J Hum Genet (2006) 51:1087–1099
DOI 10.1007/s10038-006-0070-9

ORIGINAL ARTICLE

Identification of a novel non-coding RNA, *MIAT*, that confers risk of myocardial infarction

Nobuaki Ishii · Kouichi Ozaki · Hiroshi Sato · Hiroya Mizuno · Susumu Saito · Atsushi Takahashi · Yoshinari Miyamoto · Shiro Ikegawa · Naoyuki Kamatani · Masatsugu Hori · Satoshi Saito · Yusuke Nakamura · Toshihiro Tanaka

Received: 14 August 2006 / Accepted: 31 August 2006 / Published online: 26 October 2006

Abstract Through a large-scale case-control association study using 52,608 haplotype-based single nucleotide polymorphism (SNP) markers, we identified a susceptible locus for myocardial infarction (MI) on chromosome 22q12.1. Following linkage disequilibrium (LD) mapping, haplotype analyses revealed that six SNPs in this locus, all of which were in complete LD, showed markedly significant association with MI (χ^2= 25.27, P=0.0000005; comparison of allele frequency, 3,435 affected individuals versus 3,774 controls, in the case of intron 1 5,338 C>T; rs2331291). Within this locus, we isolated a complete cDNA of a novel gene, designated myocardial infarction associated transcript (*MIAT*). *MIAT* has five exons, and in vitro translation assay showed that *MIAT* did not encode any translational product, indicating that this is likely to be a functional RNA. In vitro functional analyses revealed that the minor variant of one SNP in exon 5 increased transcriptional level of the novel gene. Moreover, unidentified nuclear protein(s) bound more intensely to risk allele than non-risk allele. These results indicate that the altered expression of *MIAT* by the SNP may play some role in the pathogenesis of MI.

Keywords Case-control association study · Myocardial infaraction associated transcript · Novel gene · SNP analysis

N. Ishii · K. Ozaki · T. Tanaka (✉)
Laboratory for Cardiovascular Diseases,
SNP Research Center, The Institute of Physical and Chemical Research (RIKEN), 4-6-1 Shirokanedai,
Minato-ku, Tokyo 108-8639, Japan
e-mail: toshitan@ims.u-tokyo.ac.jp

N. Ishii · Satoshi Saito
Division of Cardiology, Department of Medicine,
Nihon University School of Medicine, Tokyo, Japan

H. Sato · H. Mizuno · M. Hori
Department of Cardiovascular Medicine,
Osaka University Graduate School of Medicine,
Suita, Japan

Susumu Saito · Y. Nakamura
Laboratory for Genotyping, SNP Research Center,
The Institute of Physical and Chemical Research (RIKEN),
Tokyo, Japan

A. Takahashi · N. Kamatani
Laboratory for Statistical Analysis, SNP Research Center,
The Institute of Physical and Chemical Research (RIKEN),
Tokyo, Japan

Y. Miyamoto · S. Ikegawa
Laboratory for Bone and Joint Disease,
SNP Research Center, The Institute of Physical and Chemical Research (RIKEN), Tokyo, Japan

Introduction

Coronary artery diseases (CAD) including myocardial infarction (MI) have been the major cause of mortality and morbidity among late-onset diseases in many industrialized countries with a Western lifestyle (Breslow 1997; Braunwald 1997). MI often occurs without any preceding clinical signs, and is followed by severe complications, especially ventricular fibrillation and cardiac rupture which result in sudden death. Although recent advances in the treatment and diagnosis of MI have improved the quality of life for MI patients, its morbidity is still high.

Epidemiological studies revealed a variety of coronary risk factors, including type 2 diabetes mellitus,

hypercholesterolemia, hypertension, and obesity. There are also studies reporting a genetic factor of this disorder; one reported first-degree relatives of patients who have had an acute MI prior to age 55 years have 2–7 times higher risk of MI (Lusis et al. 2004). A twin study indicated an eight-fold increase in risk of death from MI when a first twin dies of MI before age 55 years (Marenberg et al. 1994). In this context, common genetic variants are considered to contribute to genetic risks of common diseases (Lander 1996; Risch and Merikangas 1996; Collins et al. 1997).

To date, various genetic variants that confer susceptibility to MI have been indicated to be present on several chromosomal loci through linkage analyses or case-control association studies using single nucleotide polymorphism (SNP) (Topol et al. 2001; Yamada et al. 2002; Ozaki et al. 2002, 2004, 2006; Stenina et al. 2003; Helgadottir et al. 2004). Case-control association study by means of genome-wide SNP analysis is one of the most powerful approaches to identify genetic variants susceptibility to common diseases.

We report here identification of SNPs in myocardial infarction associated transcript (*MIAT*) that were associated with MI through our comprehensive SNP association study. We also demonstrated the possible transcriptional effect of this variation on the expression level of *MIAT*.

Materials and methods

DNA samples

This study included 3,464 Japanese individuals with myocardial infarction who were referred to Osaka Acute Coronary Insufficiency Study Group, which involved the cardiovascular units of 25 hospitals in Osaka. The diagnosis of definite myocardial infarction has been previously described (Ohnishi et al. 2000; Ozaki et al. 2002). The control individuals consisted of 3,819 members of the general population who were recruited through several medical institutes in Japan. DNAs were prepared from these samples according to standard protocols. All individuals were Japanese, gave written informed consent to participate in the study, or their parents gave them when they were under 20 years old, according to the process approved by the relevant Ethical Committee at SNP Research Center, The Institute of Physical and Chemical Research (RIKEN) Yokohama.

SNP discovery and genotyping

Protocols for PCR primers, PCR experiments, DNA extraction, DNA sequencing and genotyping of SNPs have been previously described (Iida et al. 2001). For SNP discovery, genomic DNAs from 24 Japanese individuals were used, and direct sequencing was performed using capillary sequencer (ABI3700; Applied Biosystems, Foster City, Calif., USA).

Statistical analysis

We assessed association and Hardy–Weinberg equilibrium by χ^2-test (Yamada et al. 2001; Ozaki et al. 2002). Haplotype block and haplotype frequencies were estimated using SNPAlyze software (DYNACOM, Chiba, Japan) and Haploview v3.2 (Barrett et al. 2005).

Northern blot analysis

Human multiple-tissue Northern (MTN) blots (Clontech, Palo Alto, Calif., USA) were pre-hybridized and hybridized with α-[^{32}P]-dCTP labeled cDNA fragment prepared by PCR using primer pair shown in Table 1, as a probe. Washed membranes were autoradiographed for 7 days at –80°C.

