. We take several precautions in every step to prevent contamination. For example, before the resection margins are sampled in the operation theater, the operating field is rinsed extensively and the instruments are changed. Similarly, the pathologist's working area is cleaned before the lymph nodes are dissected from the surgical specimen, and the dissection instruments are decontaminated with 0.1 M HCl and PBS between lymph node preparations. In addition to these general sampling measurements, any sample (resection margin, lymph node) is taken in duplicate to enable a routine histopathological assessment of margin invasion and dysplasia. In the initial pilot study performed by Brennan *et al.* (12), DNA was isolated from cryosections. However, to exclude putative contamination via the cryomicrotome, we decided to sample duplicate margins as mentioned above, and used one for histological control and one for DNA isolation. A second source of contamination can occur during the process

of PCR amplification. More remarks on PCR contamination by amplimers are provided in the section dealing with RNA as marker in which much more sensitive assays have been described.

***Taq* Errors.** A false-positive result in the plaque assay or any other assays using point mutations as marker could also be caused by the enzyme used to amplify the DNA: a thermostable DNA polymerase. It can be anticipated that the enzyme used (in our case *Taq* polymerase) produces spontaneous single-base substitution errors at reported frequencies of 1 in 9,000 nucleotides polymerized (42, 43). This knowledge underlines the need to assess the degree to which random mutations introduced by the *Taq* polymerase during the amplification process play a role in this type of analysis. In particular, when archival stained and fixed DNA is analyzed for point mutations as marker, the error frequencies increase considerably. Both *Taq* errors and DNA sequence damage are potential limitations of the use of point mutations as markers, but these are controllable by an adapted experimental set-up, additional measurements, and statistical analysis.[4]

Sensitivity. The sensitivity of the *p53*-based plaque assay is determined by the amount of input DNA in the PCR amplification, the number of plaques screened, and in some applications, by the *Taq* error rate. The clinically relevant level of sensitivity is unknown. At present, we screen ~5,000 plaques with p53 inserts, corresponding to 5,000 cells. However, although the plaque assay is limited in comparison with other assays such as RT-PCR assays, a single tumor cell in 5,000 normal cells is detectable. The level of detection might be a crucial issue in MRD detection because it could determine the clinical significance of the molecular findings. The significant level of detection can be determined only if quantitative analyses in large studies have been completed and combined with (long-term) follow-up data. Ideally, to gain insight into the clinically required sensitivity of MRD detection, margin assessment should be performed in patients who will not receive postoperative radiotherapy. The recurrence rate in these patients is then a proper reflection of the presence of MRD postoperatively, assuming that MRD always develops into a clinically manifest recurrence.

Although the *p53*-based plaque assay seems a reliable, sensitive, and specific technique, it is very laborious. In addition, it has some other major limitations, one of which is that it cannot be generally applied to all patients; *e.g.*, 50% of HNSCC patients do not show a p53 mutation in their primary tumors. However, as more molecular markers become available, both for head-and-neck cancer and for other tumor types, this limitation might be overcome. Improvements in the technology will be crucial to enable implementation of these molecular techniques in the clinic. The power of the molecular approach can be demonstrated by the following example. A female patient presented with a T_3N_1 tonsil carcinoma. The tumor was treated by surgery and postoperative radiotherapy. The resection margins were tumor-free as assessed by histopathology. Molecular analysis of the HNSCC primary tumor and resection margins was performed for this patient. Sequencing of the primary tumor revealed a p53 mutation. DNA isolated from the margins was screened for mutated DNA, and the deep margin was positive (Fig. 2). The patient developed a local recurrence after 19 months.

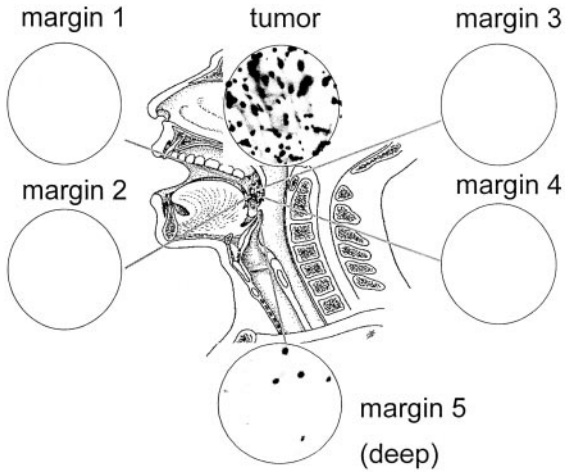


Fig. 2 Typical example of MRD detection by the plaque assay using *p53* as marker. The mutation in the *p53* gene was a nucleotide substitution located in exon 8, codon 272. DNA was isolated from resection margins, and exon 8 was amplified by PCR and cloned in lambda vectors. After host infection, the plaques were differentially hybridized with wild-type and mutant oligonucleotides as probes. The positive plaques were stabbed, replated and rescreened. Tumor DNA was used as positive control (*tumor*), and DNA from cell line UM-SCC-22A was used as negative control (not shown).

Tumor Cell Detection Using Microsatellites as Marker

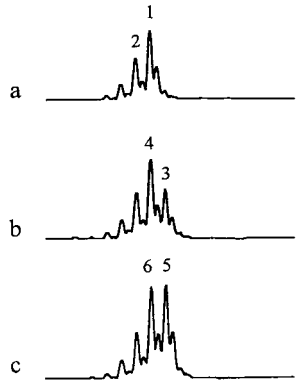
Microsatellites and LOH Analysis. In addition to the specific changes in tumor suppressor genes and oncogenes, cancer cells are genetically unstable and display extensive chromosomal changes, including amplifications, duplications, deletions, and translocations. Some of these changes can be determined by allele-specific markers such as microsatellites because these markers allow distinction of maternal and paternal alleles. Microsatellites are small repetitive sequences that often are highly polymorphic in the population. By PCR amplification and subsequent electrophoretic separation, the maternal and paternal alleles can be distinguished, at least when the number of repeats differs between the two alleles (the marker is then called "informative"). Originally, these markers were exploited for genetic analysis of tumor DNA. The loss of a specific marker in tumor DNA usually is considered as the hallmark of the loss of a specific tumor suppressor gene. Loss of the locus is thought to be the second inactivation step in the complete loss of a tumor suppressor gene, one allele inactivated by mutation and one allele inactivated by loss (44). In practice, tumor DNA is compared with normal DNA (usually isolated from blood lymphocytes) for many different microsatellite markers. The loss of a particular allele in a clinical sample, called LOH, results from allelic deletion, duplication, or amplification. The genetic instability of tumors can also be reflected in changes in the length of the microsatellite (shifts), indicative of MSI. MSI is character-

[4] Nieuwenhuis *et al.*, manuscript in preparation.

ized as a tumor-specific change of length in a microsatellite, as a result of either insertion or deletion of repeating units, when compared with matching normal DNA (45). MSI has been described in numerous human neoplasms, but it is particularly common in most hereditary nonpolyposis colorectal carcinomas and in a proportion of nonhereditary colorectal tumors (46, 47).

