MALAT-1: A long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis

CHUAN XU[*], MINHUI YANG[*], JIE TIAN[*], XIAOYAN WANG and ZUGUO LI

Department of Pathology, Nanfang Hospital, Southern Medical University,
Guangzhou 510515, Guangdong Province, P.R. China

Abstract. The human metastasis associated lung adeno-carcinoma transcript 1 (MALAT-1) as a long non-coding RNA known to be misregulated in many people who are detected with cancer. Our earlier studies found that MALAT-1 plays a pivotal role in colorectal cancer (CRC) metastasis. In this study, we analyzed the MALAT-1 gene in five fragments. We employed the sequencing process to identify MALAT-1 mutations in the following types of samples: CRC cells (SW620, SW480), normal colorectal tissues, and primary CRC tissues. We were successful in detecting the following mutations: fragment 5434 nt-6951 nt of the MALAT-1 was mutated in SW620 cells, while fragments 5434 nt-6951 nt and 6918 nt-8441 nt of MALAT-1 were mutated in SW480 cancer cells and primary CRC tissues. We over-expressed five fragments of MALAT-1 in the CRC cell line SW480; simultaneously ensuring that MALAT-1 had low expression. Our data illustrated that one of the 5 fragments (6918 nt-8441 nt) located at the 3' end of MALAT-1 plays a pivotal role in the biological processes of cell proliferation, migration and invasion. Based on these observations, we can infer that the 3' end of MALAT-1 is an important biological motif in the invasion and metastasis of CRC cells. We have successfully presented the first evidence that mutations were found on the long non-coding RNA MALAT-1 in CRC. Moreover, long non-coding RNA MALAT-1 has an important biological motif located at the 3' end of MALAT-1 (6918 nt-8441 nt) in CRC. Our study gives a new direction to research primarily focused on exploring the molecular mechanisms occurring during the invasion and metastasis of CRC.

Correspondence to: Dr ZuGuo Li, Department of Pathology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, P.R. China
E-mail: zuguole@sina.com

[*]Contributed equally

Introduction

Coding RNA and non-coding RNA (ncRNA) constitute the two major families of RNA. As compared with microRNA, long non-coding RNA research is gaining attention. In this study, several interesting facts were revealed through the sequencing of the human genome. This study was successful in discovering the following facts: there are only ~20,000 protein-coding genes, representing <2% of the total genomic sequence (1). It is still not clear if this pervasive transcription should be considered as completely useless (transcriptional 'noise') (2-4) or if non-coding RNAs (ncRNAs) with important functions exist (5). Some data indicate that these non-coding RNAs (ncRNAs) perform several significant functions, such as transcription interference, induction of chromatin remodeling and histone modifications, modulating alternative splicing patterns, generating endo-siRNAs, modulating protein activity, playing a role in structure or organization, altering protein localization, and acting as a small RNA precursor (6).

CRC is one of the most common malignant tumors, with an increasing rate of incidence that is reported throughout the world. Metastasis is the main cause for the failure of the treatment (7). We cannot exactly identify individuals who are more prone to metastasis on the basis of their clinical and histo-pathological findings, including the molecular markers that are known (8,9). We have not been successful in gaining a thorough understanding of the early molecular mechanisms (10-14). Current treatment methods are not capable of deciphering key genes and signaling cascades involved in the metastasis of human CRC thus remaining elusive. The human metastasis associated lung adenocarcinoma transcript 1 (MALAT-1) is basically defined as follows: MALAT-1 is a long non-coding RNA consisting of more than 8,000 nt expressed from chromosome 11q13. It has high expression in many cancer types, and it is related to metastasis. Because MALAT-1 lacks significant open reading frame, it cannot translate protein *in vitro* (15). Furthermore, the MALAT-1 transcript is associated with human lung adenocarcinoma metastasis: it is a long non-coding RNA that is known to be mis-regulated in many types of cancers detected in humans (16). It has been demonstrated that high expression of this gene is related to CRC metastasis (17).

In this study, MALAT-1 was divided into five segments, each overlapping and covering the entire length of the MALAT-1.

We used an RT-touchdown PCR to amplify these five fragments of the MALAT-1, and the mutations were identified on the basis of sequence analysis. For conducting further characterization of the role of these five segments of MALAT-1 in CRC invasion, we constructed an over-expression vector. In addition, functional experiments were performed using human CRC SW480 cell line - a low *in vitro* metastatic potential cell line.