Isolation of full-length cDNA

A human fetal brain cDNA library was constructed with combination of gene specific linker primers, random hexamer linker primer and oligo(dT) linker primer using ZAP cDNA synthesis kit (Stratagene, La Jolla, Calif., USA) according to the manufacturer's protocol. The library was screened with the same probe as Northern experiment. Positive clones were selected and their insert cDNAs were excised in vivo in pBluescriptIISK(-) (Stratagene) according to the manufacturer's protocol. To obtain the missing 5′- or 3′-portion, we performed a rapid amplification of cDNA ends (RACE) using BD SMART RACE cDNA Amplification Kit (Clontech) according to the manufacturer's instructions. Primers for full-length cDNA isolation were shown in Table 1.

In vitro translation assay

Four kinds of plasmids corresponding to variant 1–4, which were obtained through screening of ZAP human fetal brain cDNA library for isolation of *MIAT*

Table 1 Primer sequences used in this study

	Forward or sense	Reverse or antisense
Probe preparation	CAGGAATGCATCTCTGGCTC	CTCCTATCTCTTGCTACATTCC
cDNA library screening		
5′-gene specific linker (GSP) primer	GAGAGAGAGAGAGAGAGAGAACTAGTC-TCGAGCAGTTGATGAAGACAGCACAG	
3′-GSP linker primer	GAGAGAGAGAGAGAGAGAGAACTAGTC-TCGAGCTGTGCTGTCTTCATCAACTG	
Random hexamer	GAGAGAGAGAGAGAGAGAGAACTAGT-CTCGAGNNNNNN	
5′-RACE		
GSP-1	CCACATGGGTGGACAGATGCCTCC	
Nested GSP-1	CTAGTGCAGAGGGCTCTTTGTGCC	
GSP-2	ACGGGGTCATGGTGGCCACATGAAC	
Nested GSP-2	AACCACATGGGTGGACAGATGTCCTCC	
In vitro translation assay	ATCCTCGAGACAAAGAGCCCTCTGCACTAG	ATCGGATCCGAGCAAATG-GAGACAAAGGAC
Luciferase assay		
Intron 1 5,338 C>T FP	ATCACGCGTCCAGAGTCAGGGAAAAAGACC	ATCCTCGAGTTGAATTCTACCATTTTCT-TACATC
Exon 3 8,813 G>A and exon 3 9,186 G>A	ATCACGCGTGTCTTCTCCAGACTGGTGAC	ATCCTCGAGCTCCCTGCTGAAGAA-AGAAGG
Exon 5 11,093 G>A	ATCACGCGTCTTGATTCCTCCTCGCGTTTC	ATCCTCGAGGAGCAAATG-GAGACAAAGGAC
Exon 5 11,741_G	CGCGTCTATTCATCAGCTTGGGCTGC	TCGAGCAGCCCAAGCTGATGAATAGA
Exon 5 11,741_A	CGCGTCTATTCATCAACTTGGGCTGC	TCGAGCAGCCCAAGTTGATGAATAGA
Exon 5 12,311_C	CGCGTTGCAGGGAGGCGGAGGCCCTGC	TCGAGCAGGGCCTCCGCCTCCCTGCAA
Exon 5 12,311_T	CGCGTTGCAGGGAGGTGGAGGCCCTGC	TCGAGCAGGGCCTCCACCTCCCTGCAA
Gel shift assay		
Exon 5 11,741_G	GTGGTGTGTACTATTCATCAGCTTG-GGCTGCCA	GCTGTGGCAGCCCAAGCTGATGA-ATAGTACACA
Exon 5 11,741_A	GTGGTGTGTACTATTCATCAACT-TGGGCTGCCA	GCTGTGGCAGCCCAAGTTGATGA-ATAGTACACA

N represents a mixture of nucleotides A, C, G and T

full-length cDNA, were transcribed and translated using TNT T7 Quick Coupled Transcription/Translation Systems (Promega, Madison, USA) and Transcend Biotin-Lysyl-tRNA (Promega) according to the manufacturer's protocols. For negative control experiments, the antisense plasmids were constructed by inserting each PCR-amplified variant product in the opposite direction to pBluescript SK(+). Primers for constructing of the antisense plasmids were shown in Table 1. After SDS-PAGE and electro-blotting, the biotinylated products were visualized using the Transcend Nonradioactive Translation Detection Systems (Promega).

Luciferase assay

To investigate functions of six SNPs, we cloned genomic fragments into pGL3-promoter vector (Promega) in the 5′–3′ orientation. For intron 1 5,338 C>T and exon 5 11,093 G>A SNPs, each PCR-amplified product was used to be cloned into each of the vectors; for two SNPs in exon 3, one single PCR-amplified product containing both of the SNP loci was used; for exon 5 11,741 G>A and exon 5 12,311 C>T SNPs, double stranded oligonucleotides was cloned into pGL3-promoter vector. Primers and oligonucleotides for cloning of luciferase construct were shown in Table 1. We grew HEK293 cells (RIKEN Cell Bank, Wako, Japan) in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. We then performed luciferase assay according to the manufacturer's protocol. After 24 h of the transfection, we lysed the cells in passive lysis buffer (Promega) and then measured luciferase activity using the Dual-Luciferase Reporter Assay System (Promega).

Electrophoretic mobility-shift assay (EMSA)

We prepared nuclear extract from HEK293 cells as previously described (Andrews and Faller 1991) and then incubated it with 33-bp oligonucleotides labeled with digoxigenin-11-ddUTP using the Dig Gel Shift Kit, second generation (Roche, Mannheim, Germany). The oligonucleotides used in EMSA experiments were shown in Table 1. The reaction was done with 1/10 volume of Poly [d(I-C)]. For competition studies, we

pre-incubated nuclear extract with unlabeled oligonucleotides (200-fold excess) before adding digoxigenin-labeled oligonucleotide. We separated the protein-DNA complexes on a non-denaturing 6% polyacrylamide gel in 0.5× Tris–Borate-EDTA buffer. We transferred the gel to nitrocellulose membrane and detected the signal with a chemiluminescent detection system (Roche) according to the manufacturer's instructions.

In vitro RNA stability assay

First, we genotyped five SNP loci using DNAs extracted from 20 different human B cell lines (HEV cell lines; RIKEN cell bank). For further RNA stability assay, we used one cell line that was heterozygous for all of these SNPs. Two DNA fragments, corresponding to nt 956–1,827 and nt 1,975–3,400 of *MIAT* variant 4 (GenBank accession number; AB_263417) were amplified by PCR and used as a template. These fragments were cloned into a pBlueScript II SK(+) in the 5′–3′ orientation. We purified the vectors and performed transcription reaction using MEGAscript high yield transcription kit (T3 kit) (Ambion) according to the manufacturer's instruction. Using 3 μl of the transcription products and 3 μl of HEK293 cell extracts (diluted 20-folds with distilled water), we performed RNA stability assay as described previously (Suzuki et al. 2003).