Microsatellite analysis is a relatively easy method for genetic research of tumor DNA, but in practice there are some pitfalls: (*a*) The selection of microsatellite markers is crucial: some markers show a high frequency of LOH in a specific tumor type, and other markers are particularly unstable in human tumors and show a high frequency of MSI. (*b*) For tumor DNA analysis, it is necessary to microdissect the tumor tissue accurately because the presence of wild-type DNA will overshadow the loss of an allele. Manual methods of microdissection are laborious and require considerable dexterity, but the introduction and development of laser capture microdissection (48) has simplified this painstaking process. (*c*) The visualization of the PCR-amplified marker can be achieved by either labeling one of the PCR primers or, alternatively, incorporation of labeled nucleotides. The latter method, incorporation of labeled nucleotides, gives inferior results because both strands are labeled, which makes electrophoretic resolution and data interpretation much more difficult. In addition to these differences in labeling method, two different types of label can also be used: radioactive isotopes such as ^{32}P followed by autoradiography, or fluorescent dyes with subsequent analysis on an automated DNA sequencer (Applied Biosystems, Perkin-Elmer). (*d*) Standardization of input DNA is important to prevent artificial differences between the alleles. We use a standard 10 ng of input DNA, which corresponds to ~1,500 cells. These low amounts of DNA can be measured by fluorescence (Hoefer Dynaquant; Amersham/Pharmacia Benelux NV, Roosendaal, The Netherlands). (*e*) A difficult problem when using mono-, di-, and trinucleotide microsatellite markers for either LOH or tumor cell detection (see below) is stutter. Stutter of microsatellite markers is the presence of smaller fragments caused by polymerase slippage and 3′ nontemplate extension. In polymerase slippage, the fragments become smaller in steps of repeat length, thus a 120-bp CA repeat might show bands of 120, 118, and 116 bp, and so on. The polymorphic difference between microsatellites also consists of units of repeat length and thus might comigrate with the stutter bands. In particular, when the second allele is only one repeat unit smaller than the original band, stutter correction is needed to calculate LOH (see below). To enable stutter correction, quantitation of high-resolution gels is needed and the introduction of denaturing agents such as formamide is necessary (49). A very simple solution to solve stutter could be the use of tri- or tetranucleotide repeats, which stutter considerably less. However, these repeats are in general less widely dispersed in the genome and often are less polymorphic.

LOH is determined by comparing the allele ratio in tumor DNA to the allele ratio in normal DNA. LOH scoring is performed by visual inspection and calculation. Obviously, the latter method, calculation, is the most objective. In many cases, LOH can be scored easily by visual inspection, in particular when the tumor was homogeneous and accurately microdissected. In addition to scoring by visual inspection, LOH can be calculated, which is especially necessary when highly stuttering markers and alleles that differ only one repeat unit are used. Radioactive DNA fragments can be quantified accurately on a phosphorimager. Nonradioactive analysis always reveals quantitative information when the raw data are analyzed by an appropriate software program (GeneScan; ABI/Perkin-Elmer). Usually, LOH is calculated as the ratio between short allele-normal (Sn)/long allele-normal (Ln) and short allele-tumor (St)/long allele-tumor (Lt), in the formula: (Sn:Ln)/(St:Lt). Normally, we score an LOH when 50% of the allele is lost in the tumor (score <0.5 or >2), but other researchers use less strict criteria. Borderline data are resolved by more precise microdissection. Using this quantitation, we can also correct for stutter. For a particular marker, the relative contribution to the stutter bands is calculated from a noninformative sample and is used to calculate the relative abundance of the second allele to the first stutter band of the first allele (Fig. 3). By this method much more information can be deduced with larger panels of markers.



Peak	Size (bp)	Peak Height
1	162	1163
2	160	694
3	164	819
4	162	1336
5	164	1543
6	162	1525

Fig. 3 LOH analysis by microsatellite marker CHRNB1, using stutter-correction. *a*, example of a noninformative control sample. *Peak 1*, the two identical alleles of the sample that comigrate because they are of the same size. *Peak 2*, a so-called stutter band caused by Taq polymerase slippage (PCR artifact). The peak-height ratio of peaks 2 and 1 is 0.60 (*table* on the *right*). These stutter band ratios are highly reproducible for a given microsatellite. *b*, normal DNA of an HNSCC patient. The marker is informative for this patient and shows two alleles. The alleles differ in size by only 2 bp. *Peak 3*, the largest allele. *Peak 4* consists of the (2-bp) small allele plus the first stutter band of the large allele. To obtain the real peak height of the small allele, the contribution of the stutter band of the large allele should be subtracted: small allele = peak $4 - 0.60 \times$ peak 3. *c*, tumor DNA of the patient in *b*. LOH calculation without stutter correction is (for formula see text): (Sn:Ln)/(St:Lt) = (1336:819)/(1525:1543) = 1.65. Normally LOH is defined as a >50% loss of one allele in the tumor (score <0.5 or >2.0). In this case, the ratio is <2.0; thus the tumor sample would not be scored as LOH. However, with stutter correction, the ratio is: (Sn*:Ln)/(St*:Lt) = $[(1336 - 0.60 \times 819):819]/[1525 - 0.60 \times 1543):1543] = 2.66$. By this calculation, the tumor sample would have been scored as LOH at this locus.

Microsatellites as Marker for Tumor Cell Detection. On the basis of the strict criteria for LOH analysis in tumors and the problems with stutter, it can be concluded that LOH (imbalance in allele-specific microsatellites) is an insensitive DNA marker for tumor cell detection. The allele ratio of normal DNA

varies from 0 to 10% in different experiments;[5] thus, a clinical sample should at least contain >20% tumor cells to obtain a reliable signal. Nevertheless, in particular cases, LOH analysis was shown to reveal relevant information when serum DNA or urine sediments were analyzed (13, 14, 50, 51). As stated above, in a number of cases new microsatellite alleles arise as a result of genetic instability of tumors (MSI). These novel microsatellite alleles might serve as markers for the detection of MRD. Indeed, MSI at a given marker can be detected at a ratio of tumor to normal DNA of 1:500 (11). However, MSI data also should be interpreted with caution because artificial bands can occur when archival material of low DNA concentration is used. Again, appropriate quantitation of the amount of input DNA of the clinical sample is crucial to avoid these artificial PCR products.

In conclusion, microsatellite analysis can be used for cancer detection. However, the application greatly depends on the ratio of tumor cells to normal cells. In bladder cancer, microsatellite analysis of urine has been used successfully to identify cancer cells and to monitor patients for recurrent disease (14, 50). In urine, 50% of the cells appear to consist of tumor cells, but this ratio might be totally different in other clinical samples. For example, the percentage of tumor cells in the resection margins of patients treated for HNSCC is much lower, and LOH analysis would not be sensitive enough for tumor cell detection. Microsatellite alterations are much more useful for the detection of residual disease, but large panels of markers have to be screened on the tumor to find a suitable alteration. The number of head and neck tumors that exhibit MSI in one or more markers is not clear at present and ranges between 7 and 30% in different studies (52, 53). Finally, the detection of tumor cells using microsatellites as marker remains less sensitive compared with detection based on mutated *p53*.

RNA as Molecular Marker for Disseminated Tumor Cells

General Aspects. Circulating tumor cells have been noticed by clinicians and investigators since 1869 (54). However, the number of circulating tumor cells is in general low, and only since the development of techniques with a high sensitivity has this field obtained increasingly more attention. A major breakthrough was immunocytology, which was exploited in large numbers of studies with high clinical impact (55–57). In a recent study, Braun *et al.* (57) demonstrated the clinical significance of bone marrow assessment for patients with breast cancer. In breast cancer, the presence of axillary lymph node metastases is an important prognostic indicator. However, even in node-negative patients, up to 30% develop distant metastases, whereas 40% of node-positive patients survive 10 years or more. The presence of tumor cells in the bone marrow was very significantly correlated to a poor survival and disease-free survival. Particularly for node-negative patients, it was suggested to use bone marrow status as a criterion to select patients for chemotherapy.