Materials and methods

Nucleic acid extraction and amplification by RT-touchdown PCR. RNA was extracted from CRC cell lines SW620 and SW480, CRC tissues samples, and normal colorectal tissues samples using TRIzol Reagent (Invitrogen, Germany). One microgram from each sample was used for the complete procedure of cDNA synthesis. Because the RNA was available in limited amounts, cDNA was synthesized and amplified in accordance with the RT-touchdown PCR (all primer sequences are given in Table II). Amplification of GAPDH served as a loading control. PCR conditions that were employed in the process are briefly stated: 1 min at 95°C, 20 sec at 95°C, 20 sec at 62°C, and 1 min at 72°C. PCR was amplified for 2 PCR cycles with 1°C lower annealing temperature, until 50°C, 20 sec at 95°C, 20 sec at 50°C, 1 min at 72°C, 11 cycles, and 3 min at 72°C, a total of 35 cycles. The PCR amplified products were subjected to electrophoreses on a 1% agarose gel.

CDNA cloning and sequencing. The PCR products were purified from CRC cell lines SW620 and SW480, CRC tissues samples and normal colorectal tissues were cloned into a standard vector (PTA2 Cloning, Toyobo). Eight positive clones from each reaction were randomly selected and grown overnight in 2 ml of LB medium. Then, these plasmids were isolated. The inserted fragments of MALAT-1 were identified and analyzed using restricting enzyme digestion. Furthermore, individual clones were sequenced using standard procedures of dye terminator chemistry on a 3700 Sequencer (Applied Biosystems). Sequences were compared to sequences in GeneBank using the Blast program at the NCBI. After conducting sequencing successfully, the five fragments of the MALAT-1 gene were sub-cloned in the pEGFP-C1 vector (Clontech, USA); then these recombinant eukaryotic expression vectors were constructed using the gene recombination technique for construction of the recombinant plasmids pEGFP-C1-MALAT-1/fragments. Eight positive clones were confirmed on the basis of restricting enzyme digestion and sequence analysis.

Cell culture and transfection. SW620 and SW480 cells were maintained in RPMI medium 1640 supplemented with 10% FBS, 2 mM L-glutamine, penicillin (100 units/ml), and streptomycin (100 μg/ml) in a humidified atmosphere of 5% CO_2 and is maintained at a temperature of 37°C. Cells were plated in growth medium without antibiotics at a density of 50-60%; this was done a day before conducting transfection. In the human colon cancer SW480 cell line, transfection of these five pEGFP-C1-MALAT-1/fragment vectors was conducted using Lipofectamine 2000 (Invitrogen, Germany) according to the manufacturer's protocol for 48 h. Fluorescent microscopy was used for observing expression of GFP transfected cells.

Real-time quantitative polymerase chain reaction (RT-QPCR). After transfection, 1 μg total of RNA was reverse transcribed using SuperScript reverse transcriptase (Takara) and oligo (dT) primers. The resulting cDNA samples were amplified by real-time PCR using primers specific for MALAT-1/2997 nt-5541 nt, MALAT-1/202 nt-1880 nt, MALAT-1/5434 nt-6951 nt, MALAT-1/6918 nt-8441 nt, MALAT-1/1821 nt-3247 nt, and GAPDH (the primer sequences are given in Table III). Amplification of GAPDH served as a loading control. The experiments were performed in triplicate.

Flow cytometry assay. After transfection, cells were collected when they were in an exponential growth phase. Single-cell suspensions containing $1x10^6$ cells were fixed using 75% alcohol. The cell cycle was monitored using propidium iodide (PI) as it could successfully stain the nuclei. Finally, the fluorescence of DNA-bound PI in cells was measured using a FACScan flow cytometer (BD Biosciences): the results obtained using this device were then analyzed with ModFit 3.0 software (Verity Software House, Topsham, ME). The experiments were performed in triplicate.

In vitro cell growth assay. The cells were prepared at a concentration of $1x10^4$ cells/ml, respectively. Aliquots (100 μl) were dispensed in a 96-well microtiter plates. The cells were incubated for 1, 2, 3, 4, 5 and 6 days, respectively. Three-(4.5-dimethylthiazol-2-yl)-2.5-diphenyltetrazolium bromide (MTT) assay was performed by adding 20 μl of MTT (5 mg/ml; Promega, Madison, WI, USA) for 4 h. When MTT incubation was completed, the supernatants were removed. Then, 150 ml of dimethyl sulfoxide (DMSO Sigma, St. Louis, MO, USA) as added to each well. After 15 min, the absorbance (OD) of each well was measured and the value was recorded using a microplate reader set at a wavelength of 570 nm. The experiments were performed in triplicate.

Plate clone formation assay. From each treatment group, nearly $1x10^2$ cells were added to each well (3 cm in diameter) of a 6-well culture plate. Each cell group consisted of three wells. After incubating at 37°C for 12 days, the cells were washed twice with PBS and stained with Giemsa solution. The number of colonies containing ≥50 cells was counted under a microscope [plate clone formation efficiency = (number of colonies/number of cells inoculated) x100%]. These experiments were performed in triplicate.