Results

Large-scale SNP association study

We performed a large scale case-control association study using 188 MI patients and 752 general Japanese population by our high-throughput multiplex PCR-Invader assay method (Ohnishi et al. 2001) for 52,608 gene-based SNPs selected from the JSNP database on the basis of the haplotype block structure reported previously (Haga et al. 2002; Tsunoda et al. 2004). Detailed screening strategy and results will be described elsewhere (Ebana et al., in preparation). From this first-stage screening, we identified one SNP (rs2301523) in *FLJ25967* (GenBank accession number, AK098833) on chromosome 22q12.1 to reveal a significant association with MI (P=0.0006). Further investigation of this SNP using a total of 3,464 MI patients and 3,819 general population confirmed the association with MI with a χ^2 value of 22.71 (P=0.0000019; comparison of allele frequency) and odds ratio of 1.36 (95% confidence interval (CI); 1.20–1.55, Table 2).

Table 2 Association analyses between MI and six SNPs in *MIAT*

SNP position	MI (%)				Control (%)				Comparison of allele frequency		
	11	12	22	Sum	11	12	22	Sum	χ^2	P value	Odds ratio (95% CI)
Intron1 5,338 C>T	2,885 (84.0)	528 (15.4)	22 (0.6)	3,435	3,320 (88.0)	443 (11.7)	11 (0.3)	3,774	25.27	0.0000005	1.38 (1.22–1.57)
Exon 3 8,813 G>A	2,926 (85.1)	500 (14.5)	14 (0.4)	3,440	3,308 (88.9)	406 (10.9)	8 (0.2)	3,722	23.22	0.0000014	1.38 (1.21–1.58)
Exon 3 9,186 G>A	2,919 (84.3)	528 (15.2)	17 (0.5)	3,464	3,364 (88.1)	445 (11.7)	10 (0.3)	3,819	22.71	0.0000019	1.36 (1.20–1.55)
Exon 5 11,093 G>A	2,885 (84.3)	523 (15.3)	15 (0.4)	3,423	3,314 (88.1)	438 (11.6)	8 (0.2)	3,760	22.89	0.0000017	1.37 (1.20–1.56)
Exon 5 11,741 G>A	2,888 (84.2)	525 (15.3)	15 (0.4)	3,428	3,308 (88.1)	440 (11.7)	8 (0.2)	3,756	22.47	0.0000021	1.36 (1.20–1.55)
Exon 5 12,311 C>T	2,896 (84.3)	525 (15.3)	15 (0.4)	3,436	3,315 (88.0)	442 (11.7)	8 (0.2)	3,757	21.85	0.0000030	1.36 (1.19–1.56)

CI confidence interval

11, 12 and 22 indicate homozygote of major allele, heterozygote and homozygote of minor allele, respectively

Identification of the full-length gene

Since cDNA sequences of 1,713 bases for FLJ25967 (GenBank Accession code; AK098833) did not contain a long open reading frame (ORF), we first examined expression of *FLJ25967* in 4 fetal and 16 adult human tissues by Northern blot analyses. As shown in Fig. 1, the transcript of approximately 10 kb in length was detected in spleen, peripheral blood leukocyte, lung, liver, thymus, colon and small intestine, although predominant signals were detected in fetal brain and brain. To obtain the full-length cDNA sequence for this gene, we screened the fetal brain cDNA library and subsequently carried out 5′- and 3′-RACE experiments. From the results of the full-length cDNA sequences, this gene was considered to have four splicing variants (GenBank Accession number: AB263414, AB263415, AB263416, and AB263417), each consisting of 10,142, 10,016, 10,068 and 9,942 nucleotides, respectively. However, we still found no long ORF in either of the four cDNA sequences; the longest ORF we identified was 447-base long encoding 149 amino acids (Fig. 2b). Although we compared this ORF with protein sequecnes in the public databases using the BLAST program (Altschul et al. 1990), we found no protein showing a significant similarity. To investigate whether the transcript was really translated into protein, we carried out in vitro translation assay for each of the four variants, but no translated product was detected (Fig. 3), suggesting this gene encodes a functional RNA. We designated this gene *MIAT*, myocardial infarction associated transcript.

A comparison of genomic DNA sequences in GenBank database with each of the cDNA sequences determined a genomic structure of the *MIAT* gene. This gene consists of five exons (Fig. 2a) and all the splice junctions were considered to conform to the basic GT/AG rule (Mount 1982; Shapiro and Senapathy 1987).

Linkage disequilibrium and haplotype analysis

To search for the possibility that another SNP(s) in this locus confers risk of MI, we investigated SNPs in this region thoroughly, except for the regions corresponding to repetitive sequences. By direct sequencing of genomic DNA from 24 individuals in Japanese population, we identified a total of 60 SNPs, including 14 novel ones which were not registered in the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/index.html; build 126, Table 3). Subsequently, we selected 35 SNPs on the basis of the following two conditions; (1) SNPs should show minor allele frequencies (MAF) >5%, and (2) only one SNP can be selected from a group of SNPs that are in perfect LD among them. Then, we genotyped 96 individuals with MI and investigated a precise haplotype structure in this region using Haploview software (http://

