the distribution of mutations and in the pattern of mutational events when compared to tumors indicating that they are representative of the cells from which they were derived. Surprisingly, we found discrepancies in the p53 status in 23% of cell lines, some of which are widely used, such as MOLT-4 or CAPAN-1.

## Results and Discussion

**p53 mutations in cell lines versus tumors.** The pattern of p53 mutations can be analysed in two informative ways, either by examining the distribution of p53 mutations in the p53 protein or by scoring the various mutational events that lead to these mutations. Both types of analysis have been very informative when applied to various types of human tumors.[25] These studies demonstrate a link between exposure to various types of carcinogens and the development of specific cancers. The most striking example is that of tandem mutations, specifically induced by ultraviolet radiation, which are only observed in skin cancers.[26] The relationships between G→T transversion and lung cancer in smokers or mutation of codon 249 observed in aflatoxin B1-induced liver cancers are also very demonstrative.[25,27] The distribution of p53 mutations along the p53 protein is similar in tumors and cell lines, indicating that there is no bias in the selection of specific mutant p53 during establishment of a cell culture (Fig. 1A and data not shown). The only exception concerns colorectal cancer cell lines. p.R175H is one of the most frequent p53 mutations in tumors, but is very rare in colorectal cancer cell lines (Suppl. Fig.). This finding is specific for p.R175H and it is not observed for the other two hot spots at codons 248 and 273. The reason for this bias is not known. Comparison of the various mutational events in cell lines and tumors has been performed for all cancers together or for 8 cancer types (Fig. 1A and B and Suppl. Figs.). As previously observed, the pattern of mutations differs between various types of cancers, but there is a striking similarity when comparing tumors and cell lines from the same origin. In colorectal and brain cancer, there is a predominance of GC→AT transition at CpG dinucleotides, whereas in lung cancer or head and neck SCC, the frequency of GC→TA transversion is 30% and 20%, respectively, with only a few transitions at CpG dinucleotides. This high frequency of transversion in these cancers has been shown to be associated with tobacco smoking and will not be discussed in more detail here.[28] This similarity in the pattern of p53 mutations in primary tumors and cell lines is a strong argument suggesting that these p53 mutations did not occur de novo during the establishment of these cell lines. It also supports the small number of studies that have found matched p53 mutations in primary tumors that were used to establish cell lines and confirms that analysis of the spectrum of mutations in oncogenes or tumor suppressor genes in human cell lines accurately reflects the situation observed in primary tumors.

**Analysis of p53 mutant activity in cell lines.** Analysis of p53 mutations in human tumors has led to the discovery that at least 5% to 10% of published p53 mutations could be due to PCR or sequencing artefacts.[22] However, these mutations are not randomly distributed among the 2,500 publications reporting p53 mutations. A meta-analysis identified about 30 publications (1,600 p53 mutations) with a high concentration of unusual p53 mutations that shared the following properties: (i) multiple p53 mutations in the same tumor (3 to 14); (ii) a high frequency of synonymous mutations; (iii) a low frequency of mutations at hot spot codons; (iv) most of these mutations retained either partial or total transactivational activity.[29] The vast majority of these studies were associated with the use of nested PCR for amplification and analysis of the p53 gene. Analysis of p53 mutations in cell lines provides several advantages over analysis of tumors to minimize artefactual data: (i) DNA extracted from cell lines is available in large quantities. Analysis requires neither nested PCR nor excessive numbers of PCR cycles and can be easily repeated; (ii) The high quality of the DNA avoids PCR problems associated with DNA extracted from paraffin-embedded tissue; (iii) DNA is not contaminated by normal DNA from stroma or cells or infiltrating lymphocytes.

The UMD p53 mutation database includes functional information about the majority of p53 missense mutants, as originally published by Kato et al.,[23] (see also material and methods). Quantitative data concerning the transcriptional activity of each missense p53 mutation has been extremely useful to classify and analyse p53 mutations in the p53 database.[21,22,29] The mean and 95% confidence interval (CI) of the remaining activity of all mutant p53 proteins found in cell lines or in tumors was calculated by using the activity measured on the p21WAF1 promoter (similar results were obtained with the activity measured on seven other promoters of transcription, data not shown). The analysis shows that the mean activity was situated between -1 and -1.2. This value corresponds to a residual transcriptional activity of about 10% compared to wild-type p53. The narrower 95% CI in tumors compared to cell lines is due to the greater number of tumors used in the analysis (Fig. 2A). In the majority of cancers, residual p53 activity was lower in cell lines than in tumors, but this difference was only marginally significant in head and neck, breast and SCLC, $p = 0.03$). On the other hand, residual p53 activity has a wider distribution in tumors compared to cell lines (variance analysis, Fig. 2B). A large number of mutant p53 retain wild-type activity in tumors, but this feature is rarely observed in cell lines. This difference was highly significant for all cancer types ($p < 0.0001$) except for brain cancers and haematological malignancies. Two non-exclusive explanations can be proposed for this difference between tumors and cell lines. First, it is possible that only tumors with fully inactivated p53 are preferentially selected to establish cell lines. This hypothesis could also explain why the frequency of p53 mutations is always higher in cell lines than in tumors. It is also possible that this profile of p53 inactivation in cell lines is more representative of the true pattern of p53 inactivation and that the tumor p53 database contains passenger mutations and/or artefactual mutations with partial or fully active p53.[21]

During the course of these analyses, we also observed that 82 cell lines displayed two p53 missenses mutations. Preliminary observations suggested that the two mutations may not have the same importance and that only one mutation was the driving force selected during transformation.[22] In order to obtain more information, clustering analysis was performed on cell lines with either single (SM cell lines) or double mutations (DM cell lines). Three clusters were obtained for the two populations, corresponding to mutant p53 with wild-type activity (cluster I), intermediate residual activity (cluster II) or no activity (cluster III) (Table 1). The number of mutants in clusters I and II was significantly higher in DM cell lines than in SM cell lines, whereas mutations with total loss of activity were more frequent in SM cell lines ($p < 0.0001$, Table 1). Mutations in DM cell lines were further analysed to determine how paired mutations

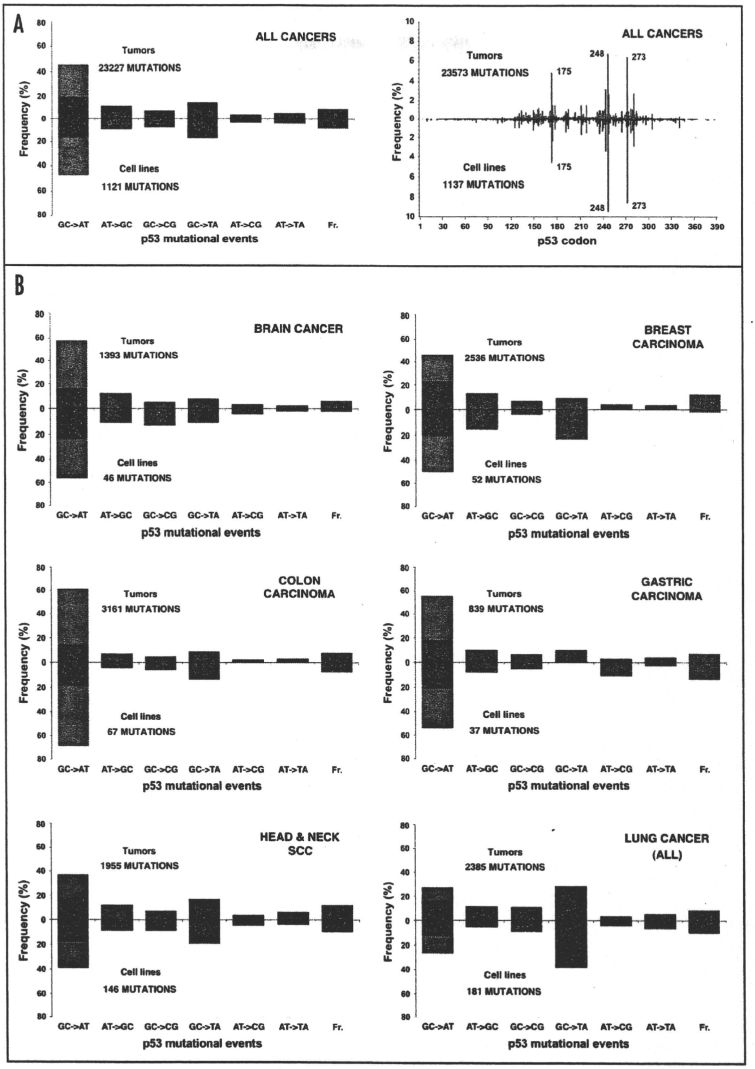


Figure 1. Mutation spectrum in tumours and cell lines: (A) Mutational events (left) and distribution of mutations (right) in all tumours (upper part) and cell lines (lower part). Data were obtained from the UMD p53 database, 2007_R1 release (http://p53/free/fr). (B) mutational events in tumours versus cell lines in various types of cancer. A similar pattern of mutational events is observed for other cancers (melanoma, ovarian carcinoma, oesophageal carcinoma or pancreatic carcinoma, data not shown). Transitions at CpG dinucleotides are shown in red.

(two mutations in a single cell line) were associated (Table 2). Only one of the 41 cell lines presented two mutations in cluster I with wt activity for the two p53 mutant alleles. This choriocarcinoma cell line (NUC-1) displays two unusual p53 mutations at codons 17 and 24 that have never been observed in any other cell lines or tumors. Among the 11 remaining mutations in cluster I, three were paired with mutations in cluster II and 8 were paired with mutations in cluster III. Among the 19 mutations in cluster II, two were paired with a mutation of the same class, 3 with class I mutations and 12 with class III mutations. The majority (30) of the 50 mutations in cluster III were paired with a mutation of the same class and 12 were paired with class II mutations (Table 2).

Double mutations can occur in two configurations, either on the same allele (DMS) or on two different alleles (DMD). Unfortunately, in the majority of cases, this status is unknown (DMU). In the p53 mutation database based on tumors, the majority of DM with a known configuration are DMD (about 90%). No cell lines with two missense mutations in the same allele have been reported and only 10 cell lines with mutations on two different alleles have been reported. All of these cell lines expressed one class III mutation associated with either another class III mutation (6), or class II (3) or class I (1) mutations.

Altogether, our results indicate that: (i) there is a higher frequency of weak mutations in DM than in SM mutations and (ii) the majority of these weak mutations are paired with a more potent mutation. This suggests that the two mutants do not have the same contribution to the transforming process. Whether or not these weak mutations are passenger mutations associated with a driving mutation or true mutations associated with selection of the transforming phenotype is an unresolved question. One of the main problems associated with p53 mutations is the possible dominant negative activity of mutant p53 via hetero-oligomerization making it very difficult to reach any definitive conclusions concerning weak p53 mutations. Weakening of the second allele could possibly accentuate the dominant negative activity of p53.

**p53 status in human tumor cell lines.** The NCI-60 panel is a good example of a series of cell lines that are widely used for both basic research and drug discovery.[1] This panel originally contained 60 cell lines from nine histological origins (Table 3). Several observations unrelated to p53 status revealed that some cell lines were either mixed up or were derived from the same donor (Table 3).[5] At least 100 studies have analysed the p53 status of a subset of the panel and in 1997, O'Connor et al., reported the p53 status of the entire NCI-60 panel.[16] This paper has been used as a reference for 10 years despite discrepancies with other data in the literature. A second analysis of the entire NCI-60 panel was performed in 2006 and the results are fairly heterogeneous compared to the 1997 study (Table 3). Inspection of the two studies leads to the detection of 19 apparent differences (Table 3). Three differences were due to typographical errors in the 1997 report (RPMI-8226, SK-MEL-28 and Hs-578-T). A more careful examination of four other discrepancies reveals that they are due to a problem of nomenclature associated with a different mutation screening strategy. In the 1997 paper, p53 mutations were analysed by cDNA sequencing, while the 2007 analysis was performed using genomic DNA as starting material. One of the disadvantages of RNA-based analysis is that it is impossible to infer whether deletions found in the cDNA are due to splicing

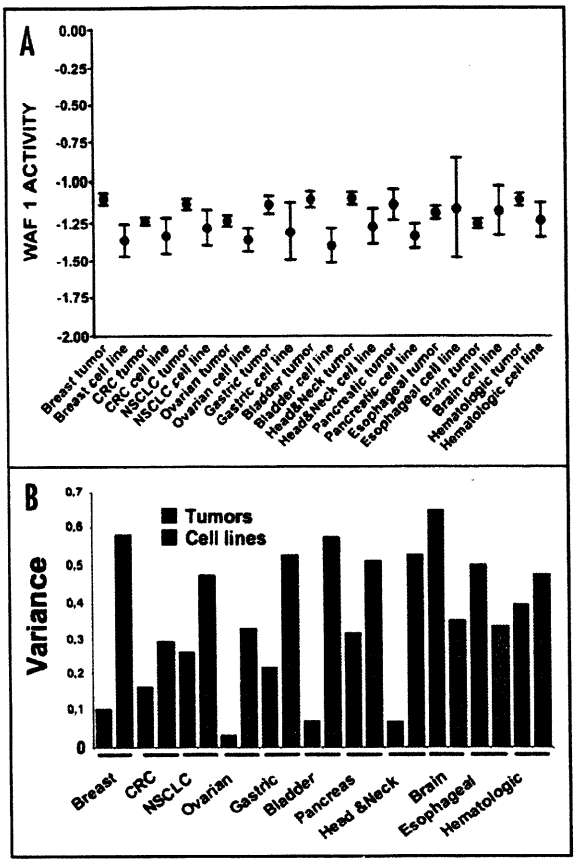


Figure 2. Analysis of the residual p53 activity of mutant p53 in tumors and cell lines. (A) Points, mean p53 activity as measured by transactivation with the p21WAF1 promoter; bars, 95% CI. A similar distribution was observed with other p53 response genes (data not shown). The y-axis corresponds to p53 transactivation activity, with a value of -1.5 for the negative control and 2.5 for wild-type p53. (B) Variance of the p21WAF1 promoter activity in tumors and cell lines. CRC, colorectal carcinoma; NSCLC, non-small cell lung cancer. Data from cell lines and tumors are displayed in black and red respectively.

mutations or intragenic deletions in the gene. On the other hand, it is always difficult to predict the consequence of mutations found in intron or splice junctions after genomic sequencing. Both methods are complementary and may be necessary to ensure an accurate genetic status.

In HOP-62, RNA-based analysis detected an insertion between codon 212–225 but no information about the insertion sequence was available. Codon 225 is at the boundary of exon 6 and intron 6 suggesting a splicing defect, as analysis at the genomic level confirms the presence of a splice mutation in the acceptor signal of exon 6 (Table 3).

In OVCAR-8, the 126–132 deletion detected by the RNA-based assay concerns the first six residues of exon 5. Genomic analysis described a mutation in the acceptor site of exon 5 and a splicing defect leading to a shift of the normal donor site of exon 5 that skips 18 nucleotides (6 aa residues) in exon 5. Examination of the DNA sequence at codon 132 reveals an AG dinucleotide sequence preceded by a pyrimidine tract similar to those found in the splice donor sequence. The same situation is observed for NCI/ADR-RES

**Table 1 Cluster analysis of p53 mutation activity**

| | SM | DM |
|---|---|---|
| Cluster I (wt activity) | 27 (3.4%) | 13 (15.9%) |
| Cluster II (low acitivity) | 73 (9.3%) | 19 (23.2%) |
| Cluster III (no activity) | 687 (87.5%) | 50 (61.0%) |
| Total | 787 (100%) | 82 (100%) |

The table entries are the number (and %) of mutants classified into the three clusters based on k-means clustering of the promoter activities of p53 target genes. There are significantly more cluster-I and cluster-II mutations among the double mutations (DM) than among the single mutations (SM) ($p = 2e-10$ using the chi-square test).

**Table 2 Discordance table of class assignment of the 82 DM mutations (from 41 pairs)**

| | Cluster I | Cluster II | Cluster III |
|---|---|---|---|
| Cluster I (wt activity) | 1 | 3 | 8 |
| Cluster II (low acitivity) | 0 | 2 | 9 |
| Cluster III (no activity) | 2 | 3 | 15 |

Majority of the weak mutations (cluster I and cluster II) are paired with strong mutations (cluster III).

that has been recently shown to be an ovarian carcinoma cell line originating from the same patient as OVCAR-8.