Tumor cell invasion assay. Warm serum-free medium was added to the top chamber of the cell invasion chamber (Chemicon, Temecula, CA, USA); thus rehydration of the ECM layer was performed for 2 h at room temperature. Tumor cells in serum-free medium (300 μl containing $1x10^5$ cells) were added to the top chamber. The bottom chamber was prepared using 10% FBS that serves as a chemo-attractant. After incubating for 48 h, we removed the non-invasive cells using a cotton swab. The cells that had migrated through the membrane and adhered to the lower surface of the membrane were fixed with methanol and stained with haematoxylin. For

Table I. Mutations found in MALAT-1 in CRC-derived cell lines and CRC tissues.

Gene	Mutation location	Cell lines or cancer tissues
MALAT-1/5434 nt-6951 nt	(6559 nt-6560 nt)delAA (6600 nt)G→A (6365 nt)T→C	SW620 cell line
MALAT-1/5434 nt-6951 nt	(6560 nt)delT	SW480 cell line
MALAT-1/6918 nt-8441 nt	(7592 nt)T→C (7930 nt)C→T (8349 nt)delT	SW480 cell line
MALAT-1/5434 nt-6951 nt	(6560 nt)delA	Primary colorectal cancer tissues
MALAT-1/6918 nt-8441 nt	(8349 nt)delT	Primary colorectal cancer tissues

quantification, the cells were counted under a microscope in 5 predetermined fields at a magnification of x200. The experiments were performed in triplicate.

Statistical analysis. The SPSS statistical software package, version 13.0 (Abbott Laboratories, North Chicago, IL, USA) was used for conducting the statistical analyses. RT-QPCR assay, flow cytometry assay, plate clone formation assay, and tumor cell invasion assay were tested using one-way ANOVA. In all samples, $P \leq 0.05$ was considered to be statistically significant. *In vitro* cell growth assay was tested using factorial design ANOVA.

Results

Identification of MALAT-1 mutation. Because MALAT-1 is very long, it was incised into five fragments (Fig. 1A). In this study, it was discovered that humans and mice had highly homologous sequence of MALAT-1 as far as the fragments 6528 nt-7960 nt and 8152 nt-8336 nt of MALAT-1 (Fig. 1B) were concerned. This observation was in good agreement with the results of bioinformatics analysis. By sequencing SW480 we explored the mutation of the five fragments of MALAT-1 in the following types of samples: CRC SW620 and SW480 cell lines (these had different metastatic potential), normal colorectal tissues and primary CRC tissues. First, RNA was extracted from the SW620 and SW480 CRC cell lines, normal tissues, and primary CRC tissues. Then, the five fragments of MALAT-1 were amplified by RT-touchdown PCR (Fig. 2A-C); the PCR products were cloned subsequently. Finally, TA cloned bacteria were sequenced. After conducting sequencing analysis and comparison with normal tissues, we found that the fragment 5434 nt-6951 nt of the MALAT-1 was mutated in SW620, whereas the other two fragments 5434 nt-6951 nt and 6918 nt-8441 nt of the MALAT-1 were mutated in SW480 cancer cells and the primary CRC tissues. Table I gives an overview of all mutations in the CRC cells and tissues.
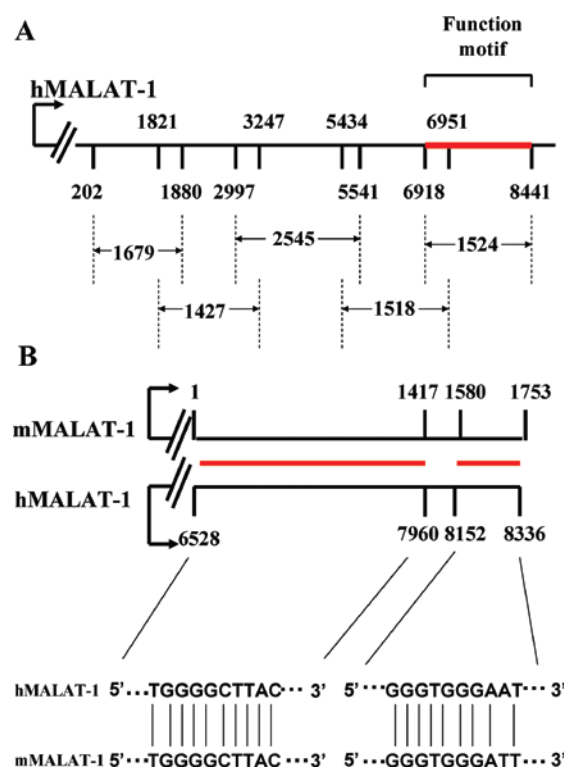

Figure 1. The location the five fragments and species homologous fragments of MALAT-1. (A) Schematic representation of the location of five fragments of MALAT-1. The red bar shows the function motif of MALAT-1. Below, the size of five fragments of MALAT-1. (B) Red bars represent the highly conservative domain in the seed-pairing sequence of MALAT-1 between human (Has) and mouse (Mmu).