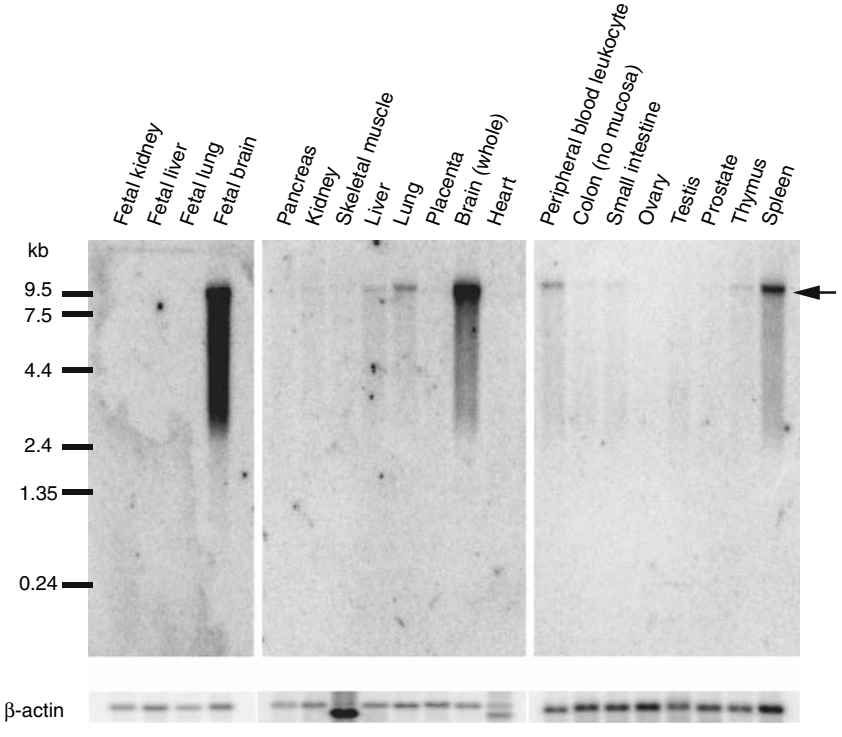
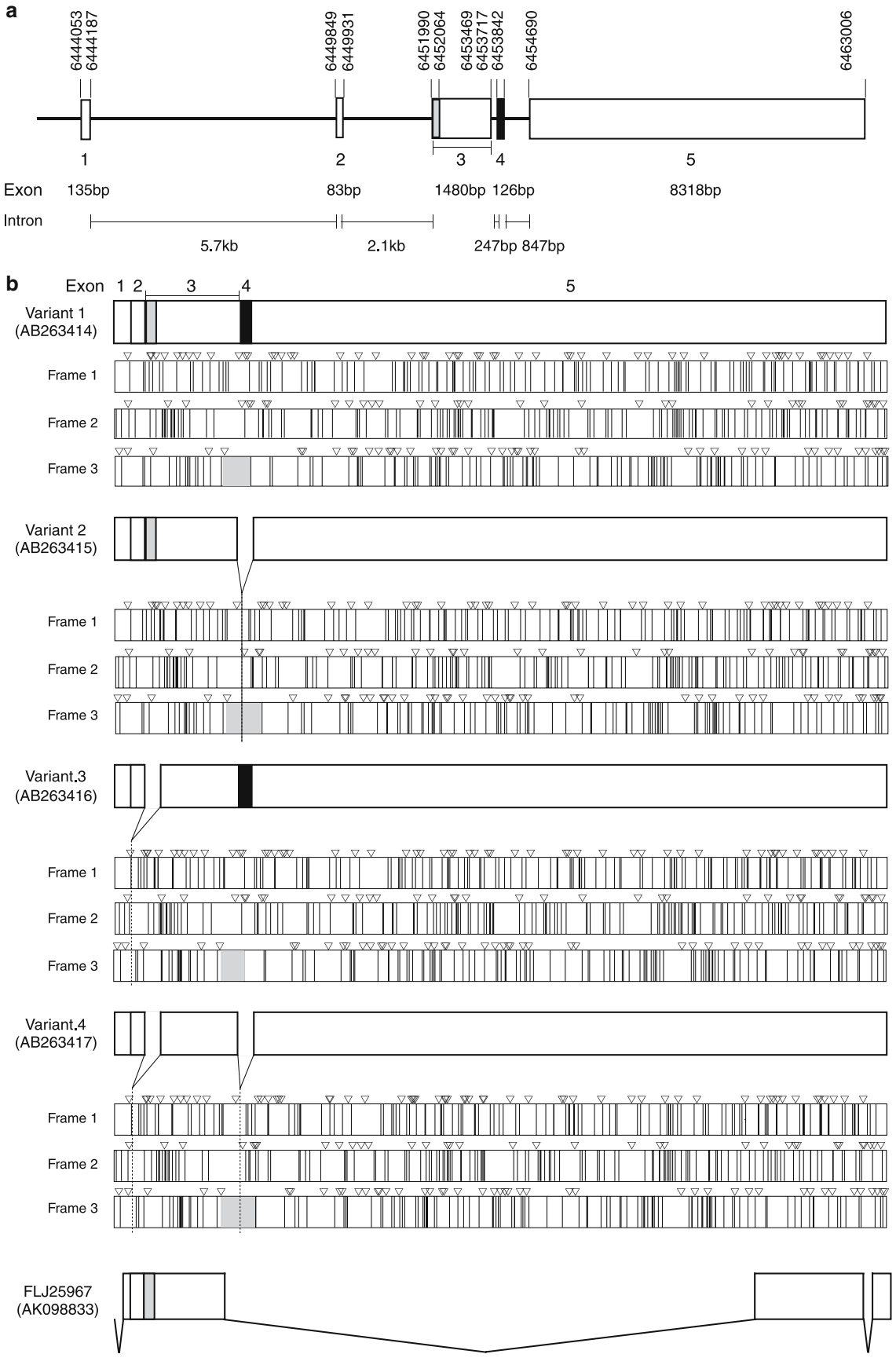


Fig. 1 Expression of *MIAT* in human tissues. *Arrows* indicate the transcript of approximately 10 kb in size. *β*-actin cDNA was used as the quantity control


a
6444053
6444187
6449849
6449931
6451990
6452064
6453469
6453717
6453842
6454690
6463006
1
2
3
4
5
Exon 135bp 83bp 1480bp 126bp 8318bp
Intron 5.7kb 2.1kb 247bp 847bp
b
Exon 1 2 3 4 5
Variant 1 (AB263414)
Frame 1
Frame 2
Frame 3
Variant 2 (AB263415)
Frame 1
Frame 2
Frame 3
Variant.3 (AB263416)
Frame 1
Frame 2
Frame 3
Variant.4 (AB263417)
Frame 1
Frame 2
Frame 3
FLJ25967 (AK098833)

◀ **Fig. 2** Genomic organization and splicing variants of *MIAT*. **a** Genomic structure of *MIAT*. Base-pair numbering at the top was based on an entry from GenBank DNA database (NT011520.10). **b** Four splicing variants of *MIAT* and their possible open reading frames. The longest ORF was shown in gray. *Inverted triangle* and *vertical line* indicate ATG initiation codon and stop codon, respectively

www.hapmap.org). As shown in Fig. 4, we identified one haplotype block including the marker SNP rs2301523, which was first identified to have an association with MI through our large-scale study. In addition, we found five SNPs (intron 1 5,338 C>T, exon 3 8,813 G>A, exon 5 11,093 G>A, exon 5 11,741 G>A, exon 5 12,311 C>T) in the block in strong linkage disequilibrium (LD) with the marker SNP. Therefore, we examined whether these five SNPs were also tightly associated with susceptibility to MI by genotyping approximately 3,400 individuals with MI and 3,700 control subjects. As shown in Table 2, all of them showed significant associations with MI. In particular, one SNP (intron 1 5,338 C>T; rs2331291) presented the most significant association with MI (χ^2=25.27, P=0.0000005; comparison of allele frequency).

To clarify a possibility that one particular haplotype in the block confers risk of MI, we selected five tag SNPs (exon 3 8,555 T>C, exon 3 9,186 G>A, exon 5 10,804 G>A, exon 5 14,569 A>T, exon 5 15,219 C>A) that could cover 90% of haplotypes in this block. We compared frequency of haplotypes and genotypes of 652 individuals with MI and 620 control subjects using these tag SNPs. Since neither haplotypes nor four other tag SNPs showed statistical significance for the association with MI (Tables 4, 5), we considered that the first marker SNP rs2301523 and additional five SNPs showing the strong LD with it were candidate SNPs associated with MI.