In EKVX, the deletion of codon 187 to 224 detected on RNA-based analysis corresponds exactly to the deletion of the entire exon 6, a strong argument for a splicing defect. Genomic analysis did not reveal a splicing defect but a tandem mutation at codons 203 and 204 in exon 6 (Table 3). If the two cell lines analysed were really EKVX, this result suggests that a mutation at either codon 203 and/or 204 could affect p53 gene splicing. This observation is not surprising, as it is now well known that exons contain exonic splicing enhancers (ESE) that regulate either alternative splicing or normal splicing.[30] These ESE are recognized by the SR proteins that regulate the various splicing events. Mutations in ESE have been identified in numerous genes including APC or NF1.[31,32] Exonic mutations that can change p53 splicing have also been described.[33,34] Taken together, the contradictions noted in the p53 status of the four cell lines, HOP-62, OVCAR-8, NCI/ADR-RES and certainly EKVX are only due to the different strategies used for their analysis and a lack of homogeneity in the nomenclature used to report p53 mutations. The problem of the nomenclature of p53 mutations as well as other gene defects is a recurrent problem in publications.[35] Despite numerous recommendations, the description of p53 mutations in the literature is highly heterogeneous and can reach a high degree of fantasy with tables that are either totally non-informative or with so many typographical errors that they cannot be interpreted. In a recent survey, the editors of 80 journals with frequent publications of p53 mutations were contacted in order to stress this problem and define certain guidelines for the publication of p53 mutations (Soussi T, Unpublished). Unfortunately, this survey was a complete failure with less than 10% of replies and no change in the trends of reporting accurate p53 mutations. In fact, the number of typographical errors or incomprehensible mutations has increased over the last five years (Soussi T, unpublished observations).

After eliminating typographical errors and possible splice mutations, the p53 status of 15 cell lines was different between the two studies. Using the UMD p53 database and the literature, we checked for other publications that have analysed the p53 status of these cell lines. For two cell lines, CCRF-CEM and HL-60, sufficiently concordant publications are available to define a consensus concerning the p53 status (Table 3). For 13 cell lines, analysis of the literature revealed a very heterogeneous situation and no consensus could be reached (Table 3, Inconclusive). Cell lines such as MOLT-4 or NCI-H226 represent an extreme situation, as multiple publications do not show any common p53 mutations. For other cell lines such as DU-145, which have been shown to display two different p53 mutations in two different alleles (p.V274F and p.P223L), the ambiguity concerns the fact that several authors have detected only one of the two mutations, either p.V274F or p.P223L. It is therefore possible that during long-term cell culture, one of the two mutant p53 alleles is lost, as no selection pressure is exerted on cell growth.

A similar situation is observed for other cell lines that do not belong to the NCI-60 panel, but with many discrepancies (Table 4 and Suppl. Table S1, see also p53 website). In many cell lines, the p53 status has been analysed in only one or two reports and the information is subsequently reproduced in the literature. This is a very dangerous situation as it could lead to erroneous phenotype-genotype correlations in various types of studies. The pancreatic carcinoma cell line CAPAN-2 is a good example of the problems raised by erroneous phenotype. This cell line has been described as either wt, mutated (p.R273H) or p53 null (Table 4 and reference within). A Pubmed literature search indicates that all three phenotypes are used in various studies.

The "p53-null" status is used in different ways in the literature. The two most common meanings are a cell line with a documented p53 gene deletion (both alleles) or a cell line with a p53 mutation. We have also observed more "unusual" situations in which this status is only based on p53 expression (RNA or protein). Unfortunately, this type of information diffuses rapidly in the literature without any verification of the original publication. The p53 status of the two cell lines SK-OV-3 (Ovarian cancer) and FRO (anaplastic thyroid carcinoma cell line) are a good example of this ambiguity. In the majority of publications, the p53 status of these two cell line is stated as "p53 null". In fact, close examination of the original manuscript shows that the p53 gene in SK-OV-3 is not deleted and did not sustain any gross rearrangement but neither p53 RNA or protein are found. In these publications, no p53 mutations were found but the recent analysis performed at the Sanger Institute detected a deletion of a single nucleotide at position 267 (codon 90).[3] It is therefore possible that nonsense-mediated mRNA decay (NMD) eliminates p53 aberrant mRNA. NMD has been observed in the human leukaemia cell line K562 where p53 is also inactivated via a 1 base pair insertion at nucleotide 136. For the FRO cell line, the original reference for the analysis of the p53 gene status is always correctly quoted, but a closer look at this original paper demonstrates a marked decrease of p53 RNA in the cell but no mutation was detected by sequencing exons 5 to 8. Either a mutation is situated outside this region leading to a decrease of RNA expression (frameshift mutation associated with Nonsense-Mediated mRNA Decay) or the altered p53 expression is due to another mechanism. Because the whole p53 gene is present, it is incorrect to define SK-OV-3 or FRO cell lines as "p53 null", as

**Table 3 p53 status in the NCI-60 panel cell lines**

| Cell line | ATCC number | Cancer | Mutation in UMD[1] | | Mutation in COSMIC[2] | | Consensus[3] | Comments |
|---|---|---|---|---|---|---|---|---|
| | | | Mutation | Reference | Protein | DNA | | |
| CCRF-CEM | CCL-119 | *Leukemia* | R248Q | 16 | p.R175H, p.R248Q | c.524G>A, c.743G>A | p.[R175H] + p.[R248Q] | Two mutations in separate alleles. Derivative CCRF-CEM-VLB100 has a third mutation |
| | | | R175H+ R248Q | 394 | | | | |
| HL-60[5] | CCL-240 | | R248L | 16 | p.M1_*394del | c.1_1182del1182 | p.M1_*394del | |
| | | | p53 Null | 40 4 | | | | |
| K-562 | CCL-243 | | ND | 16 | p.Q136fs*13 | c.406_407insC | p.Q136fs*13 | |
| | | | 136ins1 | 41 4 | | | | |
| MOLT-4[6] | CRL-1582 | | wt | 16,39 | p.R306X | c.916C>T | Inconclusive | |
| | | | L111V | 42 | | | | |
| | | | R248Q | 43 | | | | |
| RPMI-8226 | CCL-155 | | E285L[7] | 16 | p.E285K | c.853G>A | p.E285K | This mutant is temperature-sensitive |
| | | | E285K | 44 | | | | |
| SR | CRL-2262 | | wt | 16 | wt | wt | wt | |
| A549 | CCL-185 | *Non-Small Cell Lung* | wt | 16,45 | wt | wt | wt | |
| EKVX | | | del187-224 | 16 | p.V203_E204>V* | c.609_610GG>TT | Inconclusive | Possible splicing defect See text for discussion |
| HOP-62 | | | Ins212-225 | 16 | p.? | c.673-2A>G | Splicing defect | See text for discussion |
| HOP-92 | | | R175L | 16 | p.R175L | c.524G>T | p.R175L | |
| NCI-H226 | CRL-5826 | | P309A | 16 | wt | wt | Inconclusive | |
| | | | R158L | 46 | | | | |
| NCI-H23 | CRL-5800 | | M246I | 16,47 | p.M246I | c.738G>C | p.M246I | |
| NCI-H322M | | | R248L | 16,46 | p.R248L | c.743G>T | p.R248L | |
| NCI-H460 | HTB-177 | | wt | 16 | wt | wt | wt | |
| NCI-H522 | CRL-5810 | | 191delG | 16, 46 | p.P191fs*56 | c.572_572delC | p.P191fs*56 | |
| COLO-205 | CCL-222 | *Colon* | G266E | 16 | p.Y103_L111>L | c.308_333>TA | Inconclusive | |
| | | | 103del27 | 48 | | | | |
| HCC-2998 | | | R213X | 16 | p.R213x | c.637C>T | p.R213x | |
| HCT-116 | CCL-247 | | wt | 16 | wt | wt | wt | |
| HCT-15 | CCL-225 | | P153A | 16 | p.S241F, p.? | c.722C>T, c.1101-2A>C | Inconclusive | HCT-15 and DLD1 are derived from the same individual. DLD1 display a p.S241F mutation |
| | | | S241F | 49 | | | | |
| HT-29 | HTB-38 | | R273H | 16, 48 | p.R273H | c.818G>A | p.R273H | |
| KM12 | | | H179R | 16 | p.R72fs*51 | c.215delG | Inconclusive | |
| SW620 | CCL-227 | | R273H | 16, 50 | p.R273H, p.P309S | c.818G>A, c.925C>T | p.[R273H] (+) p.[P309S] | SW480 and SW620 are derived from the same individual with a similar p53 alteration. The p.P309S mutation is not always reported |
| | | | R273H, P309S | 51 | | | | |
| SF268 | | *CNS* | R273H | 16,52 | p.R273H | c.818G>A | p.R273H | |
| SF295 | | | R248Q | 16 | p.R248Q | c.743G>A | p.R248Q | |
| SF539 | | | wt | 16 | p.R342fs*3 | c.1024delC | Inconclusive | |
| SNB75 | | | E258K | 16 | p.E258K | c.772G>A | p.E258K | |
| U251/SNB19 | | | R273H | 16,53 | p.R273H | c.818G>A | p.R273H | SNB19 and U251 cell lines are derived from the same individual and are similar |

in the case of H1299 or Saos-2 cell lines in which the p53 gene is entirely deleted. These cell lines are commonly used as recipients to reintroduce either wild-type of mutant p53. Whether the presence of an endogenous p53 gene which is still transcriptionally active in the SK-OV-3 or FRO cell could interfere with this reconstitution experiment is not known, but should be carefully considered before conducting this type of experiment. The recent finding of p53 isoforms that could be expressed by alternative splicing may also increase the complexity of this problem, as the various delta133 isoforms could be theoretically expressed in this cell line.

Another reason why "p53-null" should be used cautiously to describe cell lines that express mutant p53 is the observation that p53 mutations are fairly heterogeneous in terms of loss of function and several cell lines display a normal or partial p53 response. Finally, there is now ample evidence that some mutant p53 behave as dominant oncogenes with a gain of function activity. We therefore believe that the "p53 null" status should be used only for cell lines that are totally devoid of p53 gene. Any other situation should be referred to as "mutant p53".

The UMD_p53 database (2007_R1 release) includes p53 mutations in 1,211 cell lines: 827 of these mutations have only been described once, preventing any verification. A discrepancy was detected in 88 of the remaining 384 cell lines (23%), in line with the study by Macleod et al., who showed that 18% of cell lines in the DSMZ-German Collection of Microorganisms and cell Cultures

**Table 3 p53 status in the NCI-60 panel cell lines (continued)**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| LOXIMVI | | *Melanoma* | wt | 16 | wt | wt | wt | |
| Malme-3M | HTB-64 | | wt | 16 | wt | wt | wt | |
| M14 | | | G266E | 16 | p.G266E | c.797G>A | p.G266E | |
| SK-MEL-2 | HTB-68 | | G245S | 16 | p.G245S | c.733G>A | p.G245S | |
| SK-MEL-28 | HTB-72 | | C145V[7] | 16 | p.L145R | c.434_435TG>GT | p.L145R | |
| | | | L145R | 54 | | | | |
| SK-MEL-5 | HTB-70 | | wt | 16 | wt | wt | wt | |
| UACC-257 | | | wt | 16 | wt | wt | wt | |
| UACC-62 | | | wt | 16 | wt | wt | wt | |
| MDA-MB-435 | | | G266E | 16, 55 | p.G266E | c.797G>A | p.G266E | This cell line was originally reported as a breast carcinoma cell line but recent SNP analysis indicates that it is similar to the M14 melanoma cell line |
| MDA-N | | | G266E | 16 | ND | ND | p.G266E | This cell line is a derivative of MDA-MB-435 transfected with a plasmid expressing erbB2 |
| IGROV1 | | *Ovarian* | wt | 16, 56 | p.Y126C | c.377A>G | Inconclusive | |
| OVCAR-3 | HTB-161 | | R248Q | 16, 57 | p.R248Q | c.743G>A | p.R248Q | |
| OVCAR-4 | | | wt | 16 | p.L130V | c.388C>G | Inconclusive | |
| OVCAR-5 | | | ins224 | 16, | wt | wt | Inconclusive | |
| | | | 224ins3[8] | 58 | | | | |
| OVCAR-8 | | | del 126-132 | 16 | p.? | c.376-1G>A | Splicing defect | Same as NCI/ADR-RES |
| NCI/ADR-RES | | | del 126-132 | 16 | p.? | c.376-1G>A | Splicing defect | Originally labelled as MCF-7/Adr but was later found to be different from MCF-7. SNP analysis indicates that it is similar to OVCAR-8 |
| | | | 126del21 | 59 | | | | |
| SK-OV-3[9] | HTB-77 | | H179R | 16 | p.S90fs*33 | c.267delC | Inconclusive | |
| 786-0 | CRL-1932 | *Renal* | P278A | 16 | p.P278A, p.? | c.832C>G, c.560-2A>G | Inconclusive | |
| A498 | HTB-44 | | wt | 16 | wt | wt | wt | |
| ACHN | CRL-1611 | | wt | 16 | wt | wt | wt | |
| CAKI-1 | HTB-46 | | wt | 16 | wt | wt | wt | |
| RXF393 | | | R175H | 16 | p.R175H | c.524G>A | p.R175H | |
| SN12C | | | E336X | 16 | p.E336X | c.1006G>T | p.E336X | |
| TK10 | | | L264R | 16 | p.L264R | c.791T>G | p.L264R | |
| UO31 | | | wt | 16 | wt | wt | wt | |
| DU-145 | HTB-81 | *Prostate* | P223L | 16 | p.V274F | c.820G>T | Inconclusive | See text for more information |
| | | | P223L, V274F | 60,61 | | | | |
| PC-3 | CRL-1435 | | 138del | 16 | p.K139fs*31 | c.414delC | p.K139fs*31 | |
| | | | 138del1 | 60 | | | | |
| | | | R282W | 62 | | | | |
| BT-549 | HTB-122 | *Breast* | ND | 16 | p.R249S | c.747G>C | p.R249S | |
| | | | R249S | 63 | | | | |
| Hs 578T | HTB-126 | | D157E[7] | 16 | p.V157F | c.469G>T | p.V157F | |
| | | | V157F | 64 [4] | | | | |
| MCF7 | HTB-22 | | wt | 16 | wt | wt | wt | |
| MDA-MB-231 | HTB-26 | | R280K | 16, 63 | p.R280K | c.839G>A | p.R280K | |
| T47D | HTB-133 | | L194F | 16,51 | p.L194F | c.580C>T | p.L194F | |

[1]Mutations as reported in the 2007_R1 of the UMD p53 mutation database. The description of the mutations have been left as originally published by the authors; [2]Mutations described by Ikediobi et al.,[3]; [3]A mutation consensus was defined for cell lines using the following rules: (i) at least two independent studies reporting sequencing and identifying the same mutation without any contradictory reports; (ii) at least three independent studies reporting sequencing and identifying the same mutation and one fourth contradictory report. All other possibilities were not considered to be consensual and have been assigned as uncertain. The nomenclature for TP53 mutation uses either the cDNA (RefSeqNM_000546.2) or the protein (RefSeq NP 000537) as reference: For numbering, +1 is A of the ATG initiation codon in the correct RefSeq (NM_000546.2). Mutations are described using the international nomendature[65] and http://www.hgvs.org/mutnomen/; [4]Mutation found independently by multiple authors. Only the first publication is shown; [5]HL 60(TB) was used for the analysis, but it is reported to have a p53 deletion similar to HL60; [6]The status of MOLT-4 is highly heterogeneous in the literature. The report of a wt status could be due to the fact that only exons 5 to 8 (residues 126–306) were screened in several publications; [7]Typographical error in the publication; [8]It is not clear whether these authors checked the p53 status of the cell line or report the mutation described by O' Connor et al.; [9]This cell line has been reported to be null for p53 RNA or protein. Whether this is due to a small DNA rearrangement or RNA-mediated decay associated with a frameshift mutation is unknown.

were cross-contaminated.[8] The p53 status in various cell lines is a paradigm for CLCC. (i) p53 mutation is sufficiently diverse to allow comparison of various cell lines. Statuses of other genes with fewer mutation hot spots (Ha-ras) or a lower frequency of mutations are not as useful. (ii) Due to its importance in cell phenotype, p53 status has been analysed in more than 1,200 cell lines. Although p53 mutation