Construction of eukaryotic expression vector for 5 fragments of MALAT-1 and its expression in human CRC SW480 cells. In order to generate 5 recombinant expression vectors, the cDNA sequences of the 5 MALAT-1 fragments, derived from the normal colorectal mucosa tissue, were cloned into eukaryotic expression

Table II. RT-touchdown PCR primers of five fragments of human MALAT-1.

Gene	Primer sequences (3'-5')	Amplification size (nt)
MALAT-1/202 nt-1880 nt	For: GTATATGAGCTCACCATGGCCATTCCAGGTGGTGGTATTTAG	1679
	Rev: GAACTCGGGCCACCATGCTCCAAATTGTTTCATCCTACCACTC	
MALAT-1/1821 nt-3247 nt	For: CGACGCGAGCTCACCATGACTAAACGCAGACGAAAATGGAAAG	1427
	Rev: TCGCTCGGGCCCACCATGACAAACTTGCCATCTACTATT	
MALAT-1/2997 nt-5541 nt	For: TGGCCTGAGCTCACCATGTGGTTGGTAAAAATCCGTGAGGTCG	2545
	Rev: TATATAGGGCCCACCATGTAGGGAAGGGGTCAGGAGAAAGTG	
MALAT-1/5434 nt-6951 nt	For: CATGACGAGCTCACCATGAGTTTAACTTGCATCTGCAGTATTG	1518
	Rev: CTGCAGGGCCCACCATGTTGAAGACTGTTGCTTGTTTGGAA	
MALAT-1/6918 nt-8441 nt	For: ACCTCCGAGCTACCATGCAAGAAACATTCCAAACAAGCAACAG	1524
	Rev: AGTCATGGGCACCATGCCTGAAGACAGATTAGTAGTCAAAGCA	

vector pEGFP-C1. 5 recombinant plasmids were transfected into human CRC SW480 cells using Lipofectamine-2000: these were then identified with real-time PCR (Q-PCR). This study was conducted on the human CRC SW480 cells. The data obtained in this study suggest that the five fragments of MALAT-1 exhibit over-expression with real-time quantitative PCR (Fig. 3).

Identifying the biological behavior of the five fragments of MALAT-1 in human CRC SW480 cells. To confirm the role of MALAT-1 in the biological processes of proliferation, invasion and metastasis of CRC, 5 fragments of MALAT-1 expression vectors were over-expressed in CRC cells SW480 *in vitro*. In this study, we were successful in detecting their biological function in proliferation, invasion and metastasis. As shown in the cell cycles, the cellular content in the G2 phase of the 5 segments was increased remarkably in the SW480 CRC cell line. In addition, the cellular content was the highest in the G2 phase of fragment 6918 nt-8441 nt (Fig. 4A) ($P<0.01$). Comparing the plate colony formation efficiency in the 5 fragments of the MALAT-1, we discovered that the formed colonies expressed these 5 segments of MALAT-1 in SW480 cells in a significantly different manner. It is interesting to note that the formed colonies expressing fragment 6918 nt-8441 nt were maximal (Fig. 4B) ($P<0.001$). In addition, MTT assay results consistently agree with the cell cycles and plate colony formation efficiency. Similarly, SW480 cells expressing fragment 6918 nt-8441 nt have the fastest proliferation (Fig. 4C) ($P<0.001$). To investigate the effect of MALAT-1 expression on CRC cell invasion, we conducted a Matrigel invasion assay. When compared with those transfected with the control, it was found that the quantitative analysis measuring the ability of cancer cell invasion showed that: the invasion activity was significantly increased in SW480 cells when the over-expressed 6918 nt-8441 nt fragment was located at the 3' end of MALAT-1 (Fig. 4D) ($P<0.001$).
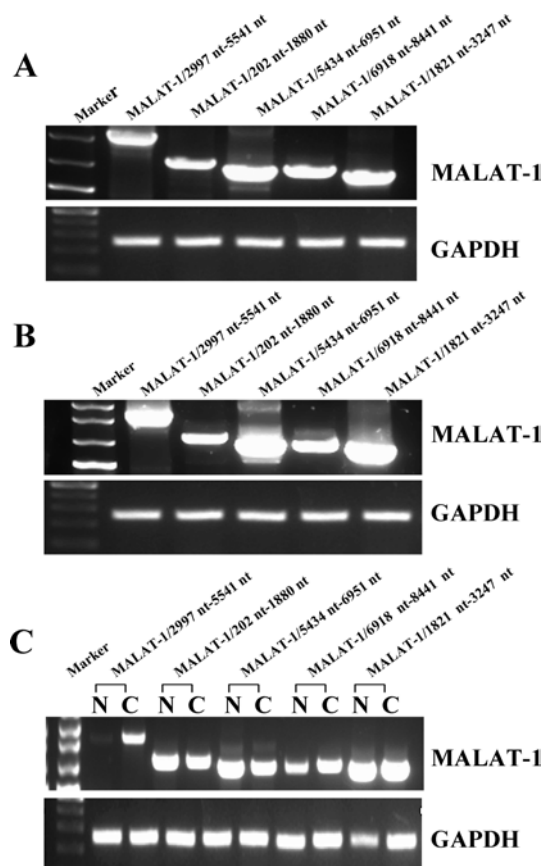