We investigated relationship between patients' genotype information and clinical profiles including diabetes, hypertension, smoking, hyperlipidemia, sex, and age by one-way ANOVA and χ^2 test. Since we could not find positive association for any of the coronary risk factors, age, or sex (data not shown), we concluded our findings are directly related to the pathogenesis of MI.

Luciferase and Gel-shift assay

Since *MIAT* was considered to be a non-coding functional RNA, we investigated the functions of these SNPs by examining their effect on transcriptional regulation by luciferase assay using HEK293 cell. As shown in Fig. 5d, the clone containing a G allele at position 11,741 in exon 5 had approximately 1.3-fold greater transcriptional activity than an A allele or the vector only. However, the remaining five SNPs did not show statistical difference in their transcriptional activity (P>0.05; Fig. 5a–c, e). Subsequently, to examine whether a nuclear factor(s) might bind to genomic sequence around the exon 5 11,741 SNP, we searched for binding motifs for known transcription factors around this SNP sequence by TFSEARCH program (http://www.cbrc.jp/research/db/TFSEARCHJ.html) based on the TRANSFAC database (Heinemeyer et al. 1998). Although no motif for binding of the transcriptional factors was predicted, we attempted to examine binding of some nuclear factor(s) in nuclear extract from HEK293 cells to oligonucleotides corresponding to the 11,741 G and 11,741 A alleles, respectively. As shown in Fig. 5f, the A-allele oligonucleotide bound more tightly to some nuclear factor present in HEK293 cells than the G-allele oligonucleotide.

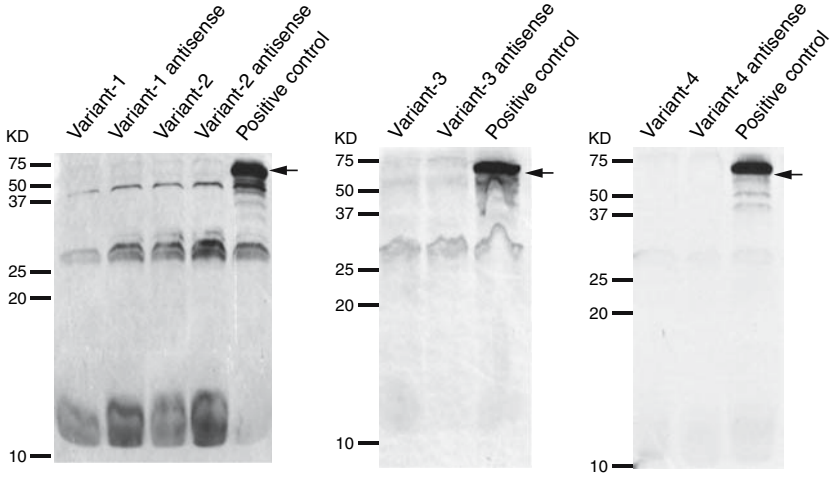


Fig. 3 In vitro translation assay The *arrows* indicate the bands for positive controls

Table 3 SNPs in *MIAT*

db SNP	SNP position	Location in Z99774.1	Allele	MAF
rs5761663	5′-flanking –1,150	22,542	G/C	0.271
–	5′-flanking –1,035	22,657	C/T	0.063
rs5752375	5′-flanking –976	22,716	C/T	0.063
–	5′-flanking –927	22,765	C/T	0.063
rs9608515	5′-flanking –902	22,790	C/T	0.063
rs2157598	5′-flanking –594	23,098	A/T	0.341
rs5761664	5′-flanking –364	23,328	C/G	0.333
rs8142890	5′-flanking –106	23,586	C/T	0.063
rs16982564	Intron 1 696	24,387	G/A	0.271
rs7284198	Intron 1 716	24,407	T/C	0.063
rs5752376	Intron 1 1,349	25,040	G/A	0.063
–	Intron 1 1,506	25,197	G/A	0.333
rs9613239	Intron 1 1,933	25,624	G/A	0.188
–	Intron 1 2,035	25,726	T/G	0.271
rs6005116	Intron 1 2,614	26,305	G/A	0.056
rs9613240	Intron 1 2,851	26,542	A/G	0.065
rs1007476	Intron 1 3,166	26,857	G/T	0.326
rs11912702	Intron 1 3,280	26,971	C/G	0.286
rs1981493	Intron 1 3,923	27,614	G/A	0.5
rs5761665	Intron 1 3,956	27,647	G/T	0.125
rs5752377	Intron 1 3,967	27,658	T/A	0.146
rs6005119	Intron 1 4,504	28,195	A/C	0.435
–	Intron 1 5,042	28,733	C/T	0.271
rs2331291	Intron 1 5,338	29,029	C/T	0.188
rs5761666	Intron 1 5,522	29,213	T/G	0.313
–	Intron 2 5,917	29,608	T/C	0.042
rs5997112	Exon 3 8,308	31,999	G/A	0.478
rs1061541	Exon 3 8,454	32,145	T/C	0.043
rs1061540	Exon 3 8,555	32,246	T/C	0.479
rs12166567	Exon 3 8,639	32,330	G/C	0.313
–	Exon 3 8,813	32,504	G/A	0.188
rs2301523	Exon 3 9,186	32,877	G/A	0.208
rs2301524	Exon 3 9,312	33,003	C/T	0.333
rs4274	Exon 5 10,804	34,495	G/A	0.292
rs3752603	Exon 5 11,011	34,702	C/T	0.417
–	Exon 5 11,093	34,784	G/A	0.188
rs6005121	Exon 5 11,634	35,325	G/A	0.021
–	Exon 5 11,741	35,432	G/A	0.167
–	Exon 5 12,311	36,002	C/T	0.182
–	Exon 5 12,388	36,079	C/T	0.048
rs7985	Exon 5 13,052	36,743	A/G	0.45
rs3180936	Exon 5 13,083	36,774	A/T	0.45
rs4280	Exon 5 13,562	37,253	G/T	0.294
–	Exon 5 14,160	37,851	A/G	0.413
rs8142049	Exon 5 14,309	38,000	C/T	0.435
rs5761670	Exon 5 14,569	38,260	A/T	0.25
–	Exon 5 14,588	38,279	G/A	0.042
rs9625066	Exon 5 15,219	38,910	C/A	0.271
–	Exon 5 15,515	39,206	C/T	0.042
rs9620625	Exon 5 15,771	39,462	G/A	0.022
rs5761672	Exon 5 15,776	39,467	G/A	0.386
rs6005127	Exon 5 16,074	39,765	C/T	0.273
rs5752378	Exon 5 16,352	40,043	G/A	0.176
rs1018834	Exon 5 16,244	40,115	T/A	0.292
rs713720	Exon 5 16,968	40,659	G/C	0.021
rs4822768	Exon 5 17,266	40,957	T/C	0.261
rs7293223	Exon 5 18,211	41,902	G/A	0.262
rs3747138	Exon 5 18,529	42,220	A/G	0.063
rs1045362	Exon 5 18,657	42,348	A/G	0.292
rs1045363	Exon 5 18,663	42,354	C/T	0.25