Table 4 **Cell lines with controversial p53 mutations**

| Cell line | ATCC number | Cancer | p53 Mutation UMD | |
|---|---|---|---|---|
| | | | AA Change | Reference* |
| HT-1197 | CRL-1473 | *Bladder carcinoma* | p.His365Arg | 66 |
| | | | wt | 2 |
| EJ | | *Bladder carcinoma* | p.Tyr126X | 67 |
| | | | p.Lys164Glu | 68 |
| | | | wt | 69 |
| RT-112 | | *Bladder carcinoma* | p.Ser183X, p.Arg248Gln | 2 |
| | | | p.Arg248Gln | 70 |
| SD | | *Bladder carcinoma* | p.Arg110Leu | 71 |
| | | | p.Ser116Cys | 67 |
| T-24 | HTB-4 | *Bladder carcinoma* | p.Tyr126X | 67 |
| | | | p.Tyr126delTAC | 66 |
| VM-CUB-1 | | *Bladder carcinoma* | p.Tyr126X, p.Arg175His | 2 |
| | | | p.Arg175His | 67 |
| VM-CUB-2 | | *Bladder carcinoma* | p.Arg158Leu, p.Tyr163Cys | 67 |
| | | | p.Arg158Leu | 71 |
| MDA-MB-436 | HTB-130 | *Breast carcinoma* | p.Glu204delinsAspfsX6 | 55 |
| | | | p.Arg273His | 72 |
| DAUDI | CCL-213 | *Burkitt lymphoma* | p.Arg213X | 73 |
| | | | p.Gly266Glu | 2 |
| A-172 | CRL-1620 | *Glioblastoma* | p.Cys242Phe | 74 |
| | | | wt | 75 |
| U-118-MG | HTB-15 | *Glioblastoma* | p.Arg213Gln | 76 |
| | | | wt | 2 |
| SK-LMS-1 | | *Leiomyosarcoma* | p.Met237Lys, p.Gly245Ser | 77 |
| | | | p.Gly245Ser | 78 |
| SK-UT-1 | HTB-114 | *Leiomyosarcoma* | p.Arg175His, p.Arg248Gln | 2 |
| | | | p.Arg175His | 78 |
| NCI-H1048 | CRL-5853 | *Lung (SCLC)* | p.Ser46fsX76, p.Arg273Cys | 2 |
| | | | p.Arg273Cys | 79 |
| MeWo | HTB-65 | *Melanoma* | p.Glu258Lys, p.Gln317X | 2 |
| | | | p.Glu258Lys | 80 |
| PA-1 | CRL-1572 | *Ovarian carcinoma* | p.Asn239Asp | 81 |
| | | | p.Pro316Pro | 57 |
| | | | wt | 2 |
| SW626 | HTB-78 | *Ovarian carcinoma* | p.Gly262Val | 58 |
| | | | p.Arg273His | 82 |
| Capan-2 | HTB-80 | *Pancreatic cancer* | p.Arg273His | 83 |
| | | | wt | 2 |

*References correspond to studies in which the p53 gene status was analysed experimentally and not deduced from other reports in the literature.

analysis cannot replace DNA fingerprinting, our finding is a strong argument to suggest that CLCC should not be ignored. We are also very concerned by the observation that the p53 status based on a cell line (either correct or false) can be reproduced from a single publication in the literature without any subsequent confirmation. Finally, we have also noticed a marked heterogeneity in the labelling of cell lines, a problem that can also lead to confusion between mislabelled cell lines with similar names.[36] CLCC includes several situations: (i) cross-contamination between two cell lines (the best example being HeLa cells); (ii) cell lines with an incorrect origin (such as the KB cell line often wrongly described as an oral cancer when it is actually a cervical cancer); and (iii) cell lines that have been contaminated during manipulation. We believe that the problem identified in the present analysis is predominantly related to confusion or incorrect labelling of cell lines. Although, the material and methods sections of published articles usually state that cell lines were derived from cell banks such as ATCC or DMSZ, it is well known that many cell lines have been exchanged between research groups, a situation that increases the probability of CLCC. These problems have already been extensively discussed over the past year, but seem to be ignored by the scientific community. We strongly encourage all scientists to comply with the various recently published guidelines for correct handling of cell lines.[37,38]

The p53 status in cell lines is now available at the p53 web site (http://p53.free.fr/Database/Cancer_cell_lines/p53_cell_lines.html). A specific section is devoted to cell lines with a controversial p53

status. We invite all scientists to update these tables with their own findings so that a consensus concerning the p53 status of each cell line can be reached. Finally, we strongly encourage those involved in studies dealing with p53 (or other p53 family members) to regularly check the p53 status of their cell lines.

## Material and Methods

**Analysis of the biological activity of mutant p53 proteins.** Data analysis. The p53 database used for this study contains 21,717 mutations derived from 1,992 publications (UMD p53 database (htpp://p53.free.fr), 2007_R1 release released in January 2007).[22] This release contains functional data for the majority of missense p53 mutants. Mutant p53 activity has been described previously.[22,23] Briefly, 2,314 haploid yeast transformants containing p53 mutations and a GFP-reporter plasmid have been constructed. Mutant p53 activity was tested by measuring the fluorescent intensity of GFP that is controlled by the p21WAF1 promoter sequence of the plasmid after 3 days of growth at 37°C. For functional analysis, frameshift and nonsense mutations were also excluded, as their biological significance has not been clearly established (see text for more information). The mean and 95% Confidence Interval (CI) of the biological activity of all mutants was calculated by using the transactivational activity measured on the p21WAF1 promoter. Similar results were obtained with the activity measured on 7 other promoters of transcription (data not shown).

Statistical analysis. To identify the distinct levels of p53 residual activities among the mutants we used the k-means clustering,[24] whose aim is to partition the data into 3 groups such that the sum of squares from each mutant to the assigned cluster centres is minimized. Three clusters were chosen to represent mutants with no, low and wild-type activity levels. The analysis was based on the measurements of promoter activities of 8 p53 target genes, including p21WAF1, MDM2, BAX, v14-3-3-σ, AIP, GADD45, NOXA and p53R2.

**Acknowledgements**

We are grateful to B. Zhivotovsky and B. Joseph for reading this manuscript. This work is supported by Cancerföreningen i Stockholm and the Swedish Research Council (VR).

**Note**

Supplementary materials can be found at:
www.landesbioscience.com/supplement/BerglindCBT7-5-Sup.pdf

**References**

1. Monks A, Scudiero D, Skehan P, Shoemaker R, Paull K, Vistica D, Hose C, Langley J, Cronise P, Vaigro Wolff A, et al. Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor cell lines. J Natl Cancer Inst 1991; 83:757-66.
2. Forbes S, Clements J, Dawson E, Bamford S, Webb T, Dogan A, Flanagan A, Teague J, Wooster R, Futreal PA, Stratton MR. COSMIC 2005. Br J Cancer 2006; 94:318-22.
3. Ikediobi ON, Davies H, Bignell G, Edkins S, Stevens C, O'Meara S, Santarius T, Avis T, Barthorpe S, Brackenbury L, Buck G, Butler A, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Hunter C, Jenkinson A, Jones D, Kosmidou V, Lugg R, Menzies A, Mironenko T, Parker A, Perry J, Raine K, Richardson D, Shepherd R, Small A, Smith R, Solomon H, Stephens P, Teague J, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Reinhold W, Weinstein JN, Stratton MR, Futreal PA, Wooster R. Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. Mol Cancer Ther 2006; 5:2606-12.
4. Nelson Rees WA, Daniels DW, Flandermeyer RR. Cross-contamination of cells in culture. Science 1981; 212:446-52.
5. Chatterjee R. Cell biology. When 60 lines don't add up. Science 2007; 315:929.
6. O'Brien, SJ. Cell culture forensics. Proc Natl Acad Sci USA 2001; 98:7656-8.
7. Markovic O, Markovic N. Cell cross-contamination in cell cultures: the silent and neglected danger. In Vitro Cell Dev Biol Anim 1998; 34:1-8.
8. MacLeod RA, Dirks WG, Matsuo Y, Kaufmann M, Milch H, Drexler HG. Widespread intraspecies cross-contamination of human tumor cell lines arising at source. Int J Cancer 1999; 83:555-63.
9. Donzelli M, Bernardi R, Negri C, Prosperi E, Padovan L, Lavialle C, Brison O, Scovassi AI. Apoptosis-prone phenotype of human colon carcinoma cells with a high level amplification of the c-myc gene. Oncogene 1999; 18:439-48.
10. Liscovitch M, Ravid D. A case study in misidentification of cancer cell lines: MCF-7/AdrR cells (re-designated NCI/ADR-RES) are derived from OVCAR-8 human ovarian carcinoma cells. Cancer Lett 2007; 245:350-2.
11. Caron de Fromentel C, Nardeux PC, Soussi T, Lavialle C, Estrade S, Carloni G, Chandrasekaran K, Cassingena R. Epithelial HBL-100 cell line derived from milk of an apparently healthy woman harbours SV40 genetic information. Exp Cell Res 1985; 160:83-94.
12. Soussi T, Beroud C. Assessing TP53 status in human tumors to evaluate clinical outcome. Nat Rev Cancer 2001; 1:233-40.
13. Oren M. Decision making by p53: life, death and cancer. Cell Death Differ 2003; 10:431-42.
14. Harris SL, Levine AJ. The p53 pathway: positive and negative feedback loops. Oncogene 2005; 24:2899-908.
15. El Deiry WS. The role of p53 in chemosensitivity and radiosensitivity. Oncogene 2003; 22:7486-95.
16. O'Connor PM, Jackman J, Bae I, Myers TG, Fan S, Muroh M, Scudiero DA, Monks A, Sausville EA, Weinstein JN, Friend S, Fornace AJJ, Kohn KW. Characterization of the p53 tumor suppressor pathway in cell lines of the National Cancer Institute anticancer drug screen and correlations with the growth-inhibitory potency of 123 anticancer agents. Cancer Res 1997; 57:4285-300.
17. Shi LM, Myers TG, Fan Y, O'Connor PM, Paull KD, Friend SH, Weinstein JN. Mining the National Cancer Institute Anticancer Drug Discovery Database: cluster analysis of ellipticine analogs with p53-inverse and central nervous system-selective patterns of activity. Mol Pharmacol 1998; 53:241-51.
18. Brachman DG, Beckett M, Graves D, Haraf D, Vokes E, Weichselbaum RR. p53 mutation does not correlate with radiosensitivity in 24 head and neck cancer cell lines. Cancer Res 1993; 53:3667-9.
19. Shiraishi K, Kato S, Han SY, Liu W, Otsuka K, Sakayori M, Ishida T, Takeda M, Kanamaru R, Ohuchi N, Ishioka C. Isolation of temperature-sensitive p53 mutations from a comprehensive missense mutation library. J Biol Chem 2004; 279:348-55.
20. Caron de Fromentel C, Soussi T. TP53 tumor suppressor gene: a model for investigating human mutagenesis. Genes Chromosomes Cancer 1992; 4:1-15.
21. Soussi T, Ishioka C, Claustres M, Beroud C. Locus-specific mutation databases: pitfalls and good practice based on the p53 experience. Nat Rev Cancer 2006; 6:83-90.
22. Soussi T, Kato S, Levy PP, Ishioka C. Reassessment of the TP53 mutation database in human disease by data mining with a library of TP53 missense mutations. Hum Mutat 2005; 25:6-17.
23. Kato S, Han SY, Liu W, Otsuka K, Shibata H, Kanamaru R, Ishioka C. Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci USA 2003; 100:8424-9.
24. Hartigan J, Wong M. *A K-means clustering algorithm*. Applied Statistics 1979; 28:100-8.
25. Wang XW, Harris CC. p53 tumor-suppressor gene: clues to molecular carcinogenesis. J Cell Physiol 1997; 173:247-55.
26. Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP, Halperin AJ, Ponten J. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc Natl Acad Sci USA 1991; 88:10124-8.
27. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. Science 1996; 274:430-2.
28. Toyooka S, Tsuda T, Gazdar AF. The TP53 gene, tobacco exposure, and lung cancer. Hum Mutat 2003; 21:229-39.
29. Soussi T, Asselain B, Hamroun D, Kato S, Ishioka C, Claustres M, Beroud C. Meta-analysis of the p53 mutation database for mutant p53 biological activity reveals a methodologic bias in mutation detection. Clin Cancer Res 2006; 12:62-9.
30. Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. Science 2002; 297:1007-13.
31. Aretz S, Uhlhaas S, Sun Y, Pagenstecher C, Mangold E, Caspari R, Moslein G, Schulmann K, Propping P, Friedl W. Familial adenomatous polyposis: aberrant splicing due to missense or silent mutations in the APC gene. Hum Mutat 2004; 24:370-80.
32. Zatkova A, Messiaen L, Vandenbroucke I, Wieser R, Fonatsch C, Krainer AR, Wimmer K. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. Hum Mutat 2004; 24:491-501.
33. Varley JM, Attwooll C, White G, McGown G, Thorncroft M, Kelsey AM, Greaves M, Boyle J, Birch JM. Characterization of germline TP53 splicing mutations and their genetic and functional analysis. Oncogene 2001; 20:2647-54.
34. Holmila R, Fouquet C, Cadranel J, Zalcman G, Soussi T. Splice mutations in the p53 gene: case report and review of the literature. Hum Mutat 2003; 21:101-2.
35. Horaitis O, Cotton RG. The challenge of documenting mutation across the genome: the human genome variation society approach. Hum Mutat 2004; 23:447-52.
36. Shimada Y. Researchers should have respect for the originator of the cell lines. Clin Cancer Res 2005; 11:4634.