Figure 2. PCR products of the five fragments of MALAT-1 (fragment 1-fragment 5) by RT-touchdown PCR in human CRC cell lines (SW620 and SW480) and tumor tissues. (A) RT-touchdown PCR of the five segments of MALAT-1 gene in human CRC cell line SW620; GAPDH was amplified as control. (B) RT-touchdown PCR of the five segments of MALAT-1 gene in human CRC cell line SW480; GAPDH was amplified as control. (C) RT-touchdown PCR of the five segments of MALAT-1 gene in human normal colorectal tissue and human CRC tissue; GAPDH was amplified as control. All results are representative of three separate experiments.

Table III. Real-time PCR primers of five fragments of human MALAT-1.

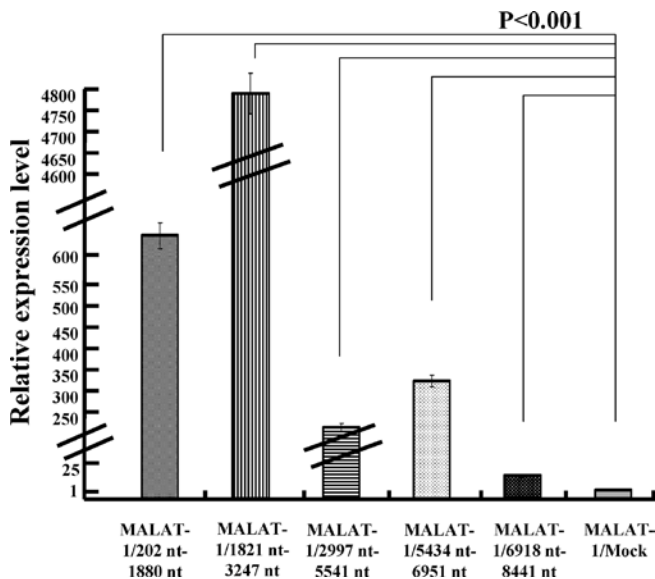| Gene	Primer sequence (3'-5')	Amplification size (nt)
MALAT-1/202 nt-1880 nt	For: TCATACCTAACCAGGCATAACA Rev: AAGTGCTCACAAGGCAAATC	206
MALAT-1/1821 nt-3247 nt	For: CGCATTTACTAAACGCAGAC Rev: TCTCTATTCTTTTCTTCGCC	133
MALAT-1/2997 nt-5541 nt	For: GTTACTCTTTTTTCCCCCCACCCCC Rev: TTCTCCCCCACCCTCTCTCTTCCCT	125
MALAT-1/5434 nt-6951 nt	For: GTTTGTCTGTAGTTCAGTGTTGGGG Rev: TAAAGTTACATTCGTTCTTCCGCTC	128
MALAT-1/6918 nt-8441 nt	For: CCAATGCTCTTCAGTAGGGTCA Rev: ACCAGGAGTGCCAACCACC	141

Figure 3. mRNA relative expression levels of the five segments of MALAT-1 after transfection and over-expression of the five segments of MALAT-1 in the human CRC cell line SW480. Each bar represents the mean ± SD. Representative data from three independent experiments are shown.

Discussion

MALAT-1 (also known as HCN, NEAT2, PRO2853 and NCRNA00 047) is located at chromosome 11q13.1: it is a poly-adenylated non-coding RNA (ncRNA) of ~8 kb (15). The aberrant expression of ncRNA has an important role in the process of cancer development (18). MALAT-1 is over-expressed in meta-stasis of the following types of cancer: non-small cell lung cancer, endometrial stromal sarcoma, and hepatocellular carcinoma (15,16). MALAT-1 is also expressed in some normal human tissues (15).