MAF minor allele frequency

'–' The variant was not registered in dbSNP database (build 126)

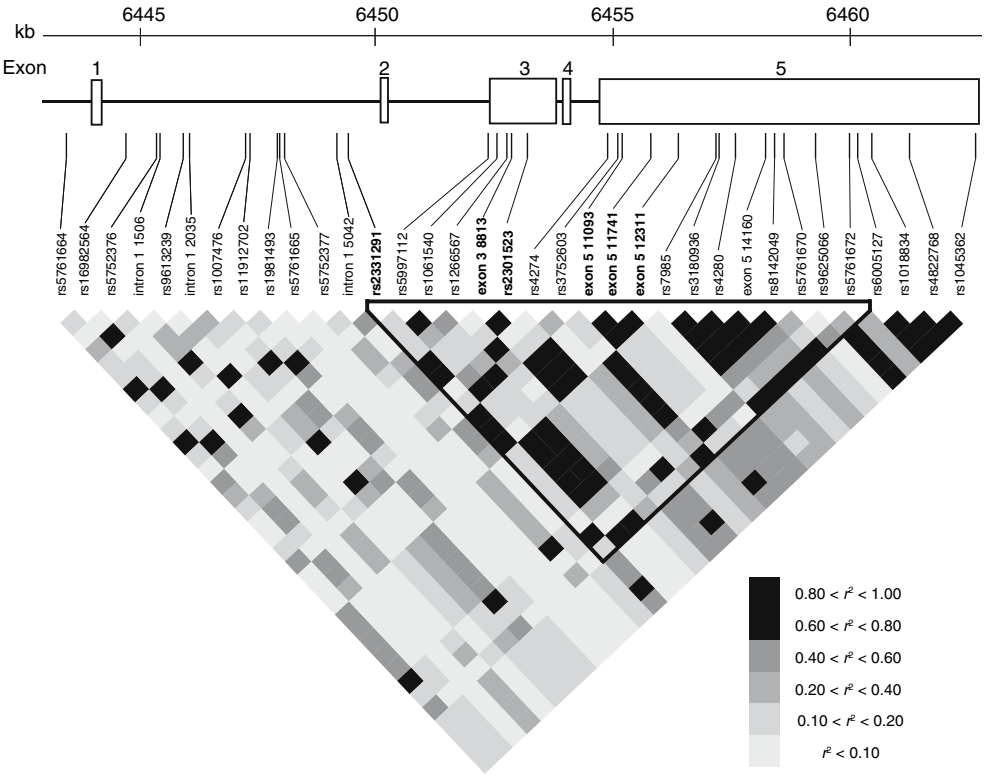


Fig. 4 Linkage disequilibrium structure at the *MIAT* locus SNPs identified by this study are shown. SNPs in *bold* indicate significant association with MI (see also Table 2). Those which were not deposited in dbSNP database (build 126) were labeled with their location in the gene (e.g., intron 1 5,042). Pairwise r^2 values for all combination of SNP pairs are shown in *gray scale*

Table 4 Association analyses of the tag SNPs in the haplotype block

dbSNP	SNP position	MI (%)				Control (%)				Comparison of allele frequency	
		11	12	22	Sum	11	12	22	Sum	χ^2	*P* value
rs1061540	exon 3 8,555 T>C	182 (27.9)	288 (44.2)	182 (27.9)	652	173 (27.9)	291 (46.9)	156 (25.2)	620	0.48	0.49
rs2301523	exon 3 9,186 G>A	534 (81.9)	115 (17.6)	3 (0.5)	652	545 (87.9)	73 (11.8)	2 (0.3)	620	8.89	0.0039
rs4274	exon 5 10,804 G>A	258 (39.6)	268 (41.1)	126 (19.3)	652	237(38.2)	285 (46.0)	98 (15.8)	620	0.31	0.57
rs5761670	exon 5 14,569 A>T	356 (54.6)	237 (36.3)	59 (9.0)	652	326 (52.6)	244 (39.4)	50 (8.1)	620	0.09	0.77
rs9625066	exon 5 15,219 C>A	437 (67.0)	188 (28.8)	27 (4.1)	652	431 (69.5)	169 (27.3)	20 (3.2)	620	1.26	0.26

11, 12 and 22 indicate homozygote of major allele, heterozygote and homozygote of minor allele, respectively

Table 5 Haplotype analyses

Haplotype ID	Tag SNPs					Haplotype frequency		Comparison of haplotype frequency	
	rs1061540	rs2301523	rs4174	rs5761670	rs9625066	MI	Control	χ^2	*P* value
Haplotype 1	T	C	A	A	C	0.361	0.383	1.30	0.25
Haplotype 2	C	C	G	T	C	0.259	0.272	0.65	0.42
Haplotype 3	C	C	G	A	C	0.154	0.160	0.19	0.66
Haplotype 4	T	T	G	A	A	0.052	0.053	0.01	0.90
Haplotype 5	T	C	G	A	A	0.049	0.062	2.06	0.15
Haplotype 6	C	C	G	A	A	0.042	0.047	0.32	0.57