37. Buehring GC, Eby EA, Eby MJ. Cell line cross-contamination: how aware are Mammalian cell culturists of the problem and how to monitor it? In Vitro Cell Dev Biol Anim 2004; 40:211-5.
38. Yoshino K, Iimura E, Saijo K, Iwase S, Fukami K, Ohno T, Obata Y, Nakamura Y. Essential role for gene profiling analysis in the authentication of human cell lines. Hum Cell 2006; 19:43-8.
39. Cheng J, Haas M. Frequent mutations in the p53 tumor suppressor gene in human leukemia T-cell lines. Mol Cell Biol 1990; 10:5502-9.
40. Wolf D, Rotter V. Major deletions in the gene encoding the p53 tumor antigen cause lack of p53 expression in HL-60 cells. Proc Natl Acad Sci USA 1985; 82:790-4.
41. Neubauer A, He M, Schmidt CA, Huhn D, Liu ET. Genetic alterations in the p53 gene in the blast crisis of chronic myelogenous leukemia: analysis by polymerase chain reaction based techniques. Leukemia 1993; 7:593-600.
42. Murai Y, Hayashi S, Takahashi H, Tsuneyama K, Takano Y. Correlation between DNA alterations and p53 and p16 protein expression in cancer cell lines. Pathol Res Pract 2005; 201:109-15.
43. Rodrigues NR, Rowan A, Smith ME, Kerr IB, Bodmer WF, Gannon JV, Lane DP. p53 mutations in colorectal cancer. Proc Natl Acad Sci USA 1990; 87:7555-9.
44. Teoh G, Tai YT, Urashima M, Shirahama S, Matsuzaki M, Chauhan D, Treon SP, Raje N, Hideshima T, Shima Y, Anderson KC. CD40 activation mediates p53-dependent cell cycle regulation in human multiple myeloma cell lines. Blood 2000; 95:1039-46.
45. Lehman TA, Bennett WP, Metcalf RA, Welsh JA, Ecker J, Modali RV, Ullrich S, Romano JW, Appella E, Testa JR, et al. p53 mutations, ras mutations, and p53-heat shock 70 protein complexes in human lung carcinoma cell lines. Cancer Res 1991; 51:4090-6.
46. Mitsudomi T, Steinberg SM, Nau MM, Carbone D, D'Amico D, Bodner S, Oie HK, Linnoila RI, Mulshine JL, Minna JD, et al. p53 gene mutations in non-small-cell lung cancer cell lines and their correlation with the presence of ras mutations and clinical features. Oncogene 1992; 7:171-80.
47. Takahashi T, Nau MM, Chiba I, Birrer MJ, Rosenberg RK, Vinocour M, Levitt M, Pass H, Gazdar AF, Minna JD. p53: a frequent target for genetic abnormalities in lung cancer. Science 1989; 246:491-4.
48. Kastrinakis WV, Ramchurren N, Rieger KM, Hess DT, Loda M, Steele G, Summerhayes IC. Increased incidence of p53 mutations is associated with hepatic metastasis in colorectal neoplastic progression. Oncogene 1995; 11:647-52.
49. Gayet J, Zhou XP, Duval A, Rolland S, Hoang JM, Cottu P, Hamelin R. Extensive characterization of genetic alterations in a series of human colorectal cancer cell lines. Oncogene 2001; 20:5025-32.
50. Liu Y, Bodmer WF. Analysis of P53 mutations and their expression in 56 colorectal cancer cell lines. Proc Natl Acad Sci USA 2006; 103:976-81.
51. Nigro JM, Baker SJ, Preisinger AC, Jessup JM, Hostetter R, Cleary K, Bigner SH, Davidson N, Baylin S, Devilee P, et al. Mutations in the p53 gene occur in diverse human tumor types. Nature 1989; 342:705-8.
52. Chen P, Iavarone A, Fick J, Edwards M, Prados M, Israel MA. Constitutional p53 mutations associated with brain tumors in young adults. Cancer Genet Cytogenet 1995; 82:106-15.
53. Fujiwara T, Mukhopadhyay T, Cai DW, Morris DK, Roth JA, Grimm EA. Retroviral-mediated transduction of p53 gene increases TGF-beta expression in a human glioblastoma cell line. Int J Cancer 1994; 56:834-9.
54. Haapajarvi T, Kivinen L, Heiskanen A, des Bordes C, Datto MB, Wang XF, Laiho M. UV radiation is a transcriptional inducer of p21(Cip1/Waf1) cyclin-kinase inhibitor in a p53-independent manner. Exp Cell Res 1999; 248:272-9.
55. Concin N, Zeillinger C, Tong D, Stimpfl M, Konig M, Printz D, Stonek F, Schneeberger C, Hefler L, Kainz C, Leodolter S, Haas OA, Zeillinger R. Comparison of p53 mutational status with mRNA and protein expression in a panel of 24 human breast carcinoma cell lines. Breast Cancer Res Treat 2003; 79:37-46.
56. Perego P, Giarola M, Righetti SC, Supino R, Caserini C, Delia D, Pierotti MA, Miyashita T, Reed JC, Zunino F. Association between cisplatin resistance and mutation of p53 gene and reduced bax expression in ovarian carcinoma cell systems. Cancer Res 1996; 56:556-62.
57. Yaginuma Y, Westphal H. Abnormal structure and expression of the p53 gene in human ovarian carcinoma cell lines. Cancer Res 1992; 52:4196-9.
58. De Feudis P, Debernardis D, Beccaglia P, Valenti M, Graniela Sire E, Arzani D, Stanzione S, Parodi S, D'Incalci M, Russo P, Broggini M. DDP-induced cytotoxicity is not influenced by p53 in nine human ovarian cancer cell lines with different p53 status. Br J Cancer 1997; 76:474-9.
59. Ogretmen B, Safa AR. Expression of the mutated p53 tumor suppressor protein and its molecular and biochemical characterization in multidrug resistant MCF-7/Adr human breast cancer cells. Oncogene 1997; 14:499-506.
60. Isaacs WB, Carter BS, Ewing CM. Wild-type p53 suppresses growth of human prostate cancer cells containing mutant p53 alleles. Cancer Res 1991; 51:4716-20.
61. Gurova KV, Rokhlin OW, Budanov AV, Burdelya LG, Chumakov PM, Cohen MB, Gudkov AV. Cooperation of two mutant p53 alleles contributes to Fas resistance of prostate carcinoma cells. Cancer Res 2003; 63:2905-12.
62. Kashii T, Mizushima Y, Monno S, Nakagawa K, Kobayashi M. Gene analysis of K-, H-ras, p53, and retinoblastoma susceptibility genes in human lung cancer cell lines by the polymerase chain reaction/single-strand conformation polymorphism method. J Cancer Res Clin Oncol 1994; 120:143-8.
63. Bartek J, Iggo R, Gannon J, Lane DP. Genetic and immunochemical analysis of mutant p53 in human breast cancer cell lines. Oncogene 1990; 5:893-9.
64. Kovach JS, McGovern RM, Cassady JD, Swanson SK, Wold LE, Vogelstein B, Sommer SS. Direct sequencing from touch preparations of human carcinomas: analysis of p53 mutations in breast carcinomas. J Natl Cancer Inst 1991; 83:1004-9.
65. den Dunnen JT, Antonarakis SE. Nomenclature for the description of human sequence variations. Hum Genet 2001; 109:121-4.
66. Cooper MJ, Haluschak JJ, Johnson D, Schwartz S, Morrison LJ, Lippa M, Hatzivassiliou G, Tan J. p53 mutations in bladder carcinoma cell lines. Oncol Res 1994; 6:569-79.
67. Grimm MO, Jurgens B, Schulz WA, Decken K, Makri D, Schmitz Drager BJ. Inactivation of tumor suppressor genes and deregulation of the c-myc gene in urothelial cancer cell lines. Urol Res 1995; 23:293-300.
68. Rieger KM, Little AF, Swart JM, Kastrinakis WV, Fitzgerald JM, Hess DT, Libertino JA, Summerhayes IC. Human bladder carcinoma cell lines as indicators of oncogenic change relevant to urothelial neoplastic progression. Br J Cancer 1995; 72:683-90.
69. Sharma S, Schwarte Waldhoff I, Oberhuber H, Schafer R. Functional interaction of wild-type and mutant p53 transfected into human tumor cell lines carrying activated ras genes. Cell Growth Differ 1993; 4:861-9.
70. Warenius HM, Jones M, Gorman T, McLeish R, Seabra L, Barraclough R, Rudland P. Combined RAF1 protein expression and p53 mutational status provides a strong predictor of cellular radiosensitivity. Br J Cancer 2000; 83:1084-95.
71. Williamson MP, Elder PA, Knowles MA. The spectrum of TP53 mutations in bladder carcinoma. Genes Chromosomes Cancer 1994; 9:108-18.
72. Takahashi M. [Analyses of p53 mutations in breast cancers with a combined use of yeast functional assay and immunohistochemical staining]. Hokkaido Igaku Zasshi 1998; 73:275-86.
73. Gaidano G, Ballerini P, Gong JZ, Inghirami G, Neri A, Newcomb EW, Magrath IT, Knowles DM, Dalla Favera R. p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proc Natl Acad Sci USA 1991; 88:5413-7.
74. Gomez Manzano C, Fueyo J, Kyritsis AP, Steck PA, Roth JA, McDonnell TJ, Steck KD, Levin VA, Yung WK. Adenovirus-mediated transfer of the p53 gene produces rapid and generalized death of human glioma cells via apoptosis. Cancer Res 1996; 56:694-9.
75. Jia LQ, Osada M, Ishioka C, Gamo M, Ikawa S, Suzuki T, Shimodaira H, Niitani T, Kudo T, Akiyama M, Kimura N, Matsuo M, Mizusawa H, Tanaka N, Koyama H, Namba M, Kanamaru R, Kuroki T. Screening the p53 status of human cell lines using a yeast functional assay. Mol Carcinog 1997; 19:243-53.
76. Russell SJ, Ye YW, Waber PG, Shuford M, Schold SCJ, Nisen PD. p53 mutations, O6-alkylguanine DNA alkyltransferase activity, and sensitivity to procarbazine in human brain tumors. Cancer 1995; 75:1339-42.
77. Smardova J, Pavlova S, Svitakova M, Grochova D, Ravcukova B. Analysis of p53 status in human cell lines using a functional assay in yeast: detection of new non-sense p53 mutation in codon 124. Oncol Rep 2005; 14:901-7.
78. Stratton MR, Moss S, Warren W, Patterson H, Clark J, Fisher C, Fletcher CD, Ball A, Thomas M, Gusterson BA, et al. Mutation of the p53 gene in human soft tissue sarcomas: association with abnormalities of the RB1 gene. Oncogene 1990; 5:1297-301.
79. D'Amico D, Carbone D, Mitsudomi T, Nau M, Fedorko J, Russell E, Johnson B, Buchhagen D, Bodner S, Phelps R, et al. High frequency of somatically acquired p53 mutations in small-cell lung cancer cell lines and tumors. Oncogene 1992; 7:339-46.
80. Zolzer F, Hillebrandt S, Streffer C. Radiation induced $G_1$-block and p53 status in six human cell lines. Radiother Oncol 1995; 37:20-8.
81. Mihara K, Miyazaki M, Kondo T, Fushimi K, Tsuji T, Inoue Y, Fukaya K, Ishioka C, Namba M. Yeast functional assay of the p53 gene status in human cell lines maintained in our laboratory. Acta Med Okayama 1997; 51:261-5.
82. Gelfi C, Righetti SC, Zunino F, Della Torre G, Pierotti MA, Righetti PG. Detection of p53 point mutations by double-gradient, denaturing gradient gel electrophoresis. Electrophoresis 1997; 18:2921-7.
83. Ruggeri B, Zhang SY, Caamano J, DiRado M, Flynn SD, Klein Szanto AJ. Human pancreatic carcinomas and cell lines reveal frequent and multiple alterations in the p53 and Rb-1 tumor-suppressor genes. Oncogene 1992; 7:1503-11.
84. Rodicker F, Putzer BM. p73 is effective in p53-null pancreatic cancer cells resistant to wild-type TP53 gene replacement. Cancer Res 2003; 63:2737-41.

*Int. J. Cancer:* **123,** 117–122 (2008)


# TNF-α-inducing protein, a carcinogenic factor secreted from *H. pylori*, enters gastric cancer cells

**Masami Suganuma[1]*, Kensei Yamaguchi[2], Yoshie Ono[2], Haruo Matsumoto[2], Tomonori Hayashi[3], Takahiko Ogawa[3], Kazue Imai[3], Takashi Kuzuhara[4], Akira Nishizono[5] and Hirota Fujiki[4]**

[1]*Research Institute for Clinical Oncology, Saitama Cancer Center, Saitama, Japan*
[2]*Saitama Cancer Center, Saitama, Japan*
[3]*Radiation Effects Research Foundation, Hiroshima, Japan*
[4]*Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Tokushima, Japan*
[5]*Department of Infectious Disease, Faculty of Medicine, Oita University, Oita, Japan*

**TNF-α inducing protein (Tipα) is secreted from *Helicobacter pylori* (*H. pylori*): it is a potent inducer of *TNF-α* and *chemokine* genes, mediated through NF-κB activation, and it also induces tumor-promoting activity in Bhas 42 cells. To investigate the carcinogenic mechanisms of *H. pylori* with Tipα, we first examined how Tipα acts on gastric epithelial cells. We found that fluorescent-Tipα specifically bound to, and then entered, the cells in a dose- and temperature-dependent manner, whereas deletion mutant of Tipα (del-Tipα), an inactive form, neither bound to nor entered the cells, suggesting the presence of a specific binding molecule. Mutagenesis analysis of Tipα revealed that a dimer formation of Tipα with a disulfide bond is required for both specific binding and induction of *TNF-α* gene expression. A confocal laser scanning microscope revealed some Tipα in the nuclei, but del-Tipα was not present, which indicated that an active form of Tipα can penetrate the nucleus and may be involved in the induction of *TNF-α* gene expression. Examination of Tipα production and secretion in 28 clinical isolates revealed that *H. pylori* obtained from gastric cancer patients secreted Tipα in significantly higher amounts than did *H. pylori* from patients with chronic gastritis, suggesting that Tipα is an essential factor in *H. pylori* inflammation and cancer microenvironment in the human stomach. Tipα is thus a new carcinogenic factor of *H. pylori* that can enter the nucleus through a specific binding molecule, and its mechanism of action is completely different from that of CagA.**


**Key words:** Tipα; gastric cancer; tumor promoter; NF-κB; TNF-α

*Helicobacter pylori* has been identified as a causative agent of chronic inflammation, chronic gastritis and peptic ulcer, and is also classified by IARC as a definitive carcinogen for gastric cancer.[1] It is well accepted that persistent *H. pylori* infection results in an inflammatory response in the stomach by high induction of proinflammatory cytokines, such as tumor necrosis factor-α (TNF-α), interleukin-1 (IL-1) and IL-8.[2] Based on our previously reported evidence that TNF-α is one of the essential cytokines for tumor promotion and an instigator of a cytokine network sequence of tumor promotion from TNF-α through IL-1 to IL-6 and back to TNF-α, we think a gene product of *H. pylori* which induces TNF-α plays a significant role in gastric cancer development in humans[3–5]: We thus cloned *TNF-α inducing protein* (*Tipα*) gene (*HP0596*) from genome sequence of *H. pylori* strain 26695. Recombinant Tipα protein strongly induces expression of *TNF-α* and various *chemokine* genes, activates NF-κB in mouse gastric epithelial cells MGT-40 and showed *in vitro* transforming activity of v-H-*ras* transfected BALB/3T3 (Bhas 42) cells, a standard model of initiated cells.[6–9] *Tipα* gene is unique for *H. pylori* genome, which does not have any obvious homologue in other species: only in *H. pylori* strain was a homologue of Tipα found as *H. pylori*-membrane protein 1 (*HP-MP1*) gene from strain SR7791 (94.3% homology) and *jph0543* from strain J99 (95.5%).[6,10–13] Thus, the *Tipα* gene family codes new carcinogenic factors of *H. pylori*.

Tipα protein with a molecular weight of 19 kDa is secreted from *H. pylori* as a homodimer form with 38 kDa. The homodimer form of Tipα is active in the induction of *TNF-α* gene expression, NF-κB activation in gastric epithelical cells and transforming activity in Bhas 42 cells.[6] Tipα has several unique features: (*i*) no similarity to other virulence factors of *H. pylori*, such as vacuolating cytotoxin, immunodominant cytotoxin-associated antigen (CagA) or urease, (*ii*) secretion from *H. pylori* in a manner independent of Type IV secretion system, (*iii*) induction of NF-κB activation in a cag pathogenicity island (cagPAI)-independent manner, (*iv*) the presence of transforming activity by treatment with a recombinant protein *in vitro*. Furthermore, we recently found that Tipα has DNA binding activity using surface plasmon resonance assay (Biacore).[14] All the results were based on the evidence that a recombinant deletion mutant of Tipα (rdel-Tipα), which lacks 6 amino acids including 2 cysteine residues in N-terminal position, was inactive. Therefore, we think that identifying the molecular mechanisms of Tipα will provide a key to understanding the carcinogenic mechanisms of *H. pylori* infection.

Using fluorescence-labeled rTipα by flow cytometry and immunocytochemical analysis, we first examined how exogenously added recombinant Tipα (rTipα) induces expression of *TNF-α* and *chemokine* genes on gastric epithelial cells. Here we report that rTipα bound to cell surface molecules specifically, and then penetrated the cytosol; some rTipα was found localized in the nuclei of gastric cancer cells MGT-40, but rdel-Tipα, an inactive mutant, showed less binding activity and negligible penetration into cytosol and nucleus.

To further investigate the significance of Tipα in gastric cancer development, we next examined Tipα production and secretion in 28 clinical isolates of *H. pylori* from patients with chronic gastritis and gastric cancer in Saitama Cancer Center, Japan. Although all isolates were positive for Tipα production, we found that *H. pylori* isolates obtained from gastric cancer patients secreted significantly larger amounts of Tipα into culture broth compared with those from chronic gastritis patients. Therefore, we think that Tipα secreted from *H. pylori* plays an essential role in inflammation and cancer development in humans.

## Material and methods

### *Preparation of recombinant Tipα and mutant proteins*

Recombinant His-tagged proteins of Tipα (rTipα), del-Tipα (rdel-Tipα), cysteine substituted with alanine mutants—Cys5Ala (C5A), Cys7Ala (C7A), Cys5Ala/Cys7Ala (C5AC7A)—were expressed in *E. coli* transfected pET28(a)+ containing each corresponding gene induced with IPTG, and by $Ni^{2+}$-loaded Hitrap

Grant sponsors: Japan Society for the Promotion of Science; Smoking Research Fund.
***Correspondence to:** Research Institute of Clinical Oncology, Saitama Cancer Center, Ina, Kitaadachi-gun. Saitama 362-0806, Japan. Fax: +81-48-722-1739. E-mail: masami@cancer-c.pref.saitama.jp
Received 19 December 2007; Accepted after revision 18 January 2008
DOI 10.1002/ijc.23484
Published online 15 April 2008 in Wiley InterScience (www.interscience.wiley.com).