MALAT-1 ncRNA lacks credible open reading frames: it does not contain a valid Kozak sequence, thereby indicating that the likelihood of translation is quite impossible (15). Major part of MALAT-1 is present in an unspliced form. Significant

signal pathways have not yet been observed outside the nucleus (19). The function of this long ncRNA is not known. Its abundance and aberrant expression in many cancers suggests that it plays a pivotal role in the development of cancer (20).

On the basis of previous reports, several intracellular functions of ncRNA have been proposed. First, the long ncRNA plays an important role in pre-mRNA metabolism; it establishes an intimate association with SC35 splicing domains that are located within the mammalian nucleus (19). Second, it is concentrated in nucleoli as a riboregulator, controlling expression of its targeted gene (16). Third, it has been identified as a regulator of tumor suppressor proteins (e.g. PTB-associated splicing factor (PSF) (20). In recent studies, a highly conserved tRNA-like small RNA of 61 nucleotides, originating from the MALAT-1 locus (7519 nt to 7576 nt), was found exclusively in the cytoplasm. The 3' end processing mechanism is significant, because it is through that mechanism of a single gene locus yields both a stable nuclear-retained non-coding RNA with a short poly (A) tail-like moiety, and a small tRNA-like cytoplasmic RNA (hMALAT-1: 7519 nt-7542 nt; mMALAT-1: 6691 nt-6714 nt) (21). This 61 nt transcript is included in the 6918 nt-8441 nt (1524 nt) transcript that was found exclusively in our study. We suggest that this special structure can determine its very important function.

As predicted by bioinformatics analysis, we found a highly homologous sequence of MALAT-1 in both humans and mice and this sequence was located at their 3' end region of MALAT-1. Conventional theory states that homologous gene sequences of species are highly conserved in the biological evolution process. DNA sequences with functional importance are often conserved among species through negative selection (22-25). However, some studies have illustrated that a group of ultra-conserved ncRNAs is altered in the case of people detected with cancer (26).

In our study, based on the results of sequencing analysis, we deduce that the MALAT-1 RNA transcripts are mutated not only in human CRC cell lines (SW620 and SW480), but also in primary CRC tissues. Furthermore, our findings suggest that fragment 5434 nt-6951 nt of the MALAT-1 is mutated in the SW620 colorectal cell line with high metastatic potential.
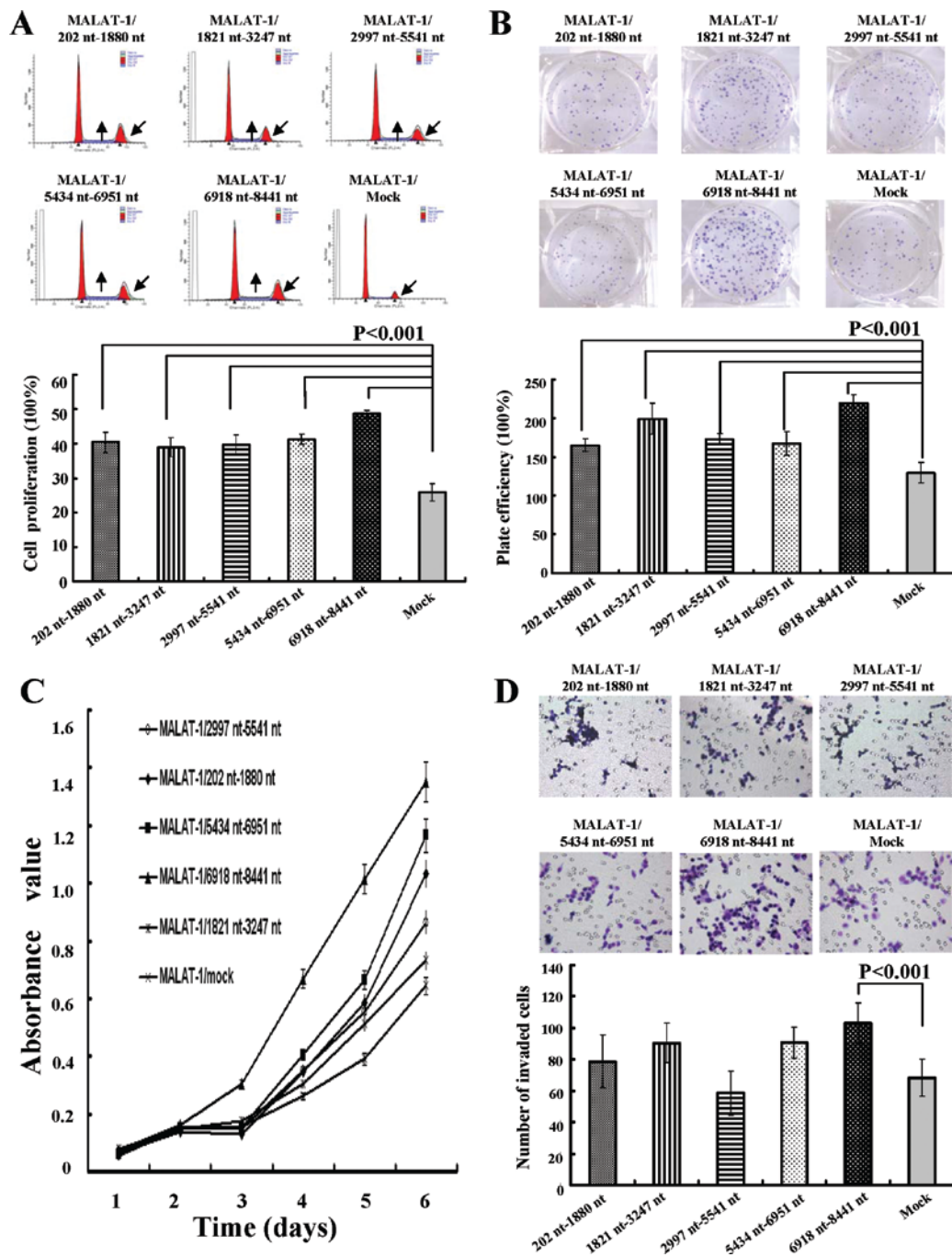
Figure 4. The proliferation effect and invasion ability of five segments of MALAT-1 after the over-expression of the five segments of MALAT-1 in the human CRC cell line SW480. (A) Analysis of the cell cycles of the transfected cells for the five segments of MALAT-1 by flow cytometry in the human CRC cell line SW480. (B) Analysis of the plate colony's formation efficiency of transfected cells for the five segments of MALAT-1 in human CRC cell line SW480. (C) MTT analysis of the proliferation rate of mock and the five fragments of MALAT-1 transfected cells in human CRC cell line SW480. Each value represents the mean ± SD of the absorbance value (OD) for cells. The difference in the proliferation ability between the control and these five segments of MALAT-1 over-expression-treated SW480 cells was significant ($^*P<0.001$). All results are representative of three separate experiments. (D) Representative images of invaded SW480 cells are shown at magnification x200. The difference in invasion ability between the control and segment 6918 nt-8441 nt of MALAT-1 over-expression-treated SW480 cells was significant ($^*P<0.001$). Each of the three independent experiments was performed in duplicate.