[5] Unpublished data.

The laboriousness of immunocytochemistry and the development of the RT-PCR technique triggered the exploration of novel methodologies for rare tumor cell detection. Most RT-PCR methods for the detection of tumor cells from solid tumors make use of differentially expressed transcripts present in the normal tissue and retained in the malignant cells but absent in the clinical sample of interest. A number of critical issues when establishing RT-PCR methodologies at a sensitivity of 1 cell in 10^7 background cells will be discussed below. Most of these methodological and technical aspects are not discussed in the many reviews on MRD, but they are crucial for the evaluation of studies or judging the literature when the data are published. It is remarkable that for RT-PCR assays, a number of cornerstones for "Good Scientific Practice," such as appropriate controls, sensitivity, specificity, reproducibility, and quantitation, often are neglected. To our belief the, uncontrolled variations in methodology might explain a large number of the contradictory findings in the literature.

Sensitivity. Most RT-PCR data on MRD are collected from the analysis of bone marrow and blood. An increasing number of RT-PCR assays focus on MRD detection in lymph nodes or other samples, but it should be noted that assays for blood or bone marrow often fail in lymph node samples because the markers suitable for tumor cell detection in blood and bone marrow are often expressed in lymph nodes of noncancer controls (see also below). It should further be realized that the frequencies of tumor cells in bone marrow and blood are very low, ranging from 1 in 10^5 (breast) to 1 in 10^7 (head and neck). This means that the techniques exploited have to approach very high levels of sensitivity. Using the E48 antigen as marker, we were able to detect reproducibly a single cell in a 7-ml tube of blood, corresponding to $2–7 \times 10^7$ WBCs, but it demanded a large investment in time to get all variables controlled. In this assay, we use hypo-osmotic lysis of RBCs, reverse transcription with a specific antisense primer, no primers other than the two PCR primers, quadruplicate assays (see also below), an antigen that is abundantly expressed on the target tissue (squamous cells) and not illegitimately expressed in blood or bone marrow cells (58), and Southern blotting using a radiolabeled cDNA as probe. Without preselection techniques (positive or negative immunoselection), we doubt whether higher sensitivities are achievable.

Intriguingly, when we used the same RT-PCR on lymph nodes, we noticed an "unwanted" positivity in samples from non-cancer controls (it was not a "false" positivity). Initially, this was considered as a deterrent to developing the assay for lymph nodes further, but recent data of others have led to the realization that it is not only the "technical" level of sensitivity that determines whether an assay leads to relevant data. Liefers *et al.* (59) described an RT-PCR assay for the assessment of MRD in lymph nodes of colon cancer patients using CEA as marker. CEA is, however, expressed in almost any tissue but usually at a low level. The authors therefore adapted the sensitivity of their assay to the level where control material was negative (to increase the specificity) and could still find clinically relevant data. This study was criticized, nevertheless, because the number of control samples tested was only seven, which is indeed too low, but they clearly showed that the suitability of a marker is not determined by the *maximum*

reachable level of sensitivity but merely by the *relevant* level of sensitivity. The *maximum reachable* level of sensitivity is determined by the expression of the marker and the technology chosen. With respect to the latter, it is obvious that RT-PCR assays using Southern blotting and hybridization are much more sensitive than those using agarose gel detection.

In contrast, the *relevant* level of sensitivity is determined by (*a*) the amount of material available for the assay, (*b*) the specificity of the assay, and (*c*) the threshold of clinically relevant data. For example, from particular lymph node aspirates we isolate only 0.1 μg of RNA, corresponding to $\sim 10^5$ cells. The *relevant* level of sensitivity is thus only 1 in 100,000. In addition, when a tumor load of at least 1 tumor cell per 1,000 cells is shown to be clinically relevant, "false" positivity at a level corresponding to 1 tumor cell-equivalent per 10,000 cells in noncancer controls will not negatively influence the suitability of the assay, assuming that the assay is performed quantitatively. Moreover, a sensitivity higher than 1 cell per 10,000 cells is not useful.

With respect to sensitivity, there are a few noteworthy aspects of the *E48* RT-PCR assay we have developed. Initially, we showed that we could reproducibly detect a single tumor cell added to 7 ml of blood (on average, $2–7 \times 10^7$ nucleated cells), but we needed to perform the assay in quadruplicate (58). From 7 ml of blood we isolated, on average, 26 μg of total RNA ($\sim$56 μg from 2 ml of bone marrow, the mean value of 29 volunteers). HNSCC cell line UM-SCC-22A, used for validation of the assay, was shown to contain 17 μg of total RNA per 10^6 cells, corresponding to 17 pg/cell. In a serial dilution, we are able to detect 0.5 pg of total RNA, corresponding to 1/30 of a cell. Usually, we add 5 μg of blood or bone marrow RNA to an RT-PCR assay, thus measuring 20% of the sample (5/26 μg). The excess of blood cell RNA thus decreases the detection limit approximately $(0.2 \times 17)/0.5 = 7$-fold. As indicated, the detection limit is 0.5 pg for UM-SCC-22A RNA. When we assume that the amount of mRNA is 5%, the number of *E48* transcripts is 0.1% (CGAP database, NCBI), the mean length of transcripts is 3,000 bp and the molecular weight of a single base is 300 D, then we detect $(0.5 \times 10^{-12} \times 0.001 \times 0.05/3{,}000 \times 300 \times 6 \times 10^{23}) = 12$ transcripts. When this amount of RNA (converted into cDNA) is divided over four tubes, it can easily be explained that reproducibility problems start to occur and that cumulative quadruplicate testing is useful.

The detection level of a single cell per 7 ml of blood cannot be reached with DNA markers. As described above, the amount of RNA isolated from 7 ml of blood is $\sim$26 μg. The WBC count in blood is on average $2–7 \times 10^7$ per 7 ml, which corresponds to 1 pg of total RNA/cell. Our detection limit was 1 HNSCC cell/$2–7 \times 10^7$ WBCs, which corresponds to 17 pg of HNSCC RNA/26 μg of WBC RNA. These calculations show that an increase in sensitivity of at least 10-fold is gained by the differences in RNA content between nucleated blood cells (particularly lymphocytes) and squamous cancer cells. When DNA markers are used, this particular advantage will be lost, and a large amount of DNA must be tested to reach these high levels of sensitivity because 26 μg of genomic DNA corresponds to "only" 0.24×10^7 blood cells. When sensitivities of 1 cell per 10^7 WBCs need to be reached by DNA markers, enrichment techniques such as positive and negative immunoselection become much more important.

As stated above, the maximum achievable level of sensitivity is determined by the expression of the marker and the chosen technology. We use Southern blotting and hybridization with radiolabeled E48 cDNA as the probe for detection. When we explored radiolabeled primers as probes, we experienced a 10-fold loss in sensitivity. Using nested PCR (a second PCR step with primers located within the amplimer), we could not further increase the level of sensitivity. However, the amplimers could be seen on ethidium-stained agarose gels and Southern blot hybridization could be omitted. We did not introduce nested PCR as routine technology for reasons of quantitation and contamination (see below).

Specificity. Notwithstanding the considerations indicated above, sensitivity is not the most crucial item for RT-PCR assays because amplifications from even single DNA molecules have been described (26). The largest problem with these assays usually is specificity, and the most valuable studies on MRD detection in clinical samples using RT-PCR assays always show data of at least 20 noncancer controls, ideally matched for relevant confounding factors.