Publication of the International Union Against Cancer
global cancer control

chelating column (Amersham Bioscience, Buckinghamshire, UK).[6] For generation of cysteine substituted with alanine mutants of Tipα, sense primers were as follows: C5A, 5′-CAGCCATATGCTGCAGGCTGCCACTTGCCCAAAC-3′; C7A, 5′-CAGCCATATGCTGCAGGCTTGCACTGCCCCAAACAC-3′; C5A/C7A, 5′-CAGCCATATGCTGCAGGCTGCCACTGCCCCAAACAC-3′, and anti-sense primer was 5′-TCTCGGATCCTACATGGCTATAGGGACTTT-3′. Purities of the 5 recombinant proteins were more than 98% on SDS-PAGE.

*Cell culture and reagents*

Mouse gastric cancer cell line MGT40 was kindly provided by Dr. Masae Tatematsu, Aichi Cancer Center Research Institute, Aichi, Japan. MGT-40 cells were maintained in DMEM with 10% fetal bovine serum (JRH Biosciences, KS) and MITO+ serum extender (Becton Dickinson Labware, MA), as described previously.[15] Anti-Tipα antibody was obtained by immunization of rabbit with 19-mer synthetic peptide, also as described previously.[6] Anti-HSP90, anti-EGF receptor and anti-lamin B antibodies were purchased from Santa Cruz (CA).

*H. pylori strains and culture conditions*

Twenty-eight *H. pylori* clinical isolates were obtained from patients with chronic gastritis (11 patients) and gastric cancer (17 patients), using M-BHM *H. pylori* selection agar plates (Nissui Pharmaceutical Co., Tokyo, Japan) at Saitama Cancer Center between 2001 and 2006. *H. pylori* strain 26695 were kindly provided by Dr. Chihiro Sasakawa (The Institute of Medical Science University of Tokyo, Tokyo Japan), and ATCC43504 was purchased from American Type Culture Collection (VA). All *H. pylori* isolates were cultured in Brucella broth (Becton Dickinson Microbiology Systems, MD) containing 10% horse serum (Nippon Bio-Test Lab, Tokyo, Japan) at 37°C, with shaking in microaerobic conditions, as described previously.[6,16]

*Western blotting*

After 3 days culture of *H. pylori* strains, culture broths with volumes (7.5–15 μl) equal to 1.0 $OD_{540}$ of bacteria amount were separated on 12% SDS-polyacrylamide gel in the presence of dithiothreitol (DTT), and blotted onto nitrocellulose membranes. Tipα protein was visualized by the ECL detection system (Amersham Bioscience) using anti-Tipα antibody.[6] Amounts of Tipα were measured by density of a band of 19 kDa (Tipα) by NIH Image, and expressed as a relative unit based on Tipα amount in culture broth of 26695 strain. The results were expressed as the average of 2 separate experiments.

*Expression of TNF-α gene*

MGT-40 cells were treated with recombinant proteins for 1 hr, and total RNA was isolated with ISOGEN reagent (Nippon Gene, Tokyo, Japan). Expression of *TNF-α* gene was determined by semiquantitative RT-PCR, as described previously.[6]

*Flow cytometry*

rTipα was labeled with FITC using EZ-label™ FITC protein labeling kit (Pierce Biotechnology, IL) according to the manufacturer's instructions. MGT-40 cells ($1 \times 10^6$ cells/ml) in PBS were incubated with various concentrations of FITC-rTipα at 4°C for 2 hr. Cells were then washed with PBS and analyzed by flow cytometry (EpicsXL, Beckman Coulter, Tokyo, Japan). FITC-labeled BSA was used as negative control. Equal absorbance of FITC-rTipα or FITC-BSA at 495 nm (1 mol FITC/ mol protein) was used in the experiments. Binding of FITC-rTipα to MGT-40 cells was measured as mean fluorescence intensity (FI).[17] Binding of FITC-rTipα (2.5 μM) to MGT-40 cells was also analyzed in the presence of nonlabeled rTipα, rdel-Tipα, C5A, C7A or C5AC7A mutant at various concentrations.

*Cytochemical analysis*

rTipα, rdel-Tipα and BSA were labeled with Alexa Fluor 488 protein labeling kit (Invitrogen, Tokyo, Japan) according to the manufacturer's instructions. MGT-40 cells were incubated with Alexa Fluor 488-labeled rTipα (AF488-Tipα), rdel-Tipα (AF488-del-Tipα) or BSA (AF488-BSA) at a concentration of 5.0 μM, with equal absorbance at 495 nm for 1 hr. After fixation with 4% paraformaldehyde containing 0.2% Triton X-100, the cells were observed using fluorescence microscope (Leica Microbiosystems, Tokyo, Japan).

*Immunocytochemical analysis*

MGT-40 cells were incubated with rTipα (100 μg/ml) for 1 hr at 37°C, and then washed with PBS. The cells were fixed with 4% paraformaldehyde containing 0.2% Triton X-100. After blocking with Block Ace (Dainippon Pharm., Osaka, Japan), the cells were treated with anti-Tipα antibody, and stained with Alexa Fluor 488-conjugated anti-rabbit IgG (Invitrogen, Tokyo, Japan). Then the cells were stained with propidium iodide and analyzed using confocal laser scanning microscope LSM 5 PASCAL (Carl Zeiss, Germany).

*Subcellular fraction analysis*

MGT-40 cells were incubated with rTipα or rdel-Tipα (5 μM) for 1 hr at 37°C and washed with PBS, and the cells were fractionated into cytosol, membrane and nuclei using Qproteome cell compartment kit (Qiagen, Düesseldorf, Germany) according to the manufacturer's instructions.[18] Seven micrograms of each fraction was subjected to Western blotting, and analyzed to detect rTipα and rdel-Tipα proteins by anti-Tipα antibody, and heat shock protein 90 (HSP90; a marker for cytosol), epidermal growth factor receptor (EGFR; for membrane) and lamin B (for nucleus) as the controls for cell fractionation by corresponding antibodies.

*Statistical analysis*

The differences in secreted amounts of Tipα from *H. pylori* obtained from patients with gastritis and gastric cancer were analyzed by Mann–Whitney analysis, with exact $p$ value using SPSS V14.0 (SPSS, Chicago, IL).

## Results

*Specific binding of FITC-labeled Tipα protein to MGT-40 cells*

We first examined specific binding of FITC-labeled rTipα to MGT-40 cells using flow cytometry: Incubation of MGT-40 cells with FITC-rTipα protein at 4°C for 2 hr significantly increased cellular fluorescence dose-dependently (Fig. 1*a*). Quantitative analysis of "mean cellular FI") indicated that binding of FITC-rTipα to MGT-40 cells was significant and saturated at concentrations of 5.0–7.5 μM. However, FI in the case of FITC-BSA was significantly low and not saturated up to 7.5 μM (Fig. 1*a*). Furthermore, the binding of FITC-rTipα to MGT-40 cells was dose-dependently inhibited with nonlabeled rTipα, as shown in Figure 1*b*. However, rdel-Tipα, an inactive protein lacking 6 amino acids, including 2 cysteine residues, showed less inhibitory activity. Concentration of 50% inhibition ($IC_{50}$) values are 1.9 μM for rTipα and 20.0 μM for rdel-Tipα. The inhibition of specific binding with nonlabeled rTipα was about 10 times stronger than with rdel-Tipα. These results correlated closely with induction of *TNF-α* gene, as well as biological activities such as induction of NF-κB activation and *chemokine* gene expressions in MGT-40 cells, and *in vitro* transformation of Bhas 42 cells, as reported previously.[6]

To clarify the significance of cysteine residues more precisely, we made 3 mutants: One cysteine substituted with an alanine at 5 or 7 position (C5A or C7A mutant), and two cysteines substituted with 2 alanines at 5 and 7 positions at the same time (C5A/C7A double mutant) (Fig. 2*a*). Each recombinant protein was analyzed by SDS-PAGE in the absence of DTT. C5A and C7A proteins formed a dimer similar to Tipα, since 1 cysteine residue still

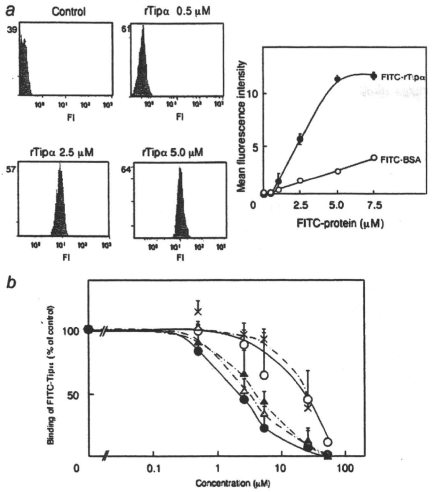


FIGURE 1 – Binding of FITC-rTipα to MGT-40 cells analyzed by flow cytometry. (*a*) MGT-40 cells were incubated with various concentrations of FITC-rTipα for 2 hr at 4°C. Fluorescence intensity (FI) was detected by Flow cytometer, as described in Material and Methods. FITC-BSA was used as a control. (*b*) MGT-40 cells were incubated with FITC-rTipα (2.5 μM) in the presence of nonlabeled rTipα (●), C5A (△), C7A (▲), C5A/C7A (X) double mutant or rdel-Tipα (○). Fluorescence intensity in the absence of nonlabeled protein is expressed as 100%. The results are the average of 2 independent experiments. Bars show SD.

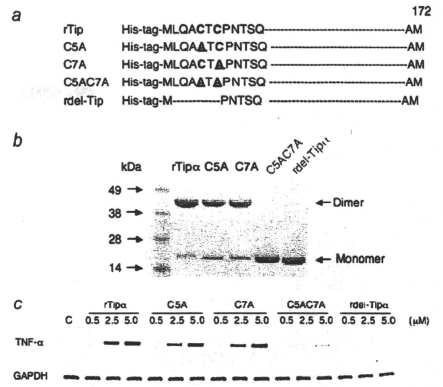


FIGURE 2 – Significance of a cysteine residue for homodimer formation and *TNF*-α gene induction. (*a*) Schematic representation of Tipα, cysteine-substituted-mutants and rdel-Tipα. (*b*) SDS-PADE analysis of recombinant proteins in the absence of DTT. rTipα, C5A and C7A formed homodimers, but C5A/C7A double mutant and rdel-Tipα did not. (*c*) Induction of *TNF*-α gene expression in MGT-40 cells. Expression of *TNF*-α gene was examined by semiquantitative RT-PCR, as described in Material and Methods.

remained, and C5A/C7A double mutant protein showed only a monomer band with a molecular weight of 21 kDa, with 2 cysteines entirely replaced (Fig. 2*b*). C5A and C7A mutant proteins dose-dependently inhibited the binding of FITC-rTipα to MGT-40 cells similar to rTipα, and C5A/C7A double-mutant protein showed weak inhibitor activity (Fig. 1*b*). $IC_{50}$ values were 2.4 μM for C5A, 2.9 μM for C7A and 21.0 μM for C5A/C7A. Their binding activities to MGT-40 cells were identical: treatment with C5A and C7A proteins strongly induced *TNF*-α gene expression, but C5A/C7A protein did not. These results indicated that homodimer formation with a disulfide bond through a cysteine residue is essential for Tipα binding to MGT-40 cells as well as TNF-α induction and carcinogenic activities (Fig. 2*c*).

The results suggest the presence of a specific binding molecule which can recognize the homodimer form of Tipα on the cell surface of MGT-40 cells. Specific binding of FITC-rTipα was also confirmed in human gastric cancer cell lines (MKN-1, MKN-45 and MKN-74) by flow cytometry (data not shown), and determination of the specific binding molecule(s) is now under investigation. When incubation of the cells with FITC-rTipα was conducted at 37°C, binding of Tipα was lower than that incubated at 4°C, and not saturated, although bindings of FITC-BSA were the same in both cases. The results led us to believe that rTipα internalizes into MGT-40 cells in a temperature-dependent manner.

### *Incorporation of Tipα into MGT-40 cells*

As the fluorescence of FITC is quite unstable for fluorescence microscope analysis, rTipα was labeled with a photostable dye, Alexa Fluor 488. After incubation of MGT-40 cells with Alexa Fluor 488-conjugated Tipα (AF488-Tipα) at 37°C, significant fluorescence was observed in the cells after 10-min incubation, it increased time-dependently until 30 min, and was sustained until 2 hr. Figure 3*a* shows strong fluorescent spots in MGT-40 cells after 1-hr incubation with AF488-rTipα at 37°C. However, the cells incubated with AF488-rdel-Tipα showed only slight fluorescence similar to those incubated with AF488-BSA, although the same "fluorescence intensities" of AF488-labeled proteins were used in these experiments. Fluorescence in the cells incubated with AF488-Tipα increased in dose-dependent and time-dependent manners. All the results show that rTipα significantly incorporated into the cells, while the inactive mutant protein rdel-Tipα did not.

To confirm incorporation of rTipα protein into the cells, we next conducted immunocytochemical analysis with anti-Tipα antibody, and observed the results under a confocal laser scanning microscope. Anti-Tipα antibody clearly recognized Tipα protein in the cytosol of the cells after treatment with Tipα for 1 hr (Fig. 3*b*), but nonimmunized rabbit IgG did not show any significant fluorescence in rTipα-treated cells. Nontreated cells were negatively stained. All results indicate that Tipα enters the gastric epithelial cells, resulting in expression of *TNF*-α and *chemokine* genes.

### *Localization of rTipα in nucleus*

Confocal laser scanning microscope analysis revealed that some fluorescent spots that had reacted with fluorescent anti-Tipα antibody were located in the nuclei of MGT-40 cells treated with rTipα, revealing the internalization of Tipα into nuclei (Fig. 4*a*). To confirm localization of rTipα in nuclei, we next conducted cell fractionation into cytosol, membrane and nuclei, after treatment with rTipα for 1 hr at 37°C. As shown in Figure 4*b*, rTipα was present in each cell fraction compared with marker proteins of each separated fraction: HSP90 for cytosol, EGFR for membrane and lamin B for nucleus. Based on the density of protein bands, we calculated that about 10% of Tipα had localized in the nuclei of the cells. However, while small amounts of rdel-Tipα, an inactive mutant, were detected in cytosol and membrane fractions—

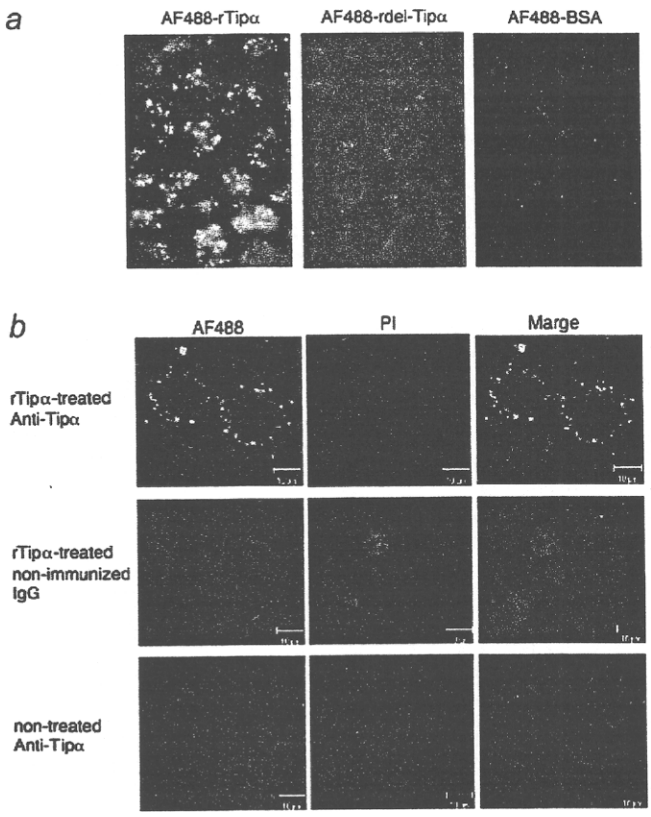


FIGURE 3 – Incorporation of Alexa Fluoro 488 (AF488)-labeled-rTipα into MGT-40 cells. (*a*) After incubation with AF488-rTipα, AF488-rdel-Tipα or AF488-BSA at a concentration of 5.0 μM for 1 hr at 37°C, MGT-40 cells were washed with PBS and then observed using fluorescence micoroscope. (*b*) Existence of rTipα in the cytosol was confirmed with confocal laser scanning microscope after immunocytochemical staining. MGT-40 cells were incubated with rTipα (5.0 μM) for 1 hr at 37°C, then subjected to anti-Tipα antibody or nonimmunized rabbit IgG (for a negative control) followed by AF488-labeled anti-rabbit IgG, as described in Material and Methods. As control, nontreated cells were subjected to anti-Tipα antibody. Nuclei were stained with propidium iodide (PI).