On the other hand, fragments 5434 nt-6951 nt and 6918 nt-8441 nt of the MALAT-1 are mutated in SW480 colorectal cell line with low metastatic potential. Moreover, fragments 5434 nt-6951 nt and 6918 nt-8841 nt of the MALAT-1 are mutated in primary CRC tissues. These mutations focus on species homologous fragments 6528 nt-7960 nt and 8152 nt-8336 nt of the MALAT-1. Our observations suggest that these mutations occurred consistently in CRC cells and tissues. This implies that these consistent mutations are not random mutations in unstable human CRC cells; instead, specific alterations play an important role in CRC development. On the basis of these observations, we speculate that these mutations occur in relatively conserved and species homologous fragments of MALAT-1; therefore, they possibly relate to CRC development and metastasis.

MALAT-1 is up-regulated in placenta percreta and is strongly associated with trophoblast-like cell invasion *in vitro* (27). MALAT-1 can promote cell migration and invasion. It also possibly plays an important role in cell invasion and metastasis. In this study, these roles have been confirmed. In order to determine the functional motif of the MALAT-1 in human CRC, 5 fragments of MALAT-1 mRNA transcripts were over-expressed in the human CRC cell line SW480 by the gene recombination technique. MALAT-1 mRNA transcripts originating from normal colorectal tissue were used for this purpose. We found that fragment 6918 nt-8441 nt, located at the 3' end of the MALAT-1, played an important role in the following biological processes: cell proliferation, migration, and invasion of human colorectal malignancies. Therefore, we suggest that this motif of the 3' end MALAT-1 gene (6918 nt-8841 nt) may be an important functional motif of the MALAT-1. In addition, this gene could possibly play a role in the development and metastasis in CRC.

On the basis of the data obtained in this study, we hypothesize that in the structure of long non-coding RNA MALAT-1, there are different functional motifs involved in the development and metastasis of CRC. The biological motif of MALAT-1, playing a pivotal role in metastasis of CRC, is generally located at the post-median of MALAT-1. The domain may have one, two, or even multiple motifs. Fragment 6918 nt-84411 nt of MALAT-1 is one of the most important regulation motifs. Its primary function is to promote cell proliferation and invasion. However, as mutations occur in this fragment, MALAT-1 will lose the ability to promote cell proliferation, invasion, and metastasis. In addition, as in a normal state, the fragment 5434 nt-6951 nt plays a pivotal role in maintaining cell's normal biological character and function. Because of mutation in this fragment, the MALAT-1 will lose normal regulation function, thereby promoting cancer development.