The specificity of RT-PCR assays is determined by (*a*) carryover contamination, (*b*) illegitimate expression, (*c*) marker expression by non-tumor cells in the sample, and (*d*) the needed *relevant* level of sensitivity. Large numbers of markers are unusable for blood or bone marrow samples because of illegitimate expression: the expression of a single transcript per 100-1000 cells (60). This level of expression is called illegitimate because at this level of expression these antigens are not expressed on the protein level and cannot be relevant for cellular functions. However, when 20×10^6 cells are assayed, there are still 20×10^4 to 20×10^3 transcripts in the sample, far above the general detection limits. In addition to illegitimate expression, samples might be "falsely" positive for a particular marker because a small subset of the normal cells might express the transcript and the antigen. Although it is generally accepted to call this phenomenon from the clinical point of view false positivity, it should be realized that it is technically not false positivity but merely "unwanted" positivity.

One of the most difficult samples to analyze for MRD is blood or bone marrow because the number of tumor cells is extremely low. We have explored a considerable number of markers for HNSCC tumor cell detection in bone marrow and blood. The suitability of a number of these markers on blood and/or bone marrow samples from noncancer patients is listed in Table 2. As can be seen, all of these markers failed. The signal in control samples is too high to enable detection of the few tumor samples present. Interestingly, in a number of cases these markers were claimed to be highly specific (*e.g.*, cytokeratin 19; Ref. 61), which in fact appeared to rely on limitations in the experimental setup (61–63). For some markers, including cytokeratin 19, the development of reliable RT-PCR assays is very difficult because a number of processed pseudogenes are present in the genome (64, 65). These processed pseudogenes are thought to have integrated into the genome during evolution from reverse-transcribed mRNAs and are, therefore, most often highly homologous to the original gene sequence. Because in

Table 2 Illegitimate expression of RT-PCR markers in bone marrow and blood

Suitability of the indicated markers for the detection of circulating tumor cells as assessed on blood and bone marrow samples of noncancer controls. The number of positive samples/total number of samples analyzed is indicated. For all markers, a 40-cycle RT-PCR was performed using the optimized annealing temperature, except for the cytokeratin-19 RT-PCR. In this case the original protocol as reported by Datta *et al.* (61) was used. The amplimers were detected by Southern blot hybridization using the cloned radiolabeled amplimer as a probe. The primers for CEA amplification were taken from Gerhard *et al.* (66), and those for cytokeratin-19 amplification from Datta *et al.* (61). Other primers sets were developed using PCR-plan software. The suitability of particular CD44 variants as squamous cell-specific markers was based on Van Hal *et al.* (73). Many markers appear to be illegitimately expressed; similar data on a number of these markers were reported by others earlier (74).

Marker	Blood	Bone marrow
CD44v6 variants		
Exon 10 sense/11 antisense	3/3	2/2
Exon (9–10)[a] sense/(10–11)[a] antisense	2/2	2/2
Exon (5–7)[a] sense/(10–11)[a] antisense	2/2	2/2
Cytokeratin 19	3/3	NE[b]
CEA	2/2	2/2
Involucrin	5/5	NE
Telomerase (hTERT)	2/2	2/2

[a] Primers spanning the exon borders.
[b] NE, not evaluated.

general some genomic DNA is co-isolated in most, if not all, standard RNA isolation procedures, these pseudogenes are amplified by the *Taq* polymerase, producing a false-positive signal. In this case, the term false-positive is correct because the signal is not derived from RNA transcripts but from the genomic DNA. This problem is easily identified by leaving out the reverse transcriptase step in an RT-PCR assay. If the signal remains, it is most likely the result of amplification of genomic sequences and often points to the presence of processed pseudogenes in the genome.

In addition to illegitimate expression and carryover contamination (see below), the specificity of an RT-PCR assay is determined by marker expression of normal cells in the clinical sample, *i.e.*, the source of the sample and the *relevant* level of sensitivity. It is obvious that the source of the sample plays a crucial role: markers usable in bone marrow samples might be unusable in lymph node samples or resection margins. Moreover, the specificity is related to the sensitivity of the assay: a marker that seems to be highly specific when using agarose gel detection might fail when using Southern blot hybridization, a more sensitive detection method. As stated, the *E48* RT-PCR is highly specific for blood and bone marrow but not for lymph nodes. In fact, when we tried to enhance the sensitivity of the *E48* RT-PCR further using *AmpliTaq* Gold (ABI/Perkin-Elmer), we ended up with "unwanted" positivity in blood and bone marrow samples of noncancer controls (data not shown). It is possible that when the sensitivity of any RT-PCR assay for MRD detection is pushed to the maximum reachable limits, the specificity might in every case drop below 100%.

When thousands of cancer cells are present in patients, the sensitivity of the assay can be limited, and subsequently, specificity is warranted with any marker; however, this obviously is not the case. Because of difficulties in quantitation in RT-PCR assays, we selected the most suitable markers on the basis of absolutely negative results on noncancer control samples. As real-time RT-PCR approaches improve the quantitation of the assays enormously, many more markers that are not completely negative on control samples might be usable. A background value can be determined in noncancer control samples, and a "cutoff" value to call a sample positive can be established. On the other hand, the establishment of a cutoff level might be hampered by the fact that the number of transcripts in each tumor cell may vary from cell to cell or from patient to patient. This would make it difficult to distinguish illegitimate expression in normal cells in the sample from the "true-positive" expression in tumor cells.

Carryover Contamination. The best-defined problem in RT-PCR assays is carryover contamination. The exponential amplification method effectively amplifies a few DNA copies 1 million-fold to amounts of DNA that are easily detectable on agarose gels. Contamination of a few of these amplimers in the following samples will lead to amplification and subsequent false-positive results. There are a number of measurements that are required to prevent carryover contamination. We use (*a*) three different laboratories, one for reagent stock preparation, one for DNA/RNA handling of clinical material and setting up the amplification reactions, and one for PCR amplification and handling of amplimers and cloned material; (*b*) reagent stock preparation on days when no amplimers are handled; (*c*) different pipettes for reagents and DNA/RNA; (*d*) aerosol tips for pipetting DNA/RNA; (*e*) RNA dilutions as positive control (see also below); (*f*) adapted cDNA probes missing one of the primer sites; (*g*) randomly added negative controls; (*h*) RT-PCR-tested water to dissolve the RNA samples; and (*i*) hand washing after every visit to the amplimer laboratory (which obviously also is good laboratory practice). Using wipe-test experiments, we could show that amplimers are present on the hands, pencils, note books, benches, and clothes of the researchers.

Interestingly, and in fact logically, contamination detection is dependent on the sensitivity of the assay. In a period when the sensitivity of our assay had declined 5–10-fold, we performed various experiments to elucidate the basis of this decreased sensitivity. When we discovered that it was because of recalibration of a PCR machine, we could easily reestablish the original assay. However, when the usual high sensitivity was reestablished, all reagents appeared to be contaminated, which had not been detected with the regular controls because it was beyond the sensitivity level at that time. We also made the mistake of preparing a large volume of ready-to-use PCR mixtures in that period, and although they were tested and shown to be contamination free, in retrospect they all appeared to be contaminated. It took a long time to find the contamination source, solve the problem, and prepare and test novel stocks. Obviously, every experienced laboratory has its own nightmares in this respect.