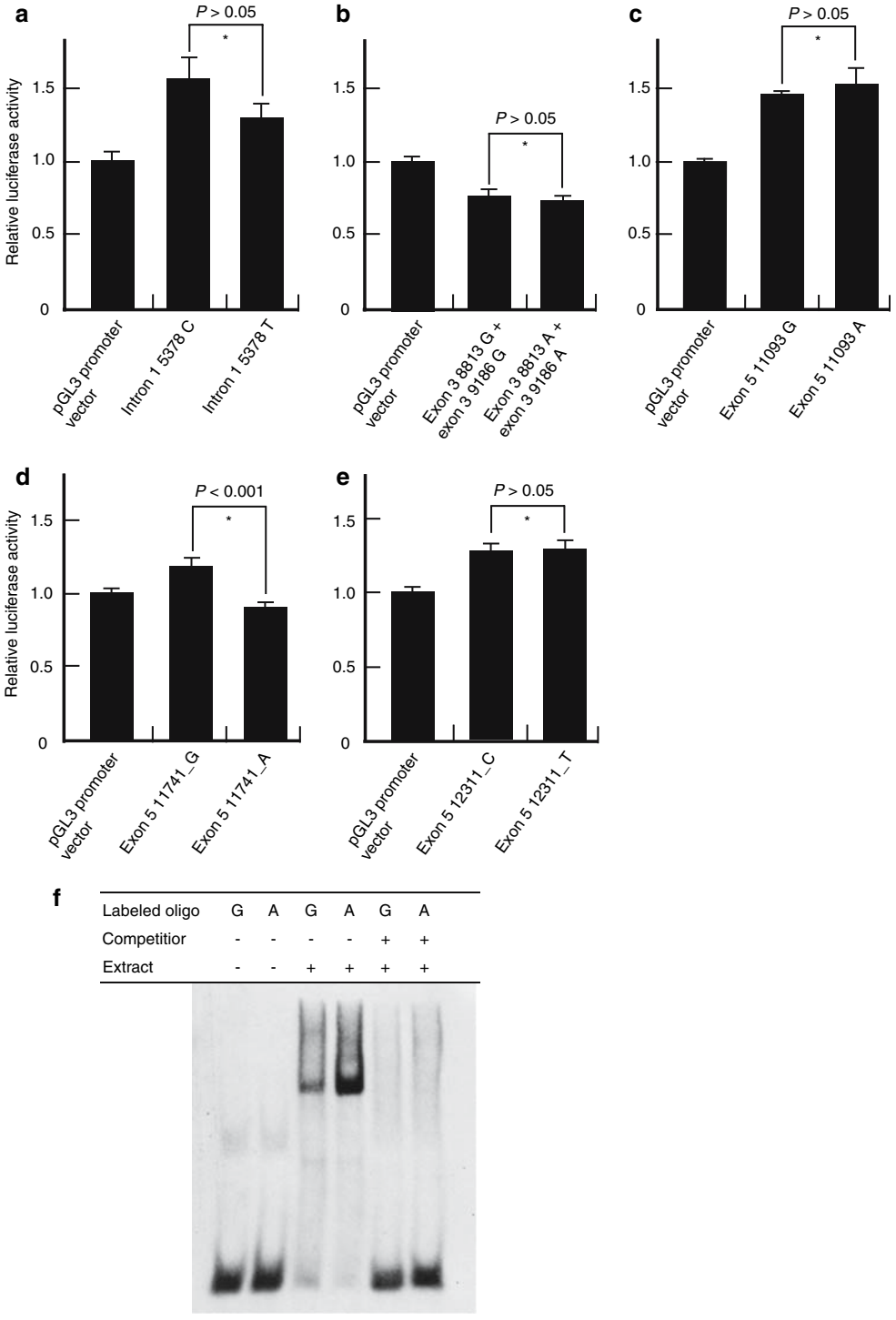


Fig. 5 Transcriptional effect of the six SNPs in *MIAT*. **a–e** Luciferase assays in HEK293 cells. Only exon 5 11,741 G>A SNP affected transcriptional activity (**d**). *Student's *t* test. We repeated each experiment three times and studied each sample in triplicate or duplicate. **f** Binding of an unknown nuclear factor(s) to exon 5 of *MIAT*. The experiments were repeated three times with similar results

Effect of SNP on RNA stability

In two previous studies, SNPs and haplotypes in exons have been implicated in having some roles in the stability of mRNA (Suzuki et al. 2003; Yang et al. 2003). To examine a possibility that the five exonic SNPs showing the significant association with MI (exon 3 8,813 G>A, exon 3 9,186 G>A, exon 5 11,093 G>A, exon 5 11,741 G>A and exon 5 112,311 C>T) might influence the stability of mRNA, we carried out RNA stability assay using HEK293 cells (Suzuki et al. 2003). Although the mRNA corresponding to the minor haplotype tended to

be degraded more rapidly as compared with that corresponding to the major one, the difference was not statistically significant (data not shown).

Discussion

Through a large scale case-control association study using gene-based genome-wide tag SNPs, we found that six SNPs in *MIAT*, a novel gene encoding a possible non-coding functional RNA, might confer the genetic risk of MI. Since the function of this gene was not known, it was not possible to identify this gene as associated with MI by means of candidate gene approach. An advantage of the genome-wide association without any hypothesis is to find genetic variations, even in genes encoding functional RNAs, associated with various diseases. Hence, we are confident that a comprehensive association study using genome-wide tag SNPs is a powerful tool to fully understand genetic backgrounds of common diseases.

Our findings indicated that *MIAT* is a non-coding functional RNA. Biological functions of several non-coding functional RNAs have been investigated intensively; for example, H19 was shown to be involved in imprinting (Pfeifer et al. 1996). Xist was the first example of the functional RNA and was proven to regulate inactivation of the X chromosome (Brown et al. 1991). Hoxa11s is an antisense RNA for Hoxa11 and regulates transcription of Hoxa11 (Hsieh-Li et al. 1995). Recently, Willingham et al. (2005) identified a functional long non-coding RNA (termed NRON), which acts as a repressor of the nuclear factor of activated T cell (NFAT), and showed that specific ncRNAs as NRON may play a role in regulating the complexity of intracellular trafficking. Carninci et al. (2005) found over 23,000 non-coding RNA species through comparison between full-length cDNA sequences and genome sequences, indicating the complexity of mammalian transcriptional landscape. Thus, a large number of non-coding functional RNAs seem to play important roles in a variety of biological functions. Another aspect of functional RNA is a micro RNA (miRNA), one of the sequence-specific post-transcriptional regulators of gene expression (Tang 2005). It is generated by Dicer, a multidomain enzyme of the RNase III family. Dicer cuts precursor miRNAs with hairpin structure into miRNAs. However, sequence comparison of sense strand with antisense one using BLASTN program (Altschul et al. 1990) did not reveal complementary segments within *MIAT*, suggesting no possible hairpin structure that might give rise to double stranded RNA by Dicer. This indicates that *MIAT* is unlikely to contain an miRNA precursor. Although it is very difficult to reveal the function of *MIAT* with the present knowledge, we think the increasing attention to non-coding RNA and subsequent progress will help to solve this problem.

In the present study, we identified SNPs in *MIAT* conferring susceptibility to MI through a large-scale case-control association study. Although function of *MIAT* remains unclear, we believe that knowledge of genetic factors contributing to the pathogenesis of MI as presented here, will lead to improved diagnosis, treatment and prevention.