Because MALAT-1 has a significant role in CRC development and metastasis, MALAT-1 can possibly be used as a target for future drug development, and for the purpose of early diagnosis and prognosis of human CRC. Further research is needed to identify the signal pathways involved in the pathogenesis and metastasis of human CRC that are supposedly regulated by MALAT-1.

Acknowledgements

References

1. Thigpen KG: International sequencing consortium. Environ Health Perspect 112: A406, 2004.
2. Wang J, Zhang J, Zheng H, *et al*: Mouse transcriptome: neutral evolution of 'non-coding' complementary DNAs. Nature 431: 757, 2004.
3. Struhl K: Transcriptional noise and the fidelity of initiation by RNA polymerase II. Nat Struct Mol Biol 14: 103-105, 2007.
4. Ebisuya M, Yamamoto T, Nakajima M and Nishida E: Ripples from neighbouring transcription. Nat Cell Biol 10: 1106-1113, 2008.
5. Mattick JS: RNA regulation: a new genetics? Nat Rev Genet 5: 316-323, 2004.
6. Wilusz JE, Sunwoo H and Spector DL: Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 23: 1494-1504, 2009.
7. Christofori G: New signals from the invasive front. Nature 441: 444-450, 2006.
8. Beahrs OH: Staging of cancer of the colon and rectum. Cancer 70: 1393-1396, 1992.
9. Duffy MJ, van Dalen A, Haglund C, *et al*: Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use. Eur J Cancer 43: 1348-1360, 2007.
10. Fearon ER and Vogelstein B: A genetic model for colorectal tumorigenesis. Cell 61: 759-767, 1990.
11. Bernards R and Weinberg RA: A progression puzzle. Nature 418: 823, 2002.
12. Vogelstein B and Kinzler KW: Cancer genes and the pathways they control. Nat Med 10: 789-799, 2004.
13. Takayama T, Miyanishi K, Hayashi T, Sato Y and Niitsu Y: Colorectal cancer: genetics of development and metastasis. J Gastroenterol 41: 185-192, 2006.
14. Cardoso J, Boer J, Morreau H and Fodde R: Expression and genomic profiling of colorectal cancer. Biochim Biophys Acta 1775: 103-137, 2007.
15. Ji P, Diederichs S, Wang W, *et al*: MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene 22: 8031-8041, 2003.
16. Lin R, Maeda S, Liu C, Karin M and Edgington TS: A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. Oncogene 26: 851-858, 2007.
17. Chang JL, Li ZG, Wang XY and Yang MH: Detection of p53, MALAT1, ki-67 and β-catenin mRNA expression and its significance in molecular diagnosis of colorectal carcinoma. World Chinese J Digestol 16: 3849-3854, 2008.
18. Perez DS, Hoage TR, Pritchett JR, *et al*: Long, abundantly expressed non-coding transcripts are altered in cancer. Hum Mol Genet 17: 642-655, 2008.
19. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB and Chess A: A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics 8: 39, 2007.
20. Garen A and Song X: Regulatory roles of tumor-suppressor proteins and noncoding RNA in cancer and normal cell functions. Int J Cancer 122: 1687-1689, 2008.
21. Wilusz JE, Freier SM and Spector DL: 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell 135: 919-932, 2008.
22. Smith NG, Brandstrom M and Ellegren H: Evidence for turnover of functional noncoding DNA in mammalian genome evolution. Genomics 84: 806-813, 2004.
23. Pennacchio LA, Ahituv N, Moses AM, *et al*: In vivo enhancer analysis of human conserved non-coding sequences. Nature 444: 499-502, 2006.
24. Sironi M, Menozzi G, Comi GP, Cagliani R, Bresolin N and Pozzoli U: Analysis of intronic conserved elements indicates that functional complexity might represent a major source of negative selection on non-coding sequences. Hum Mol Genet 14: 2533-2546, 2005.
25. Pollard KS, Salama SR, King B, *et al*: Forces shaping the fastest evolving regions in the human genome. PLoS Genet 2: E168, 2006.
26. Calin GA, Liu CG, Ferracin M, *et al*: Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell 12: 215-229, 2007.
27. Tseng JJ, Hsieh YT, Hsu SL and Chou MM: Metastasis associated lung adenocarcinoma transcript 1 is up-regulated in placenta previa increta/percreta and strongly associated with trophoblast-like cell invasion in vitro. Mol Hum Reprod 15: 725-731, 2009.