In summary, carryover contamination is a real problem, and reports on RT-PCR assays should preferably contain a sentence in the Materials and Methods section that adequate measurements were taken to prevent carryover contamination, which demonstrates that the authors are aware of this problem.

Reproducibility and Quantitation. A cornerstone of reliable measurement is reproducibility. It is therefore remarkable that in most studies on MRD, reproducibility is not tested. Often the measurements are performed only once. The underlying reason might be that the tumor cell frequency is so low that duplicate analyzes might be negative for statistical reasons (Poisson distribution). Although this might be true for immunocytological assays, it is certainly not true for RT-PCR assays because the cells are lysed and, in fact, a number of target transcripts are analyzed. We discussed earlier that problems with reproducibility might reflect low abundance of target transcripts: low numbers of cells and/or low expression of the target gene. In serial dilutions, it was shown that the signal does not fade away, but instead problems with reproducibility start to occur when the lowest level of detection is reached (58). However, when using quadruplicate assays (a 20-μl reverse transcription reaction divided over four tubes for PCR amplification), any sample to which a single UM-SCC-22A cell is added remains positive in an *E48* RT-PCR assay. Further optimization of the test conditions in fact showed that this on/off effect can be used for quantitation. At present, we use a scoring system of these "irreproducible bands" after visual inspection of a standardized autoradiogram (scores are 0–3 for every lane). To our surprise, these serial analyzes and this scoring system was quite reproducible, which enables its use as calibration curve (Fig. 4). We therefore decided to introduce the "UM-SCC-22A picogram-unit," in which we link the data of a clinical sample to a calibration curve of a serial dilution run in parallel. By subsequent analysis of *E48* expression on the primary tumor, we have an estimate of the actual tumor cell load. Obviously, this quantitation is in fact an estimate because the *E48* expression of the cells in the bone marrow is not known, and it might differ from the original tumor. Moreover, this scoring system does not substitute for duplicate assays, which implies that to accurately assess the tumor cell load in a bone marrow sample in duplicate, two reverse transcription reactions with both quadruplicate PCRs must be carried out!

In summary, instead of trying to solve the problem of reproducibility when reaching the limits of sensitivity, we decided to use it for quantitative estimation. On the one hand, the serial dilution indicates whether a particular experiment has reached the relevant level of sensitivity (at least one positive signal when 0.5 pg of RNA is used), and on the other hand, it serves as calibration curve for the clinical samples. It should be emphasized that this calibration can be used only when the number of tumor cells is 1–5 cells/2 $\times$ 10^7 nucleated blood cells. When the tumor cell load exceeds five cells, the quantitation estimate is very inaccurate. Obviously, the introduction of real-time RT-PCR measurements will help enormously in the quantitation of this type of assay. With this approach, clinical samples are always related to a calibration curve, and reproducibility will be improved. Furthermore, these techniques will allow the use of markers that are not negative for samples from noncancer controls, although for blood and bone marrow the high sensitivity needed will still require a high-specificity marker. The only drawback of the real-time RT-PCR approach is that on the level of only one to a few cells, the technique might not be sufficiently sensitive. At present, we have established a real-time PCR system for the *E48* marker, and indeed,

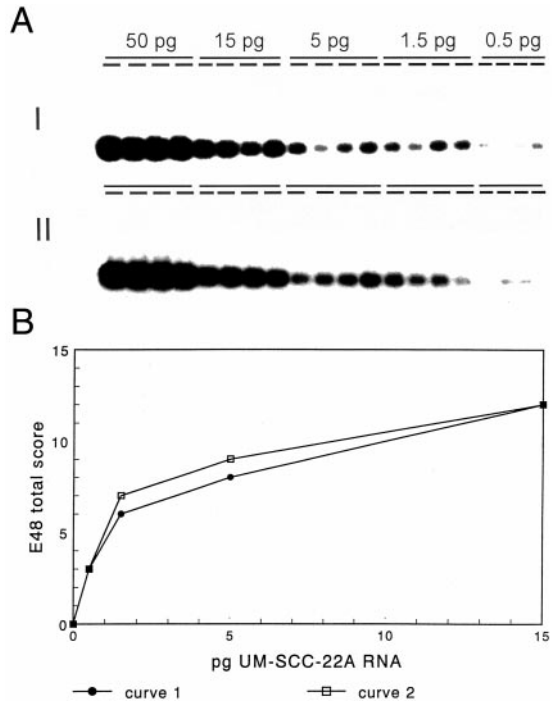


Fig. 4 *E48* RT-PCR reproducibility and calibration curve. *A*, serial dilutions of RNA isolated from HNSCC cell line UM-SCC-22A were subjected to *E48* RT-PCR. Data of duplicate experiments (*I* and *II*) are shown. Autoradiograph exposure was standardized by visual estimation of 50-pg sample intensity. The data were quantified by scoring the intensity of the hybridized fragments (range, 0–3) multiplied by the number of positive PCRs (range, 0–4). The minimum value is 0, and the maximum is 12. Most patient samples are positive in the range 1–12. Note that a single squamous cancer cell contains ~15 pg of total RNA. *B*, graphical representation of the score *versus* the dilution. There is no linear correlation, but the curves of duplicate analyses are unexpectedly comparable.

reproducibility problems occur in the sensitivity range of a few cells, which might be solved by further optimization but most likely reflects the problem of too few transcripts.

Positive Controls. RT-PCR assays, like all other assays, need to be checked by positive and negative controls. We always include negative controls, often including controls without reverse transcriptase, without RNA template, and so forth. In most reports, the results of reconstruction experiments have also been shown, which is crucial to obtain reliable estimates of the sensitivity of a particular assay. However, the routinely used positive controls run in parallel with every assay usually consist of a single analysis of a large amount (100 ng or even 1 μg) of total RNA isolated from a relevant cell line (66–68) or tumor-infiltrated tissue (59). In some cases, positive controls are omitted (69). Obviously, these high amounts of RNA are insufficient as positive control. When RT-PCR reactions are used at a level of sensitivity of only a few cells, positive controls should be on the level of a few cells: 1–10 pg. It is crucial to use a serial dilution up to the level of the relevant sensitivity and calibrated in a reconstruction experiment. For bone marrow and blood samples, these controls would be in the range of a few tumor

cells. Small disturbances in the assay leading to a decrease in the sensitivity can then easily be detected.

Reagents, Protocols, and Interlaboratory Variation. Among the main variables in RT-PCR technology are the differences in protocols and reagents. Most laboratories have invested substantial time to determine the best reagents to enhance reproducibility and sensitivity. Below are listed a few details that unexpectedly appeared to play a role in our studies. More standard variables, such as primers, annealing temperatures, magnesium concentration, and type of enzymes, are not listed. Note that these observations are based on our standard assay. Different assays might lead to other relevant variables.

• The reverse transcription reaction should have a duration of at least 2 h. A longer incubation period is not useful, but incubation for only 1 h resulted in a decrease of signal.

• Addition of other primers to the PCR reaction resulted in a decrease in signal. We therefore perform control RT-PCR reactions to check the quality of the RNA in a separate tube. Oligo-dT or random priming in the reverse transcription reaction results in decreased sensitivity.

• Freezing-thawing of the cDNA reaction results in loss of sensitivity because of absorption to the plastic of the vials. Use of specialized vials with less-adsorbent walls solves the problem in part, but not completely. We therefore complete the RT-PCR amplification in a single day.

• Commercially available kits work well but are still less sensitive.

• To enable detection of a single cell, between 2 and 10 μg of total RNA can be added to the reverse transcription reaction (25% is assayed in the PCR reaction). Higher and lower amounts of RNA result in a decrease in the level of sensitivity.