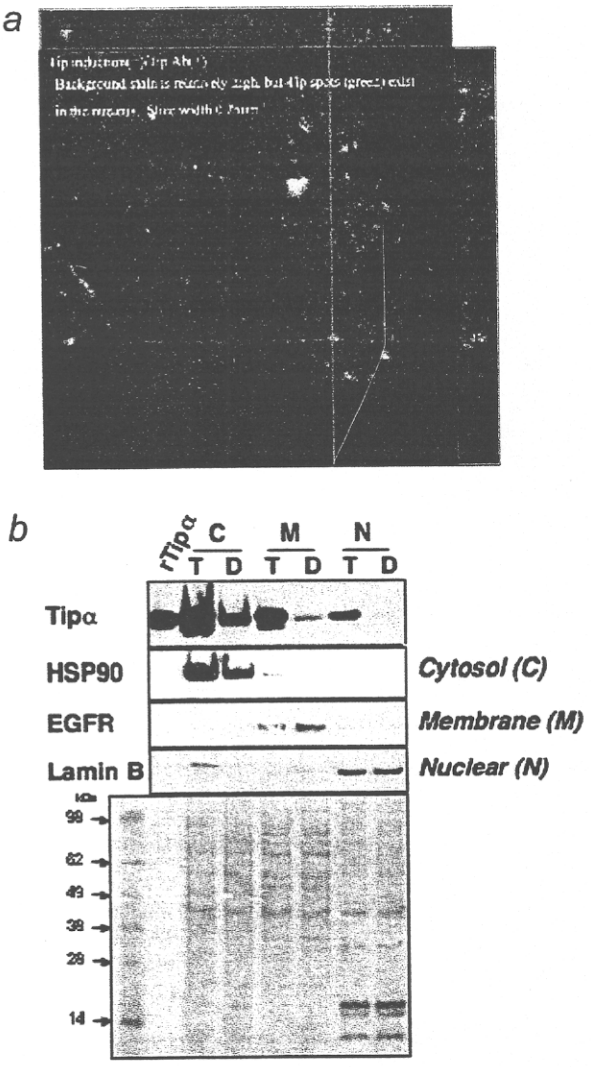


FIGURE 4 – Nuclear localization of rTipα. (*a*) Some yellow fluorescent spots were observed in the nuclei using confocal laser scanning microscope. (*b*) After cellular fractionation of rTipα or rdel-Tipα-treated cells into the cytosol, membrane and nuclei, and each fraction was analyzed by Western blotting with anti-Tipα antibody. Each fraction was confirmed by detection of specific marker proteins—HSP90 for cytosol, EGFR for membrane and lamin B for nuclei—by anti-HSP90, anti-EGFR and anti-lamin B antibodies, respectively. The gel stained with quick CBB is also indicated.

about 13.5–17.5% of rTipα—none in the nuclei fraction was found. These results strongly supported the results by confocal laser scanning microscope. Recently, we found that Tipα directly bound to DNA, but del-Tipα did not.[14] We therefore think that Tipα in the nuclei plays an essential role in the induction of *TNF-α* and *chemokine* genes and carcinogenic activity. This is the first report that a protein secreted from *H. pylori* enters the nuclei, and these results will provide a deep insight in understanding the mechanisms of gastric cancer development by *H. pylori* infection.

*Large amounts of Tipα protein secreted from H. pylori from gastric cancer patients*

To further understand the significance of Tipα in human gastric cancer development, we examined differences of production and secretion of Tipα among various *H. pylori* clinical isolates from biopsy samples of gastric mucosa obtained from patients with chronic gastritis and gastric cancer. Twenty-eight *H. pylori* isolates, from 17 gastric cancer patients and 11 chronic gastritis patients, were cultured in Brucella, broth containing 10% horse serum at 37°C in microaerobic condition for 3 days after inoculation of $5 \times 10^8$ CFU. All isolates produced CagA, which coincided with previous results in which *H. pylori* isolated from Japanese patients were more than 98%-positive cagPAI.[19] All *H. pylori* isolates produced Tipα protein, and Tipα in bacterial extracts was consistently present (data not shown). In contrast, the amounts of Tipα in culture broth varied, and clinical isolates obtained from gastric cancer patients secreted Tipα protein significantly more than did those from gastritis patients (Figs. 5*a* and 5*b*).

*H. pylori* 26695 strain, from which genome *Tipα* gene was cloned, secreted about 1.0 ng of Tipα/$10^9$ CFU/ml, and thus was expressed as 1 relative unit. Clinical isolates from cancer patients secreted Tipα protein at 1.4–13.4 relative units, and those from gastritis patients secreted at 0.8–6.7 relative units (Fig. 5*b*). Interestingly, *H. pylori* isolated from 3 of 11 gastritis patients who developed gastric cancer also secreted larger amounts of Tipα, similar to those from cancer patients. Their median values are 2.0 for *H. pylori* from 8 patients with gastritis, 3.8 from 3 patients who later developed cancer and 4.5 from 17 patients with gastric cancer. The difference of secreted amounts of Tipα between 17 gastric cancer patients and 8 chronic gastritis patients (excluding those that developed cancer) is statistically significant ($p = 0.004$), although the difference was not so great. These results

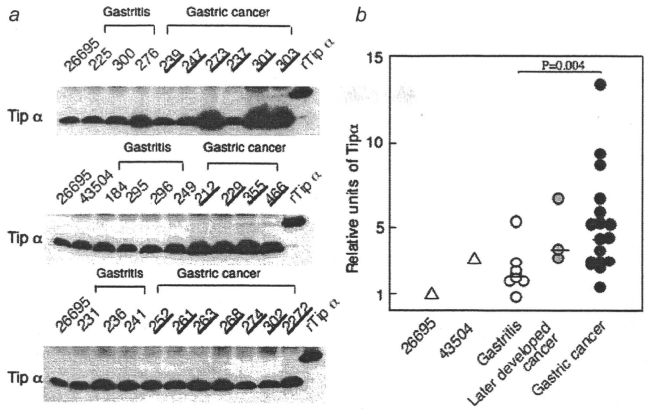


FIGURE 5 – Large amounts of Tipα were secreted from *H. pylori* clinical isolates obtained from gastric cancer patients. (*a*) Eleven clinical isolates from chronic gastritis patients, including 3 patients, who later developed gastric cancer, and 17 isolates from cancer patients were cultured for 3 days in microaerobic conditions, after inoculation of 0.5 McFalrand in 8-ml Brucella broth with 10% horse serum. Aliquots of culture broth equivalent to 1.0 $OD_{540}$ were analyzed by Western blotting with anti-Tipα antibody, as described in Material and Methods. (*b*) Relative units of Tipα in culture broth were measured for intensity of bands by NIH image, with Tipα in culture broth of 26695 strain used as a control, and expressed as 1 relative unit. Each dot corresponds to Tipα in culture broth from each clinical isolate (○ from chronic gastritis patients, gray circle from patients later developed gastric cancer, ● from gastric cancer patients). Bars indicate the median value of each group. Difference between gastritis patients (excluding those that developed cancer) and gastric cancer patients was statistically significant ($p = 0.004$).

indicate that secreted Tipα plays a significant role in gastric cancer micorenvironment during *H. pylori* infection in humans.

## Discussion

This manuscript demonstrated that Tipα protein is a new protein secreted from *H. pylori*: It can enter the gastric epithelial cells, where it then localizes in the nuclei of the cells. As for the specific function of Tipα protein in the nucleus, we recently reported that Tipα protein has direct DNA binding activity as determined by surface plasmon resonance (Biacore) assay, and that a homodimer of Tipα bound to DNA oligomer more strongly than did a monomer of del-Tipα.[14] Homodimer formation of Tipα is thus necessary for penetration into nucleus and also for the DNA binding activity. Interestingly, a dimer of Tipα dose-dependently bound to 9 bases of DNA oligomers in TNF-α promoter sequence. Pretreatment with *N*-acetylcysteine significantly inhibited penetration of Tipα into MGT-40 cells and induction of *TNF*-α gene expression (unpublished results). Therefore, we think that Tipα protein in nucleus regulates expression of *TNF*-α and *IL-6* genes as well as *chemokine* genes, and thus plays a crucial role in carcinogenic activity.

How Tipα protein secreted from *H. pylori* enters the gastric epithelial cells remains to be clarified. We think that receptor-mediated endocytosis is involved, based on results showing that homodimers of Tipα, C5A and C7A proteins specifically bound to the MGT-40 cells more strongly than monomers of del-Tipα and C5A/C7A proteins did. Recently it was reported that decay-accelerating factor (DAF), a glycoprotein, might act as a receptor for *H. pylori*: a mediator of gastric inflammation and genetic deficiency of DAF attenuated the development of inflammation among *H. pylori*-infected mice.[20]

After Tipα binds to the protein, Tipα trafficking may occur by caveolae or by lipid raft. It has been shown that some proteins—termed translocatory proteins—such as human immunodeficiency virus Tat, can enter the cell by endocytosis into endocytic vesicles and translocate into the nuclei, but the mechanisms are not yet clear.[18,21] We believe that determining the mechanisms of Tipα trafficking will provide new insight into carcinogenic mechanisms with *H. pylori* infection and various virulence factors.

Our findings showing that *H. pylori* obtained from gastric cancer patients secreted Tipα protein significantly more than those from chronic gastritis patients did indicate that larger amounts of Tipα protein induce stronger expression of *TNF*-α and *chemokine* gene expression in the gastric mucosa. Recently, Inoue *et al.* reported nasal vaccination with Tipα significantly reduced the inflammatory cytokines TNF-α and IL-12 in gastric mucosa of mice infected with *H. pylori*.[22] Cao *et al.*, using Mongolian gerbils, reported a severity of chronic gastritis associated with higher levels of mRNA of TNF-α and IL-1β, with pathogen *H. pylori* linked to glandular gastric carcinogenesis.[23] Thus, Tipα may induce cancer development through a unique mechanism that is completely different from that of CagA with Type IV secretion system.[24]

All the results show that Tipα is a unique carcinogenic factor of *H. pylori* and is both a suitable marker for detection of high risk in gastric cancer and a molecular target for cancer prevention.

## Acknowledgements

We thank Drs. Kei Nakachi and Hiroyasu Yamanaka for fruitful discussion, and Ms. Miki Kurusu, Kaori Suzuki and Ikuko Shiotani, Saitama Cancer Center, Research Institute for Clinical Oncology, and Ms. Maki Kimoto, Oita University, for their technical assistance.

## References


1. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Infection with *Helicobacter pylori*. In: IARC, ed. IARC monographs on the evaluation of carcinogenic risks to humans, Vol. 61: Schistosomes, liver flukes and *Helicobacter pylori*. Geneva: WHO, 1994. pp 1–241.
2. Peek RM, Jr, Blaser MJ. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. Nat Rev Cancer 2002;2:28–37.
3. Suganuma M, Okabe S, Marino MW, Sakai A, Sueoka E, Fujiki H. Essential role of tumor necrosis factor α (TNF-α) in tumor promotion as revealed by TNF-α-deficient mice. Cancer Res 1999;59:4516–18.
4. Suganuma M, Okabe S, Kurusu M, Iida N, Ohshima S, Saeki Y, Kishimoto T, Fujiki H. Discrete roles of cytokines. TNF-α, IL-1, IL-6 in tumor promotion and cell transformation. Int J Oncol 2002;20:131–6.
5. Moore RJ, Owens DM, Stamp G, Arnott C, Burke F, East N, Holdsworth H, Turner L, Rollins B, Pasparakis M, Kollias G, Balkwill F. Mice deficient in tumor necrosis factor-α are resistant to skin carcinogenesis. Nat Med 1999;5:828–31.
6. Suganuma M, Kurusu M, Suzuki K, Nishizono A, Murakami K, Fujioka T, Fujiki H. New tumor necrosis factor-α-inducing protein released from *Helicobacter pylori* for gastric cancer progression. J Cancer Res Clin Oncol 2005;131:305–13.
7. Suganuma M, Kuzuhara T, Yamaguchi K, Fujiki H. Carcinogenic role of tumor necrosis factor-α inducing protein of *Helicobacter pylori* in human stomach. J Biochem Mol Biol 2006;39:1–8.
8. Kuzuhara T, Suganuma M, Kurusu M, Fujiki H. *Helicobacter pylori*-secreting protein Tipα is a potent inducer of chemokine gene expressions in stomach cancer cells. J Cancer Res Clin Oncol 2007;133: 287–96.
9. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, et al. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature 1997;388:539–47.
10. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, et al. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature 1999;397:176–80.
11. Yoshida M, Wakatsuki Y, Kobayashi Y, Itoh T, Murakami K, Mizoguchi A, Usui T, Chiba T, Kita T. Cloning and characterization of a novel membrane-associated antigenic protein of *Helicobacter pylori*. Infect Immun 1999;67:286–93.
12. Suganuma M, Kurusu M, Okabe S, Sueoka N, Yoshida M, Wakatsuki Y, Fujiki H. *Helicobacter pylori* membrane protein 1: a new carcinogenic factor of *Helicobacter pylori*. Cancer Res 2001;61:6356–9.
13. Kuzuhara T, Suganuma M, Tsuge H, Fujiki H. Presence of a motif conserved between *Helicobacter pylori* TNF-α inducing protein (Tipα) and penicillin-binding proteins. Biol Pharm Bull 2005;28: 2133–7.
14. Kuzuhara T, Suganuma M, Oka K, Fujiki H. DNA-binding activity of TNF-α inducing protein from *Helicobacter pylori*. Biochem Biophys Res Commun 2007;362:805–10.
15. Ichinose M, Nakanishi H, Fujino S, Tatematsu M. Establishment and characterization of two cell lines from *N*-methyl-*N*-nitrosourea-induced mouse glandular stomach carcinomas. Jpn J Cancer Res 1998; 89:516–24.
16. Sumie A, Yamashiro T, Nakashima K, Nasu M, Watanabe M, Nishizono A. Comparison of genomic structures and antigenic reactivities of orthologous 29-kilodalton outer membrane proteins of *Helicobacter pylori*. Infect Immun 2001;69:6846–52.
17. Gómez MI, Lee A, Reddy B, Muir A, Soong G, Pitt A, Cheung A, Prince A. *Staphylococcus aureus* protein A induces airway epithelial inflammatory responses by activating TNFR1. Nat Med 2004;10:842–8.
18. Wang Y, Li D, Fan H, Tian L, Zhong Y, Zhang Y, Yuan L, Jin C, Yin C, Ma D. Cellular uptake of exogenous human PDCD5 protein. J Biol Chem 2006;281:24803–17.
19. Maeda S, Ogura K, Yoshida H, Kanai F, Ikenoue T, Kato N, Shiratori Y, Omata M. Major virulence factors, VacA and CagA, are commonly positive in *Helicobacter pylori* isolates in Japan. Gut 1998;42:338–43.
20. O'Brien DP, Israel DA, Krishna U, Romero-Gallo J, Nedrud J, Medof ME, Lin F, Redline R, Lublin DM, Nowicki BJ, Franco AT, Ogden S, et al. The role of decay-accelerating factor as a receptor for *Helicobacter pylori* and a mediator of gastric inflammation. J Biol Chem 2006;281:13317–23.
21. Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. Cell 1988;55:1189–93.
22. Inoue K, Suganuma M, Shiota S, Murakami K, Fujioka T, Nishizono A. Effects of nasal vaccination with TNF-α inducing protein (Tipα) against *Helicobacter*-infected mice. In: Proceedings of the 13th Annual Meeting of Japan Society of Helicobacter Research, 2007; p 122.
23. Cao X, Tsukamoto T, Nozaki K, Tanaka H, Cao L, Toyoda T, Takasu S, Ban H, Kumagai T, Tatematsu M. Severity of gastritis determines glandular stomach carcinogenesis in *Helicobacter pylori*-infected Mongolian gerbils. Cancer Sci 2007;98:478–83.
24. Odenbreit S, Puls J, Sedlmaier B, Gerland E, Fischer W, Haas R. Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. Science 2000;287:1497–500.