Acknowledgments We thank Maki Takahashi, Mayumi Yoshii, Saori Abiko, Wataru Yamanobe, Rumiko Ohishi, Makiko Watabe, Kaori Tabei and Saori Manabe for their assistance. We also thank all the other members of SNP Research Center, RIKEN and OACIS for their contribution to the completion of our study. We are also grateful to members of The Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting our study. This work was supported in part by a grant from the Takeda science foundation, the Uehara science foundation and the Japanese Millennium Project.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Andrews NC, Faller DV (1991) A rapid micropreparation teqhnique fore extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res 19:2499

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

Breslow JW (1997) Cardiovascular disease burden increases, NIH funding decreases. Nat Med 3:600–601

Braunwald E (1997) Shattuck lecture-cardiovascular medicine at the turn of the millennium: triumphs, concerns and opportunities. N Engl J Med 337:1360–1369

Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349:38–44

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu

J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schonbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoja K, Tan SL, Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y, FANTOM Consortium RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Project Core Group) (2005) The transcriptional landscape of the mammalian genome. Science 309:1559–1563

Collins FS, Guyer MS, Charkravarti A (1997) Variations on a theme: cataloging human DNA sequence variation. Science 278:1580–1581

Haga H, Yamada R, Ohnishi Y, Nakamura Y, Tanaka T (2002) Gene-based SNP discovery as part of the Japanese Millennium Genome Project: identification of 190,562 genetic varidations in the human genome. J Hum Genet 47:605–610

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA (1998) Databases on transcriptional regulation: TRANSFAC, TRRD, and COMPEL. Nucleic Acids Res 26:364–370

Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, Thorsteinsdottir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G, Sveinbjornsdottir S, Valdimarsson EM, Matthiasson SE, Johannsson H, Gudmundsdottir O, Gurney ME, Sainz J, Thorhallsdottir M, Andresdottir M, Frigge ML, Topol EJ, Kong A, Gudnason V, Hakonarson H, Gulcher JR, Stefansson K (2004) The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. Nat Genet 36:233–239

Hsieh-Li HM, Witte DP, Weinstein M, Branford W, Li H, Small K, Potter SS (1995) Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility. Development 121:1373–1385

Iida A, Saito S, Sekine A, Kitamoto T, Kitamura Y, Mishima C, Osawa S, Kondo K, Harigae S, Nakamura Y (2001) Catalog of 258 single-nucleotide polymorphisms (SNPs) in genes encoding three organic anion transporters, three organic anion-transporting polypeptides, and three NADH: ubiquinone oxidoreductase flavoproteins. J Hum Genet 46:668–683

Lander ES (1996) The new genomics: global views of biology. Science 274:536–539

Lusis AJ, Mar R, Pajukanta P (2004) Genetics of atherosclerosis. Annu Rev Genomics Hum Genet 5:189–218

Marenberg ME, Rish N, Berkman LF, Floderus B, de Faire U (1994) Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med 330:1041–1046

Mount SM (1982) A catalogue of splice junction sequences. Nucleic Acids Res 10:459–472

Ohnishi Y, Tanaka T, Yamada R, Suematsu K, Minami M, Fujii K, Hoki N, Kodama K, Nagata S, Hayashi T, Kinoshita N, Sato H, Sato H, Kuzuya T, Takeda H, Hori M, Nakamura Y (2000) Identification of the 187 single nucleotide polymorphisms (SNPs) among 41 candidate genes for ischemic heart disease in the Japanese population. Hum Genet 106:288–292

Ohnishi Y, Tanaka T, Ozaki K, Yamada R, Suzuki H, Nakamura Y (2001) A high-throughput SNP typing system for genome-wide association studies. J Hum Genet 46:471–477

Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Sato H, Hori M, Nakamura Y, Tanaka T (2002) Functional SNPs in the lymphotoxin-a gene that are associated with susceptibility to myocardial infarction. Nat Genet 32:650–654

Ozaki K, Inoue K, Sato H, Iida A, Ohnishi Y, Sekine A, Odashiro K, Nobuyoshi M, Hori M, Nakamura Y, Tanaka T (2004) Functional variation in *LGALS2* confers risk of myocardial infarction and regulates lymphotoxin-a secretion in vitro. Nature 429:72–75

Ozaki K, Sato H, Iida A, Mizuno H, Nakamura T, Miyamoto Y, Takahashi A, Tsunoda T, Ikegawa S, Kamatani N, Hori M, Nakamura Y, Tanaka T (2006) A functional SNP in PSMA6 confers risk of myocardial infarction in the Japanese population. Nat Genet (Epub ahead of print)

Pfeifer K, Leighton PA, Tilghman SM (1996) The structural H19 gene is required for transgene imprinting. Proc Natl Acad Sci USA 93:13876–13883

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Shapiro MB, Senapathy P (1987) RNA splice junctions of different classes of eukaryotes: Sequence statistics and functional implications in gene expression. Nucleic Acids Res 15:7155–7174

Stenina OI, Desai SY, Krukovets I, Kight K, Janigro D, Topol EJ, Plow EF (2003) Thrombospondin-4 and its variants: expression and differential effects on endothelial cells. Circulation 108:1514–1519

Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) Functional haplotypes of PADI4, encording citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34:395–402

Tang G (2005) siRNA and miRNA: an insight into RISCs. Trends Biochem Sci 30:106–114

Topol EJ, McCarthy J, Gabriel S, Moliterno DJ, Rogers WJ, Newby LK, Freedman M, Metivier J, Cannata R, O'Donnell CJ, Kottke-Marchant K, Murugesan G, Plow EF, Stenina O, Daley GQ (2001) Single nucleotide polymorphisms in multiple novel thrombospondin genes may be associated with familial premature myocardial infarction. Circulation 104:2641–2644

Tsunoda T, Lathrop GM, Sekine A, Yamada R, Takahashi A, Ohnishi Y, Tanaka T, Nakamura Y (2004) Variation of the gene-based SNPs and linkage disequilibrium patterns in the human genome. Hum Mol Genet 13:1623–1632

Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, Schultz PG (2005) A strategy for probing the functional of noncording RNAs finds a repressor of NFAT. Science 309:1570–1573

Yamada R, Tanaka T, Unoki M, Nagai T, Sawada T, Ohnishi Y, Tsunoda T, Yukioka M, Maeda A, Suzuki K, Tateishi H, Ochi T, Nakamura Y, Yamamoto K (2001) Association between a single-nucleotide polymorphism in the promoter of the human interleukin-3 gene and rheumatoid arthritis in Japanese patients, and maximum-likelihood estimation of combinatorial effect that two genetic loci have on susceptibility to the disease. Am J Hum Genet 68:674–685

Yamada Y, Izawa H, Ichihara S, Takatsu F, Ishihara H, Hirayama H (2002) Prediction of the risk of myocardial infarction from polymorphisms in candidate genes. N Engl J Med 347:1916–1923

Yang T, McNally BA, Ferrone S, Liu Y, Xheng P (2003) A single nucleotide deletion leads to rapid degradation of TAP-1 mRNA in a melanoma cell line. J Biol Chem 278:15291–15296