It is obvious that all of the different protocols from the various laboratories using all kinds of reagents results in a considerable variety in sensitivity of the assays. Because the assays need to be controlled thoroughly, we would suggest the inclusion in every experiment of a serial dilution of RNA from a standard cell line up to the lowest limit of detection. This might be LNCaP for prostatic cancer, MCF7 for breast cancer, UM-SCC-22A and/or UM-SCC-14C for head-and-neck cancer, and COLO357 for colon cancer. If researchers prefer to use their own standard cell lines, they might include a comparison with one of the "standard" cell lines.

Immunocytology *versus* RT-PCR. Although this review in principal deals with RT-PCR technology, most clinically relevant data have been gathered using immunocytology with cytokeratins as markers (55–57). Immunocytochemistry is hampered by the fact that "unwanted" binding to cytokeratins expressed in hematopoietic cells or "nonspecific" binding to Fc-receptor-positive cells lead to false-positive results (70). Intriguingly, the false-positive rates in immunocytological approaches appear to be very dependent on the laboratory, indicating that subtle differences in the methodology might have a large impact on the reliability of the data (56). To correct for false-positive staining, many researchers have introduced additional morphological assessment of positively stained cells, which often is subjective and prone to errors. In fact, a considerable number of the immunocytological studies might have suffered from false-positive results (71). Notwithstanding these limitations in specificity, clinically very relevant correlations have been found (55–57). Reduction of the false-positive rate can only improve the clinical value.

A second drawback of immunocytology and screening by hand is the labor-intensiveness that makes the technology too expensive for routine implementation in health care. However, recent advances in automated imaging techniques could solve this problem for immunocytological analysis. There are also a number of advantages when using immunocytology. Unwanted positivity because of illegitimate expression in hematopoietic cells is not an issue because the low numbers of transcripts per cell do not result in detectable protein expression. A second large advantage is that the cells remain intact and can be characterized further. Possibly, combinations of immunotechnologies and nucleic acid techniques might produce assays combining the best of both: the sensitivity of immunocytology and the specificity of nucleic acid analysis (72).

In the past few years, much progress has been made in our understanding of the molecular progression of primary human cancers. Characterizing the molecular progression of a particular tumor type is critical to providing markers for diagnostic molecular assays. Use of the appropriate marker or panel of markers would then provide information about the presence or absence of (minimal) residual disease, theoretically the best prognosticator of relapse. Point mutations in oncogenes and tumor suppressor genes, microsatellite markers, and tissue-specific transcripts have all been used as markers to detect the presence of small numbers of tumor cells from solid tumors. Recent advances in the field of molecular genetics have provided extremely sensitive RT-PCR methods that enable the sensitive detection of tumor cells through the amplification of tumor-specific or tumor-associated (marker) nucleotide sequences in DNA or RNA. PCR-based techniques are much more sensitive in identifying cancer cells than standard techniques, and at present, we and others have pointed to two molecular markers (p53 point mutations and E48 antigen expression) that seem promising for the detection of MRD in HNSCC patients. However, the precise cutoffs for clinically relevant levels of MRD have not yet been determined and will require large prospective studies in patients. These prospective studies have to show whether MRD as detected by molecular analysis is indeed a cause of locoregional or distant relapse and influences prognosis of the individual patient. Once the prognostic value of MRD has been established, molecular techniques might find acceptance in the clinic for improved staging and more individualized treatment planning for cancer patients. In the meantime, the identification of new molecular markers and the development of new technologies are ongoing processes to provide the tools necessary for widespread implementation of molecular techniques in clinical routine.

References

1. Leemans, C. R., Tiwari, R., Nauta, J. J., van der Waal, I., and Snow, G. B. Recurrence at the primary site in head and neck cancer and the significance of neck lymph node metastases as a prognostic factor. Cancer (Phila.), *73:* 187–190, 1994.

2. Leemans, C. R., Tiwari, R., Nauta, J. J., van der Waal, I., and Snow, G. B. Regional lymph node involvement and its significance in the development of distant metastases in head and neck carcinoma. Cancer (Phila.), *71:* 452–456, 1993.

3. Harrison, L. B., Lee, H. J., Pfister, D. G., Kraus, D. H., White, C., Raben, A., Zelefsky, M. J., Strong, E. W., and Shah, J. P. Long term results of primary radiotherapy with/without neck dissection for squamous cell cancer of the base of tongue. Head Neck, *20:* 668–673, 1998.

4. Fearon, E. R., and Vogelstein, B. A genetic model for colorectal tumorigenesis. Cell, *61:* 759–767, 1990.

5. Cairns, P., and Sidransky, D. Molecular methods for the diagnosis of cancer. Biochim. Biophys. Acta, *1423:* C11–C18, 1999.

6. Sidransky, D. Nucleic acid-based methods for the detection of cancer. Science (Washington DC), *278:* 1054–1059, 1997.

7. Weber, J. L., and May, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet., *44:* 388–396, 1989.

8. Califano, J., van der Riet, P., Westra, W., Nawroz, H., Clayman, G., Piantadosi, S., Corio, R., Lee, D., Greenberg, B., Koch, W., and Sidransky, D. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res., *56:* 2488–2492, 1996.

9. Sidransky, D., Von Eschenbach, A., Tsai, Y. C., Jones, P., Summerhayes, I., Marshall, F., Paul, M., Green, P., Hamilton, S. R., Frost, P., and Vogelstein, B. Identification of p53 gene mutations in bladder cancers and urine samples. Science (Washington DC), *252:* 706–709, 1991.

10. Boyle, J. O., Mao, L., Brennan, J. A., Koch, W. M., Eisele, D. W., Saunders, J. R., and Sidransky, D. Gene mutations in saliva as molecular markers for head and neck squamous cell carcinomas. Am. J. Surg., *168:* 429–432, 1994.

11. Mao, L., Lee, D. J., Tockman, M. S., Erozan, Y. S., Askin, F., and Sidransky, D. Microsatellite alterations as clonal markers for the detection of human cancer. Proc. Natl. Acad. Sci. USA, *91:* 9871–9875, 1994.

12. Brennan, J. A., Mao, L., Hruban, R. H., Boyle, J. O., Eby, Y. J., Koch, W. M., Goodman, S. N., and Sidransky, D. Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. N. Engl. J. Med., *332:* 429–435, 1995.

13. Nawroz, H., Koch, W., Anker, P., Stroun, M., and Sidransky, D. Microsatellite alterations in serum DNA of head and neck cancer patients. Nat. Med., *2:* 1035–1037, 1996.

14. Steiner, G., Schoenberg, M. P., Linn, J. F., Mao, L., and Sidransky, D. Detection of bladder cancer recurrence by microsatellite analysis of urine. Nat. Med., *3:* 621–624, 1997.

15. Gath, H., and Brakenhoff, R. H. Minimal residual disease in head and neck cancer. Cancer Metastasis Rev., *18:* 109–126, 1999.

16. Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. P53 mutations in human cancers. Science (Washington DC), *253:* 49–53, 1991.

17. Greenblatt, M. S., Bennett, W. P., Hollstein, M., and Harris, C. C. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res., *54:* 4855–4878, 1994.

18. Ahrendt, S. A., Halachmi, S., Chow, J. T., Wu, L., Halachmi, N., Yang, S. C., Wehage, S., Jen, J., and Sidransky, D. Rapid p53 sequence analysis in primary lung cancer using an oligonucleotide probe array. Proc. Natl. Acad. Sci. USA, *96:* 7382–7387, 1999.