| CARCIN | bgp017 | VD |
|---|---|---|
| Journal Name | Art. No. | CE Code |

Carcinogenesis vol.0 no.0 pp.1–6, 2009
doi:10.1093/carcin/bgp017
*Advance Access publication January 15, 2009*

# Microsatellite instability-low colorectal cancer acquires a *KRAS* mutation during the progression from Dukes' A to Dukes' B

**Shin-ichi Asaka[1,2], Yoshiko Arai[1], Yoji Nishimura[2], Kensei Yamaguchi[3], Tsutomu Ishikubo[3], Toshimasa Yatsuoka[2], Yoichi Tanaka[2] and Kiwamu Akagi[1,*]**

[1]Division of Molecular Diagnosis and Cancer Prevention, [2]Division of Gastroenterological Surgery and [3]Division of Gastroenterology, Saitama Cancer Center, 818, Komuro, Ina-machi, Kitaadati-gun, Saitama 362-0806, Japan

*To whom correspondence should be addressed. Tel: +81 48 722 1111; Fax: +81 48 723 5197;
Email: akagi@cancer-c.pref.saitama.jp

**The classification of colorectal cancer (CRC) by microsatellite instability (MSI) status is important for effective clinical management. In fact, microsatellite instability-high (MSI-H) cancer has distinctive clinicopathological and molecular features. However, microsatellite instability-low (MSI-L) cancer is not clearly defined. The objective of this study was to further clarify the characteristics of MSI-L CRC. A consecutive series of 940 primary CRCs were subdivided into three groups according to the level of MSI and analyzed the clinicopathological features and genetic changes in the *KRAS*, *BRAF* and *p53* mutation and the loss of heterozygosity (LOH) of adenomatous polyposis coli (*APC*) gene and methylation status of the *O*$^6$-methylguanine-DNA methyltransferase (*MGMT*) and *MLH1* promoter. Of the 940 CRCs, 5.9% were MSI-H, 7.1% were MSI-L and 87% were microsatellite stable (MSS). *KRAS* and *BRAF* mutations were detected in 39.4 and 4.6% of the CRCs, respectively. The frequency of *KRAS* mutations in MSI-H, MSI-L and MSS cancer was 30, 48 and 39%, respectively. The proportion of *KRAS* mutations in MSI-L cancer increased from 16 to 63% accompanying the progression from Dukes' A to Dukes' B. While the LOH of D5S346, which is located near the *APC* gene, and *p53* mutation was observed in 75 and 67% of MSI-L CRC at Dukes' A, respectively. These results indicated that the LOH of *APC* and *p53* mutation has already occurred by the Dukes' A lake 'suppressor pathway' but not the *KRAS* mutation in MSI-L CRCs. The genes involving MSI-L carcinogenesis are similar to MSS but the timing and frequency of the *KRAS* mutation is different.**

## Introduction

There are two types of genomic instability, microsatellite instability (MSI or MIN) or chromosomal instability associated with the carcinogenesis process of colorectal cancer (CRC) (1). The great majority of CRCs develop through the chromosomal instability pathway (also called the 'suppressor pathway'), which arise from adenomas and is initiated by the inactivated adenomatous polyposis coli (*APC*) gene and followed by the well-established genetic steps involved in the adenoma–carcinoma sequence (2,3). While another type of genomic instability, MSI caused by a failure of the DNA mismatch repair (MMR) system, is observed in ~10% of all CRCs (4–7). DNA MMR deficiency leads mutations in the target genes that are implicated in tumor progression such as *TGFbetaR2* (8), *IGF2R* (9), *CDX2* (10) and *BAX* (11) and it is known as the 'mutator pathway'. MSI can be subdivided into three groups, microsatellite instability-high (MSI-H), microsatellite instability-low (MSI-L) and microsatellite stable (MSS), according to the degree of instability. The recommended method to distinguish these subgroups is to analyze paired tumor and normal tissue DNAs using a panel of five microsatellite markers known as the Bethesda panel (12).

The MSI-H CRCs phenotype is more likely to occur at a proximal site, to occur in women, to be associated with a favorable prognosis (5,7,13–15) and severe inflammatory cell infiltration into the tumor tissue (16,17). A large percent of MSI-H CRC is sporadic and demonstrates somatic promoter methylation of the *hMLH1* gene (18,19), whereas a germ line mutation of the MMR genes, such as *hMSH2*, *hMLH1*, *hMSH6* and *hPMS2*, is found in the majority of MSI-H CRC without *hMLH1* promoter methylation and is known as Lynch syndrome/hereditary non-polyposis colorectal cancer (20–23). Recent morphological and molecular studies have proposed the existence of a serrated pathway, thus suggesting that serrated polyps may serve as a precursor of the MSI-positive cancers (24–26).

On the other hand, most studies have found no obvious clinicopathological or molecular differences between MSI-L and MSS cancers (27). The DNA MMR genes *hMLH1* and *hMLH2* do not appear to be implicated in the MSI-L subset (28). Some studies have reported that MSI-L is associated with cancers from individuals with germ line mutations of *hMLH6* (29), but genetic alteration of this gene is infrequent in MSI-L CRC patients. Furthermore, some researchers deny the presence of MSI-L cancers because most non-MSI-H cancers exhibit MSI-L when large numbers of microsatellite loci are tested (27,30).

Meanwhile, there is evidence indicating that the MSI-L phenotype could reflect a distinct pathway of tumor development with a different clinical behavior and different genetic and epigenetic changes. For example, a high frequency of a *KRAS* mutation (31,32) that is associated with loss of expression of the *O*$^6$-methylguanine-DNA methyltransferase (*MGMT*) gene by methylation of its promoter region (33), lower frequency of 5qLOH (31), a high frequency of *APC* mutation (34) and reduced expression of Bcl-2 protein (35) are global molecular phenotypes by which MSI-L cancers are distinguished from non-MSI-L cancers.

Therefore, MSI-L CRC is still controversial. This study investigated the genetic changes and clinicopathological features of MSI-L CRCs using a series of 940 CRCs.

**Abbreviations:** APC, adenomatous polyposis coli; CI, confidence interval; CRC, colorectal cancer; HR, hazard ratio; LOH, loss of heterozygosity; MGMT, *O*$^6$-methylguanine-DNA methyltransferase; MMR, mismatch repair; MSI, microsatellite instability; MSI-H, microsatellite instability-high; MSI-L, microsatellite instability-low; MSS, microsatellite stable; PCR, polymerase chain reaction.

## Materials and methods

### *Patients and tissue samples*

A consecutive series of 940 primary CRCs excised surgically at Saitama Cancer Center from January 1998 to May 2006 were investigated after obtaining the informed consent. Any patients who were treated by preoperative radiotherapy or chemotherapy were excluded. Furthermore, patients with inflammatory bowel disease or a known history of familial adenomatous polyposis were also excluded. This study was approved by the Ethics Committee of the Saitama Cancer Center.

### *Analysis of MSI*

Primary CRCs and paired normal colorectal mucosa obtained by surgery were immediately frozen at −80°C. The genomic DNA was extracted from fresh frozen specimens using standard methods. The Bethesda five markers, BAT25, BAT26, D5S346, D2S123 and D17S250, were used to classify the MSI status of the tumors. Polymerase chain reaction (PCR) and subsequent analyses were performed as reported previously (5). CRCs were subdivided into three groups according to the degree of MSI; MSI-H if two or more of the five markers show instability, MSI-L if only one marker shows instability and MSS if absence of MSI in all five markers. MSI-positive markers were re-examined at least twice to confirm the result. Loss of heterozygosity (LOH) was defined by at least a 30% reduction in the relative intensity of one allele in the tumor in comparison with normal levels.

*Analysis of KRAS, BRAF and p53 mutation*

The mutations in exon 1 and 2 of the *KRAS* gene were analyzed by denaturing gradient gel electrophoresis as described previously (36).

The *BRAF* V600E mutation was examined using PCR combined with restriction enzyme digestion. DNA fragments containing exon 15 of the *BRAF* gene were amplified by PCR using the following oligonucleotide primers: *BRAF* forward primer 5′-CTGTTTTCCTTTACTTACTACACC-3′ and *BRAF* reverse primer 5′-CTGTTCAAACTGATGGGACC-3′. PCR amplification was carried out with 100 ng of genomic DNA in a volume of 20 μl containing 0.2 μM deoxynucleoside triphosphate, 0.1 μM each of primers and 1 U Taq Gold. Thermal cycling was initiated with denaturation at 94°C for 10 min followed by 37 three-step cycles at 94°C for 30 s, 59°C for 30 s and 72°C for 45 s and followed by a final incubation for 7 min at 72°C. PCR products were digested with HpyCH4III at 37°C for 1 h and analyzed on 8% polyacrylamide gels.

The mutations in exon 5–8 of p53 gene were analyzed by denaturing gradient gel electrophoresis as described previously (37).

*Analysis of hMLH1 and MGMT promoter methylation*

*hMLH1* and *MGMT* promoter methylation was analyzed in 55 of MSI-H cancer and 67 of MSI-L cancer samples. The methylation status of each gene was determined by the methods previously reported (5). The primers were *hMLH1* methylation specific, 5′-AACGAATTAATAGGAAGAGGCGGATAGCG-3′ and 5′-CGTCCCTCCCTAAAACGACTACTACCC-3′; *hMLH1* unmethylation specific, 5′-TAAAAATGAATTAATAGGAAGAGTGGATAGTG-3′ and 5′-AATCTCTTCATCCCTCCCTAAAACA-3′; *MGMT* methylation specific, 5′-TTTCGACGTTCGTAGGTTTTCGC-3′ and 5′-GCACTCTTCCGAAAACGAAACG-3′ and *MGMT* unmethylation specific, 5′-TTTGTGTTTTGATGTTTGTAGGTTTTTGT-3′ and 5′-AACTCCACACTCTTCCAAAAACAAAACA-3′.

*Statistical analysis*

Differences were assessed using the chi-square or Fisher's exact test for categorical variable and unpaired Student's *t*-test for continuous factors. The overall survival was defined as the interval from the date of resection until the date of death from any cause, censored patients being those alive at the close of the study or lost to follow-up. Survival was measured from the date of the resection of the CRCs until death or until the censor date of 1 July 2006. The distribution of survival time was compared with the use of the log-rank test; survival distribution curves were estimated by the method of Kaplan–Meier. Multivariate analyses were performed with the use of the Cox proportion hazard model. The independent prognostic factors for survival were determined by a stepwise backward conditional selection in which the non-significant factors ($P > 0.1$) were successively rejected. All statistical analyses were performed using the StatView 5.5 program. $P < 0.05$ was considered to be statistically significant in all cases.

## Results

*MSI status*

The Bethesda panel, BAT25, BAT26, D5S346, D2S123 and D17S250, was used to classify the MSI status of the tumors. Of the 940 CRCs, 55 (5.9%) were MSI-H, 67 (7.1%) were MSI-L and 818 (87%) were MSS (Table I). Mononucleotide marker BAT25 and BAT26 exhibited

**Table I.** Clinicopathological and genetic features of CRCs

| | MSS, *n* (%) | MSI-L, *n* (%) | MSI-H, *n* (%) | *P* value | | |
|---|---|---|---|---|---|---|
| | | | | MSI-L versus MSS | MSI-H versus MSS | MSI-L versus MSI-H |
| Patient | 818 (87.0) | 67 (7.1) | 55 (5.9) | 0.3723 | 0.0011 | 0.0661 |
| Men | 509 (62.2) | 38 (56.7) | 22 (40) | | | |
| Women | 309 (37.8) | 29 (43.3) | 33 (60) | | | |
| Mean (±SE) age | 63.6 ± 10.3 | 63.3 ± 9.95 | 60.5 ± 13.4 | 0.785 | 0.035 | 0.199 |
| Location | | | | 0.44 | <0.0001 | 0.0003 |
| Proximal | 217 (26.5) | 22 (32.8) | 36 (65.5) | | | |
| Distal | 243 (29.7) | 16 (23.9) | 12 (21.8) | | | |
| Rectum | 358 (43.8) | 29 (43.3) | 7 (12.7) | | | |
| Tumor size | | | | 0.23 | <0.0001 | <0.0001 |
| Mean ± standard error (mm) | 45.42 ± 24.1 | 41.8 ± 23.0 | 61.35 ± 32.3 | | | |
| Histologic feature | | | | >0.999[a] | <0.0001[a] | 0.0018[a] |
| Well differentiated | 73 (8.9) | 10 (14.9) | 4 (7.3) | | | |
| Moderately differentiated | 706 (86.3) | 54 (80.6) | 38 (69.1) | | | |
| Poorly differentiated | 16 (2.0) | 3 (4.5) | 10 (18.2) | | | |
| Mucinous | 21 (2.6) | 0 (0) | 2 (3.6) | | | |
| Others | 2 (0.2) | 0 (0) | 1 (1.8) | | | |
| Mucinous component | | | | 0.73 | <0.0001 | 0.001 |
| + | 84 (10.3) | 6 (9.0) | 18 (32.7) | | | |
| − | 734 (89.7) | 61 (91.0) | 37 (67.3) | | | |
| Dukes' stage | | | | 0.22 | 0.012 | 0.039 |
| A | 150 (18.3) | 19 (28.4) | 12 (21.8) | | | |
| B | 246 (30.1) | 16 (23.9) | 26 (47.3) | | | |
| C | 246 (30.1) | 20 (29.9) | 13 (23.6) | | | |
| D | 176 (21.5) | 12 (17.9) | 4 (7.3) | | | |
| Depth of tumor invasion | | | | 0.27 | 0.073 | 0.42 |
| T1 | 68 (8.3) | 8 (11.9) | 5 (9.1) | | | |
| T2 | 123 (15.0) | 14 (20.9) | 7 (12.7) | | | |
| T3 | 571 (69.8) | 39 (58.2) | 34 (61.8) | | | |
| T4 | 56 (6.8) | 6 (9.0) | 9 (16.4) | | | |
| Extramural venous invasion | | | | 0.99 | 0.69 | 0.76 |
| + | 586 (71.6) | 48 (71.6) | 38 (69.1) | | | |
| − | 232 (28.4) | 19 (28.4) | 17 (30.9) | | | |
| KRAS mutation | | | | 0.1836 | 0.1798 | 0.051 |
| + | 311 (39.5) | 32 (47.8) | 16 (30.2) | | | |
| − | 477 (60.5) | 35 (52.2) | 37 (69.8) | | | |
| BRAF mutation | | | | 0.3885 | <0.0001 | <0.0001 |
| + | 21 (2.7) | 3 (4.5) | 17 (32.1) | | | |
| − | 767 (97.3) | 64 (95.5) | 36 (67.9) | | | |

[a]Well and moderately differentiated versus mucinous and poorly differentiated.

instability in 95% of MSI-H CRCs whereas in 1–3% of MSI-L CRCs. Therefore, BAT25 and BAT26 were identified to be the most specific and sensitive markers to detect MSI-H CRCs. The dinucleotide marker D2S123 exhibited instability not only in 95% of the MSI-H CRCs but also in 56.7% of MSI-L. D2S123 is the most sensitive but not specific for MSI-L (Figure 1A).

*Clinicopathological features*

The association of MSI status with the clinicopathological features in the 940 CRCs is shown in Table I. Consistent with the findings of previous studies, MSI-H cancers are observed more frequently in females, in the proximal colon and in poorly differentiated or mucinous CRCs in comparison with MSS. While some differences were observed between MSI-L and MSS cancer, with regard to the female to male ratio, the site of the tumor and the stage did not reach significance.

The prognosis was assessed based on the MSI status (Figure 1B). Since no Dukes' A patients died during the follow-up period, these patients were excluded from the overall survival analysis. In total, 155 of the 731 Dukes' B–D patients (21.2%) died during a mean follow-up period of 30.3 ± 19 months after surgery. The prognosis of patients with MSI-H tumors was significantly better than that of patients with MSS tumors (log-rank test, $P = 0.0335$). The prognosis of patient with MSI-L tumors had an intermediate tendency among the three groups (Figure 1B).