19. Takeda, S., Ichii, S., and Nakamura, Y. Detection of K-ras mutation in sputum by mutant-allele-specific amplification (MASA). Hum. Mutat., *2:* 112–117, 1993.

20. Landegren, U., Kaiser, R., Sanders, J., and Hood, L. A ligase-mediated gene detection technique. Science (Washington DC), *241:* 1077–1080, 1988.

21. Somers, V. A., Moerkerk, P. T., Murtagh, J. J., Jr, and Thunnissen, F. B. A rapid, reliable method for detection of known point mutations: POINT-EXACCT. Nucleic Acids Res., *22:* 4840–4841, 1994.

22. Kahn, S. M., Jiang, W., Culbertson, T. A., Weinstein, I. B., Williams, G. M., Tomita, N., and Ronai, Z. Rapid and sensitive nonradioactive detection of mutant K-ras genes via enriched PCR amplification. Oncogene, *6:* 1079–1083, 1991.

23. Levi, S., Urbanoispizua, A., Gill, R., Thomas, D. M., Gilbertson, J., Foster, C., and Marshall, C. J. Multiple K-*ras* codon-12 mutations in cholangiocarcinomas demonstrated with a sensitive polymerase chain reaction technique. Cancer Res., *51:* 3497–3502, 1991.

24. Behn, M., Qun, S., Pankow, W., Havemann, K., and Schuermann, M. Frequent detection of ras and p53 mutations in brush cytology samples from lung cancer patients by a restriction fragment length polymorphism-based "Enriched PCR" technique. Clin. Cancer Res., *4:* 361–371, 1998.

25. Ward, R., Hawkins, N., O'Grady, R., Sheehan, C., O'Connor, T., Impey, H., Roberts, N., Fuery, C., and Todd, A. Restriction endonuclease-mediated selective polymerase chain reaction—a novel assay for the detection of K-ras mutations in clinical samples. Am. J. Pathol., *153:* 373–379, 1998.

26. Vogelstein, B., and Kinzler, K. W. Digital PCR. Proc. Natl. Acad. Sci. USA, *96:* 9236–9241, 1999.

27. Fu, D. J., Tang, K., Braun, A., Reuter, D., Darnhoferdemar, B., Little, D. P., O'Donnell, M. J., Cantor, C. R., and Koster, H. Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. Nat. Biotechnol., *16:* 381–384, 1998.

28. Hultman, T., Stahl, S., Hornes, E., and Uhlen, M. Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. Nucleic Acids Res., *17:* 4937–4946, 1989.

29. Hedrum, A., Ponten, F., Ren, Z., Lundeberg, J., Ponten, J., and Uhlen, M. Sequence-based analysis of the human p53 gene based on microdissection of tumor biopsy samples. Biotechniques, *17:* 118–119, 1994.

30. Brennan, J. A., and Sidransky, D. Molecular staging of head and neck squamous carcinoma. Cancer Metastasis Rev., *15:* 3–10, 1996.

31. Sakai, E., and Tsuchida, N. Most human squamous cell carcinomas in the oral cavity contain mutated p53 tumor-suppressor genes. Oncogene, *7:* 927–933, 1992.

32. Burns, J. E., McFarlane, R., Clark, L. J., Mitchell, R., Robertson, G., Soutar, D., and Parkinson, E. K. Maintenance of identical p53 mutations throughout progression of squamous cell carcinomas of the tongue. Eur. J. Cancer B Oral Oncol., *30B:* 335–337, 1994.

33. Tjebbes, G. W. A., Van der Straat, F. G. J. L, Tilanus, M. G. J., Hordijk, G. J., and Slootweg, P. J. P53 tumor suppressor gene as a clonal marker in head and neck squamous cell carcinoma: p53 mutations in primary tumor and matched lymph node metastases. Oral Oncol., *35:* 384–389, 1999.

34. Zariwala, M., Schmid, S., Pfaltz, M., Ohgaki, H., Kleihues, P., and Schafer, R. P53 gene mutations in oropharyngeal carcinomas: a comparison of solitary and multiple primary tumours and lymph-node metastases. Int. J. Cancer, *56:* 807–811, 1994.

35. Ahomadegbe, J. C., Barrois, M., Fogel, S., Le Bihan, M. L., Douc-Rasy, S., Duvillard, P., Armand, J. P., and Riou, G. High incidence of p53 alterations (mutation, deletion, overexpression) in head and neck primary tumors and metastases, absence of correlation with clinical outcome. Frequent protein overexpression in normal epithelium and in early non-invasive lesions. Oncogene, *10:* 1217–1227, 1995.

36. Kropveld, A., van Mansfeld, A. D., Nabben, N., Hordijk, G. J., and Slootweg, P. J. Discordance of p53 status in matched primary tumours and metastases in head and neck squamous cell carcinoma patients. Eur. J. Cancer B Oral Oncol., *32B:* 388–393, 1996.

37. Offner, S., Schmaus, W., Witter, K., Baretton, G. B., Schlimok, G., Passlick, B., Riethmuller, G., and Pantel, K. P53 gene mutations are not required for early dissemination of cancer cells. Proc. Natl. Acad. Sci. USA, *96:* 6942–6946, 1999.

38. Kropveld, A., Rozemuller, E. H., Leppers, F. G. J., Scheidel, K. C., de Weger, R. A., Koole, R., Hordijk, G. J., Slootweg, P. J., and Tilanus, M. G. J. Sequencing analysis of RNA and DNA of exons 1 through 11 shows p53 gene alterations to be present in almost 100% of head and neck squamous cell cancers. Lab. Invest., *79:* 347–353, 1999.

39. Van Houten, V., Denkers, F., van Dijk, M., van den Brekel, M., and Brakenhoff, R. Labeling efficiency of oligonucleotides by T4 polynucleotide kinase depends on 5′-nucleotide. Anal. Biochem., *265:* 386–389, 1998.

40. Sambrook, J., Fritsch, E. F., and Maniatis T. Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1989.

41. Sommer, S. S., Groszbach, A. R., and Bottema, C. D. K. PCR amplification of specific alleles (PASA) is a general-method for rapidly detecting known single-base changes. Biotechniques, *12:* 82–87, 1992.

42. Tindall, K. R., and Kunkel, T. A. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. Biochemistry, *27:* 6008–6013, 1988.

43. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science (Washington DC), *239:* 487–491, 1988.

44. Knudson, A. G., Jr. Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. USA, *68:* 820–823, 1971.

45. Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., Meltzer, S. J., Rodriguez-Bias, M. A., Fodde, R., Ranzani, G. N., and Srivastava, S. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res., *58:* 5248–5257, 1998.

46. Aaltonen, L. A., Peltomaki, P., Leach, F. S., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., Powell, S. M., Jen, J., Hamilton, S. R., Peterson, G. M., Kinzler, K. W., Vogelstein, B., and de la Chapelle, A. Clues to the pathogenesis of familial colorectal cancer. Science (Washington DC), *260:* 812–816, 1993.

47. Kinzler, K. W., and Vogelstein, B. Lessons from hereditary colorectal cancer. Cell, *87:* 159–170, 1996.

48. Bonner, R. F., Emmert-Buck, M., Cole, K., Pohida, T., Chuaqui, R., Goldstein, S., and Liotta, L. A. Laser capture microdissection: molecular analysis of tissue. Science (Washington DC), *278:* 1481–1483, 1997.