In a stepwise multivariate analysis, age {hazard ratio (HR) 1.627 [confidence interval (CI) 1.216–2.301]; $P = 0.0016$}, men sex [HR 1.429 (CI 1.019–2.004); $P = 0.0388$], low-grade pathology [HR 2.029 (CI 1.231–3.343); $P = 0.0055$], *KRAS* [HR 1.69 (CI 1.215–2.351); $P = 0.0018$], *BRAF* [HR 3.593 (CI 1.933–6.678); $P < 0.0001$] and Dukes' stage [Dukes' B versus Dukes' C: HR 1.636 (CI 0.964–2.775); $P = 0.068$ and Dukes' B versus Dukes' D: HR 10.406 (CI 6.548–16.537); $P < 0.0001$] were independent variables. However, MSI was not an independent variable.

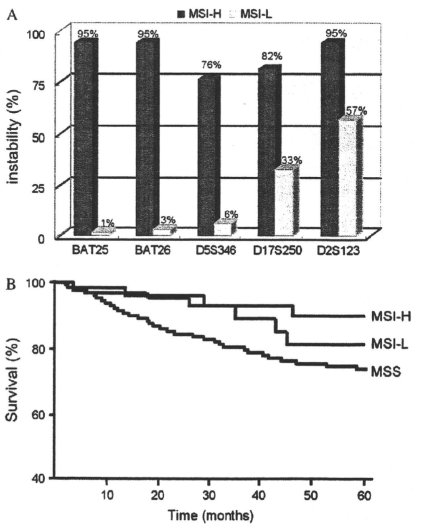


**Fig. 1.** (**A**) Frequency of markers demonstrated instability for MSI-H and MSI-L CRCs. (**B**) Overall survival of patients according to the MSI status. In total, 155 of the 731 Dukes' B–D patients (21.2%) died during a mean follow-up period of 30.3 ± 19 months after surgery.

*Mutation analysis of the KRAS and BRAF genes*

A *KRAS* mutation was detected in 39.4% and a *BRAF* V600E mutation in 4.6% of the 905 CRCs that were examined. The *BRAF* mutation was found more frequently in MSI-H cancer (32%) in comparison with MSS (3%) and MSI-L cancers (4%; $P < 0.0001$, Figure 2A). The frequency of *BRAF* mutation decreased accompanying the tumor progression in MSI-H cancer, whereas it increased in MSI-L and MSS cancers (Figure 2D–F).

The *KRAS* mutation analysis in CRCs demonstrated that MSI-L cancer showed higher frequency of the *KRAS* mutation than MSS and MSI-H cancers: MSS 39% (311/788), MSI-H 30% (16/53) and MSI-L 48% (32/67; MSI-L versus MSI-H; $P = 0.066$, MSI-L versus MSS; $P = 0.244$, MSI-H versus MSS; $P = 0.180$; Figure 2A). However, accompanying the progression from Dukes' A to Dukes' B, the frequency of the *KRAS* mutation in MSI-L cancer drastically increased from 16 to 63% (Figure 2E, MSI-L; *KRAS* mutation in Dukes A versus *KRAS* mutation in Dukes B–D, $P = 0.045$, Fisher's exact test) and was significantly higher than that in MSS or MSI-H cancers at Dukes' B–D (MSI-L versus MSS or MSI-H; $P = 0.014$, $P = 0.0394$, respectively; Figure 2C). MSI-H cancer also demonstrated an increased proportion of the *KRAS* mutation accompanying the progression from Dukes' A to Dukes' B (Figure 2F), but the number of MSI-H cases was too small to find significance (MSI-H; *KRAS* mutation in Dukes' A versus *KRAS* mutation in Dukes' B–D, $P = 0.08$, Fisher's exact test). The ratio of the *KRAS* mutation in MSI-H cancer was the same as that in MSS cancer after Dukes' B stage (Figure 2C). The ratio of tumors having either the *KRAS* or *BRAF* mutation at Dukes' B–D in MSI-H and MSI-L cancers was statistically higher than that in MSS cancer [MSS (40%) versus MSI-L (66%) or MSI-H (63%); $P = 0.0034$, $P = 0.0108$, respectively; Figure 2C].

*Type of KRAS mutation*

Of the 321 tumors with *KRAS* mutations, 196 (61%) were a G to A transition, 107 (33%) were a G to T transversion and 18 (6%) were a G to C transversion. The type of KRAS mutation was investigated in each MSI status. This revealed that 93% (14 of 15 tumors) of the *KRAS* mutations were a G to A transition in MSI-H cancer and there were significant differences between the types of *KRAS* mutations among the three MSI groups ($P = 0.0152$, chi-square test). The frequency of G to A transition mutations in MSI-L cancer was lower than in MSI-H but higher than in MSS cancer (Figure 2G). To investigate whether the high frequency of G to A transition mutation of *KRAS* gene in MSI-H and MSI-L cancer is involved in the inactivation of *hMLH1* or *MGMT*, the methylation status of the *hMLH1* and *MGMT* promoter was analyzed. Of the 30 MSI-L tumors with *KRAS* mutations, 13 (43%) had *MGMT* promoter methylation, whereas among 35 MSI-L tumors without *KRAS* mutations, 10 (29%) had it (Table II). Furthermore, 53% (10 of 19) of MSI-L tumors with G to A transition mutations in *KRAS* harbored *MGMT* promoter methylation, whereas 30% (three of 10) of MSI-L tumors with G to C or T transversion mutations in *KRAS* and 29% (10 of 35) of MSI-L tumors without a *KRAS* mutation showed *MGMT* promoter methylation (G to A versus G to C or T and wild-type, $P = 0.0705$; Table III).

These results suggest that G to A transition mutations in MSI-L tumors seem to correlate with *MGMT* promoter methylation ($P = 0.0705$). On the other hand, the frequency of *MGMT* methylation was observed in 38% of MSI-H tumors with *KRAS* mutation and 56% of MSI-H tumor without KRAS mutations. This result suggests that there is an inverse correlation between *KRAS* mutations and *MGMT* methylation in MSI-H tumors ($P = 0.2026$; Table II).

The frequency of *hMLH1* methylation was observed 25% of MSI-H tumors with *KRAS* mutations and 59% of MSI-H tumors without KRAS mutations. This result clearly shows a significant inverse correlation between *KRAS* mutations and *hMLH1* methylation in MSI-H tumors ($P = 0.0366$). The frequency of *hMLH1* methylation was

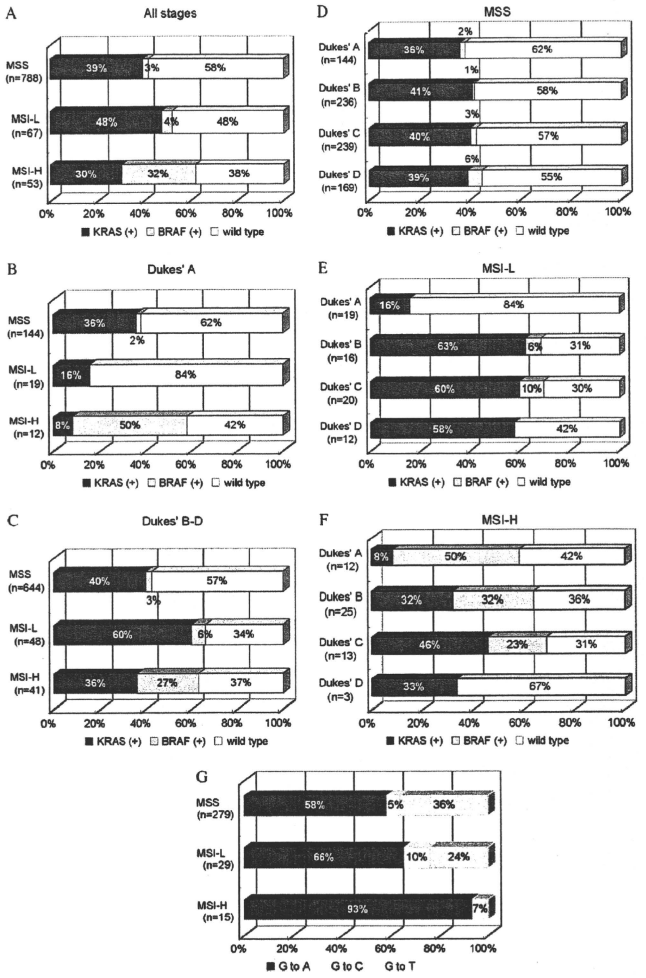


**Fig. 2.** *KRAS* and *BRAF* mutations in each MSI status. **(A)** *KRAS* and *BRAF* mutation of all CRCs. **(B)** *KRAS* and *BRAF* mutation of Dukes' A CRCs. **(C)** *KRAS* and *BRAF* mutation of Dukes' B–D CRCs. Frequency of *KRAS* and *BRAF* mutations at each stage according to the MSI status; **(D)** MSS, **(E)** MSI-L and **(F)** MSI-H. **(G)** Spectrum of KRAS mutations in each MSI status.

significantly higher in MSI-H tumors (50%) than in MSI-L tumor (6.7%; $P < 0.0001$; Table II).

*LOH of D5S346 and p53 mutation in MSI-L CRCs*

Since D5S346, one of the MSI makers, locates near the *APC* gene, MSI analysis with the Bethesda panel can also assess the LOH of *APC* gene, simultaneously. LOH of the D5S346 and *p53* mutation was detected in 75% (9/12) and 67% (12/18) of MSI-L CRC at Dukes' A, respectively (Table IV). In addition, the frequency of LOH of D5S346 and *p53* mutations in MSI-L at Dukes' B–D were 55 and 61%, respectively. These results indicate that LOH of *APC* and *p53* mutations has already occurred before Dukes' A.

**Table II.** Promoter methylation and KRAS mutation

| | | KRAS mut, *n* (%) | Wt, *n* (%) | *P* value |
|---|---|---|---|---|
| | MGMT | | | |
| MSI-L | M | 13 (43) | 10 (29) | 0.2147 |
| | U | 17 (57) | 25 (71) | |
| MSI-H | M | 6 (38) | 22 (56) | 0.2026 |
| | U | 10 (62) | 17 (44) | |
| | hMLH1 | | | |
| MSI-L | M | 2 (7) | 2 (6) | >0.9999 |
| | U | 28 (93) | 33 (94) | |
| MSI-H | M | 4 (25) | 23 (59) | 0.0366 |
| | U | 12 (75) | 16 (41) | |

M, methylated; U, unmethylated; mut, mutation; Wt, wild-type.

**Table III.** Type of KRAS mutation according to MGMT methylation

| | MGMT | KRAS mutation | | Wt, *n* (%) | *P* value |
|---|---|---|---|---|---|
| | | G to A, *n* (%) | G to C, T; *n* (%) | | |
| MSI-L | M | 10 (53) | 3 (30) | 10 (29) | 0.0705 |
| | U | 9 (47) | 7 (70) | 25 (71) | |
| MSI-H | M | 6 (43) | 0 (0) | 22 (56) | 0.4339 |
| | U | 8 (57) | 1 (100) | 17 (44) | |

M, methylated; U, unmathylated; Wt, wild type. *P* : G to A versus G to C, T + Wt.

## Discussion

### *Molecular features*

Although the *BRAF* and *KRAS* mutations are found more frequently in MSI-H and MSI-L CRC, respectively (31,34,38), the frequency of *KRAS* and *BRAF* mutation changed between each tumor stage in this study.

The development of CRC requires a multistep process characterized by the accumulation of genetic alterations. According to the well-known genetic model for colorectal tumorigenesis proposed by Fearon and Vogelstein, *KRAS* mutations occur in the early to intermediate adenomas (3). However, the frequency of *KRAS* mutations was significantly lower (16%) at Dukes' A and higher (60%) at Dukes' B–D in MSI-L CRCs. This means that most *KRAS* mutations occurred at different times in MSI-L CRC, namely, during the progression from Dukes' A to Dukes' B but not in early to intermediate adenomas. It has been reported previously that the *KRAS* mutation is found more frequently in MSI-L CRCs (31,32), but according to the current detailed study, it depends on the tumor stage. Since a large number of specimens were collected in an unbiased manner for this study, the results demonstrate representative findings of CRCs in Japan.

Meanwhile LOH of D5S346, which is located near the *APC* gene, and the *p53* mutation was observed in 75% (9/12) and 67% (12/18) of MSI-L CRC at Dukes' A, respectively. These frequencies were almost the same at Dukes' B–D in MSI-L CRC (Table IV).

Taken together, these findings indicated that LOH of *APC* and *p53* mutations have already occurred by the Dukes' A lake suppressor pathway but not the *KRAS* mutation in MSI-L CRCs.

MSI-L CRC may develop through a mild mutator pathway, which differs from the suppressor and mutator pathway and show different clinical features (31,39). However, some studies doubt the presence of the MSI-L group (27,30). In the current study, the involved genes such as LOH of *APC*, *KRAS* and *p53* mutation in MSI-L CRCs are similar to those in MSS CRCs, but at least the timing and frequency of the *KRAS* mutation is different. This may explain why the clinicopathological features of MSI-L tumors are similar to those of MSS tumors but not completely identical.

**Table IV.** LOH of D5S346 and p53 mutation in MSI-L CRCs

| Dukes' stage | D5S346 LOH | p53 mutation |
|---|---|---|
| A | 75% (9/12) | 67% (12/18) |
| B | 46% (6/13) | 57% (12/21) |
| C | 67% (8/12) | 45% (10/22) |
| D | 50% (3/6) | 92% (12/13) |

On the other hand, considering the presence of the *BRAF* mutation and methylation of the *hMLH1* promoter at the early stage in MSI-H CRC, these genetic changes should occur in the precursor of MSI-H CRC. This is not inconsistent with the concept of serrated pathways resulting from serrated polyps that were revealed in recent morphological and molecular studies (24–26,40–42).

The mechanism of the *KRAS* mutation was also analyzed and the results showed that a G to A transition mutation of *KRAS* occurs more frequently in MSI-L than MSS. Some reports demonstrated that *MGMT* inactivation by promoter methylation causes a G to A transition mutation of *KRAS* (33,43,44) and *p53* (45) and such a mutation is frequently observed in MSI-L CRCs. We attempted to determine whether or not the inactivation of *MGMT* by promoter methylation is associated with the type and frequency of *KRAS* mutation. Our results indicated that *MGMT* promoter methylation seems to affect the G to A transition and frequency of the KRAS mutation in MSI-L CRC. However, most *KRAS* mutations in MSI-H CRC show a G to A transition, *MGMT* inactivation was inversely related and the *BRAF* mutation often observed in MSI-H CRC shows a T to A transversion. Considering these results, a different mechanism might therefore be involved in mutation between MSI-L and MSI-H CRC.

### *Clinical feature*

As mention above, the genes associated with developing MSI-L CRC are similar to the suppressor pathway but the frequency and timing of KRAS mutations is different; thus, there may be different clinical and pathological features in MSI-L.

Comparing the stage distribution for each MSI status, the distribution of Dukes' B in MSI-H CRC is significantly larger than in MSS and MSI-L CRC (7). Gyef *et al.* (15) demonstrated with a logistic analysis that MSI-H CRC is less metastatic to the regional lymph nodes and distant organs than MSS CRC, even though their depth of tumor invasion is same. The same result was observed in the current study, but MSI-L CRC did not show this characteristic. This suggests that there is a mechanism restricting the progression from Dukes' B to C in MSI-H cancer. Although the precise explanation for this mechanism is still unknown, tumor-infiltrating lymphocytes, apoptosis, proliferative activity (46,47) or a mutation of *p53* (7) may lead to the 'restraining effect'.

On the other hand, the distribution of Dukes' A in MSI-L CRC is larger than MSS. Considering the *KRAS* mutation during the progression from Dukes' A to B in MSI-L CRC, tumor progression may be hindered until the occurrence of the *KRAS* mutation in MSI-L CRC.

Various studies have reported the prognosis of each MSI status. Some investigations show that the patients with MSI-H cancer demonstrate a better prognosis and the patients with MSI-L cancer have a poorer survival than patients with MSS in stage C (48,49). Considering the high frequency of the *KRAS* mutation after Dukes' B in MSI-L cancer, the worse prognosis of such patients may therefore be reasonable (50).

Although the number of cases in the current study was not sufficient to study the prognosis at each stage, among all patients with Dukes' B to D CRCs, MSI-H patients showed significantly better survival than MSS (log-rank test, $P = 0.0335$) while MSI-L patients had a slightly better prognosis than MSS. These findings may result from the fact that the proportion of Dukes' D for each MSI status is smaller, in order, MSI-H, MSI-L and MSS.