49. Litt, M., Hauge, X., and Sharma, V. Shadow bands seen when typing polymorphic dinucleotide repeats: some causes and some cures. Biotechniques, *15:* 280–284, 1993.

50. Chen, X. Q., Stroun, M., Magnenat, J. L., Nicod, L. P., Kurt, A. M., Lyautey, J., Lederrey C., and Anker, P. Microsatellite alterations in plasma DNA of small cell lung cancer patients. Nat. Med., *2:* 1033–1035, 1996.

51. Mao, L., Schoenberg, M. P., Scicchitano, M., Erozan, Y. S., Merlo, A., Schwab, D., and Sidransky, D. Molecular detection of primary bladder cancer by microsatellite analysis. Science (Washington DC), *271:* 659–662, 1996.

52. El-Naggar, A. K., Hurr, K., Huff, V., Clayman, G. L., Luna, M. A., and Batsakis, J. G. Microsatellite instability in preinvasive and invasive head and neck squamous carcinoma. Am. J. Pathol., *148:* 2067–2072, 1996.

53. Piccinin, S., Gasparotto, D., Vukosavljevic, T., Barzan, L., Sulfaro, S., Maestro, R., and Boiocchi, M. Microsatellite instability in squamous cell carcinomas of the head and neck related to field cancerization phenomena. Br. J. Cancer, *78:* 1147–1151, 1998.

54. Ashworth, T. A. A case of cancer in which cells similar to those in the tumors were seen in the blood after death. Aust. Med. J., *14:* 146, 1869.

55. Lindemann, F., Schlimok, G., Dirschedl, P., Witte, J., and Riethmüller, G. Prognostic significance of micrometastatic tumor cells in bone marrow of colorectal cancer. Lancet, *340:* 685–689, 1992.

56. Pantel, K., Izbicki, J., Passlick, B., Angstwurm, M., Häussinger, K., Thetter, O., and Riethmüller, G. Frequency and prognostic significance of isolated tumour cells in bone marrow of patients with non-small-cell lung cancer without overt metastases. Lancet, *347:* 649–653, 1996.

57. Braun, S., Pantel, K., Muller, P., Janni, W., Hepp, F., Kentenich, C. R. M., Gastroph, S., Wischnik, A., Dimpfl, T., Kindermann, G., Riethmuller, G., and Schlimok, G. Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. N. Engl. J. Med., *342:* 525–533, 2000.

58. Brakenhoff, R. H., Stroomer, J. G. W., ten Brink, C., de Bree, R., Weima, S. M., Snow, G. B., and Van Dongen, G. A. M. S. Sensitive detection of squamous cells in bone marrow and blood of head and neck cancer patients by E48 reverse transcriptase polymerase chain reaction. Clin. Cancer Res., *5:* 725–732, 1999.

59. Liefers, G. J., Cleton-Jansen, A. M., van de Velde, C. J., Hermans, J., van Krieken, J. H., Cornelisse, C. J., and Tollenaar, R. A. Micrometastases and survival in stage II colorectal cancer. N. Engl. J. Med., *339:* 223–228, 1998.

60. Chelly, J., Concordet, J. P., Kaplan, J. C., and Kahn, A. Illegitimate transcription: transcription of any gene in any cell type. Proc. Natl. Acad. Sci. USA, *86:* 2617–2621, 1989.

61. Datta, Y. H., Adams, P. T., Drobyski, W. R., Ethier, S. P., Terry, V. H., and Roth, M. S. Sensitive detection of occult breast cancer by the reverse-transcriptase polymerase chain reaction. J. Clin. Oncol., *12:* 475–482, 1994.

62. Krismann, M., Todt, B., Schroder, J., Gareis, D., Muller, K. M., Seeber, S., and Schutte, J. Low specificity of cytokeratin 19 reverse transcriptase-polymerase chain reaction analyses for detection of hematogenous lung cancer dissemination. J. Clin. Oncol., *13:* 2769–2775, 1995.

63. Dingemans, A-M. C., Brakenhoff, R. H., Postmus, P. E., and Giaccone, G. Detection of cytokeratin 19 transcripts by reverse transcriptase polymerase chain reaction in lung cancer cell lines and blood of lung cancer patients. Lab. Invest., *77:* 213–220, 1997.

64. Savtchenko, E. S., Schiff, T. A., Jiang, C. K., Freedberg I. M., and Blumenberg, M. Embryonic expression of the human 40-kd keratin—evidence from a processed pseudogene sequence. Am. J. Hum. Genet., *43:* 630–637, 1988.

65. Ruud, P., Fodstad, O., and Hovig, E. Identification of a novel cytokeratin 19 pseudogene that may interfere with reverse transcriptase-polymerase chain reaction assays used to detect micrometastatic tumor cells. Int. J. Cancer., *80:* 119–125, 1999.

66. Gerhard, M., Juhl, H., Kalthoff, H., Schreiber, H. W., Wagener, C., and Neumaier, M. Specific detection of carcinoembryonic antigen-expressing tumor cells in bone marrow aspirates by polymerase chain reaction. J. Clin. Oncol., *12:* 725–729, 1994.

67. Soeth, E., Vogel, I., Roder, C., Juhl, H., Marxsen, J., Kruger, U., Henne-Bruns, D., Kremer, B., and Kalthoff, H. Comparative analysis of bone marrow and venous blood isolates from gastrointestinal cancer patients for the detection of disseminated tumor cells using reverse transcription PCR. Cancer Res., *57:* 3106–3110, 1997.

68. Peck, K., Sher, Y-P., Shih, J-Y., Roffler, S. R., Wu, C-W., and Yang, P-C. Detection and quantitation of circulating cancer cells in the peripheral blood of lung cancer patients. Cancer Res., *58:* 2761–2765, 1998.

69. Eschwège, P., Dumas, F., Blanchet, P., Le Maire, V., Benoit, G., Jardin, A., Lacour, B., and Loric, S. Haematogenous dissemination of prostatic epithelial cells during radical prostatectomy. Lancet, *346:* 1528–1530, 1995.

70. Borgen, E., Beiske, K., Trachsel, S., Nesland, J. M., Kvalheim, G., Herstad, T. K., Schlichting, E., Qvist, H., and Naume, B. Immunocytochemical detection of isolated epithelial cells in bone marrow: non-specific staining and contribution by plasma cells directly reactive to alkaline phosphatase. J. Pathol., *185:* 427–434, 1998.

71. Funke, I., and Schraut, W. Meta-analyses of studies on bone-marrow micrometastases: an independent prognostic impact remains to be substantiated. J. Clin. Oncol., *16:* 557–566, 1998.

72. Klein, C. A., Schmidt-Kittler, O., Schardt, J. A., Pantel, K., Speicher, M. R., and Riethmuller, G. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells Proc. Natl. Acad. Sci. USA, *96:* 4494–4499, 1999.

73. Van Hal, N. L. W., Van Dongen, G. A. M. S, Stigter-van Walsum, M., Snow, G. B., and Brakenhoff, R. H. Characterization of CD44v6 isoforms in head and neck squamous cell carcinoma. Int. J. Cancer, *82:* 837–845, 1999.

74. Zippelius, A., Kufer, P., Honold, G., Kollermann, M. W., Oberneder, R., Schlimok, G., Riethmuller, G., and Pantel, K. Limitations of reverse-transcriptase polymerase chain reaction analyses for detection of micrometastatic epithelial cancer cells in bone marrow. J. Clin. Oncol., *15:* 2701–2708, 1997.