Identification of a Glutamic Acid Repeat Polymorphism of *ALMS1* as a Novel Genetic Risk Marker for Early-Onset Myocardial Infarction by Genome-Wide Linkage Analysis

Sahoko Ichihara, MD, PhD*; Ken Yamamoto, MD, PhD*; Hiroyuki Asano, MD, PhD; Masahiro Nakatochi, PhD; Mayo Sukegawa, PhD; Gaku Ichihara, MD, PhD; Hideo Izawa, MD, PhD; Akihiro Hirashiki, MD, PhD; Fumimaro Takatsu, MD, PhD; Hisashi Umeda, MD, PhD; Mitsunori Iwase, MD, PhD, FACP; Haruo Inagaki, MD, PhD; Haruo Hirayama, MD, PhD; Takahito Sone, MD, PhD; Kazuhiko Nishigaki, MD, PhD; Shinya Minatoguchi, MD, PhD; Myeong-Chan Cho, MD, PhD; Yangsoo Jang, MD, PhD; Hyo-Soo Kim, MD, PhD; Jeong E. Park, MD, PhD; Saeko Tada-Oikawa, PhD; Hidetoshi Kitajima, MD, PhD; Tatsuaki Matsubara, MD, PhD; Kenji Sunagawa, MD, PhD; Hiroaki Shimokawa, MD, PhD; Akinori Kimura, MD, PhD; Jong-Young Lee, PhD; Toyoaki Murohara, MD, PhD; Ituro Inoue, MD, PhD; Mitsuhiro Yokota, MD, PhD, FACC

Background—Myocardial infarction (MI) is a leading cause of death worldwide. Given that a family history is an independent risk factor for coronary artery disease, genetic variants are thought to contribute directly to the development of this condition. The identification of susceptibility genes for coronary artery disease or MI may thus help to identify high-risk individuals and offer the opportunity for disease prevention.

Methods and Results—We designed a 5-step protocol, consisting of a genome-wide linkage study followed by association analysis, to identify novel genetic variants that confer susceptibility to coronary artery disease or MI. A genome-wide affected sib-pair linkage study with 221 Japanese families with coronary artery disease yielded a statistically significant logarithm of the odds score of 3.44 for chromosome 2p13 and MI. Further association analysis implicated Alström syndrome 1 gene (*ALMS1*) as a candidate gene within the linkage region. Validation association analysis revealed that representative single-nucleotide polymorphisms of the *ALMS1* promoter region were significantly associated with early-onset MI in both Japanese and Korean populations. Moreover, direct sequencing of the *ALMS1* coding region identified a glutamic acid repeat polymorphism in exon 1, which was subsequently found to be associated with early-onset MI.

Conclusions—The glutamic acid repeat polymorphism of *ALMS1* identified in the present study may provide insight into the pathogenesis of early-onset MI. **(*Circ Cardiovasc Genet*. 2013;6:569-578.)**

Key Words: genetic association studies ■ genetics ■ genome-wide association study ■ linkage mapping ■ myocardial infarction ■ susceptibility, genetic

Myocardial infarction (MI) is a leading cause of death worldwide.[1] Given that a family history is an independent risk factor for coronary artery disease (CAD),[2] genetic variants are thought to contribute directly to the development of this condition. The identification of susceptibility genes may thus help to identify individuals at high risk for developing CAD or MI and to offer an opportunity for disease prevention. We previously identified candidate susceptibility genes for MI in a case–control association study with single-nucleotide polymorphisms (SNPs).[3] More recent genome-wide association studies (GWASs) have resulted in the identification of several novel loci associated with CAD or MI.[4–8] Moreover, meta-analyses of large GWASs performed with Europeans or both Europeans and South Asians have identified additional loci associated with CAD.[9–11] Such previous observations have suggested the existence of risk alleles with population-specific effects on susceptibility to CAD.[8]

Editorial see p 531
Clinical Perspective on p 578

Two major approaches to the identification of susceptibility genes for CAD or MI have been adopted: genome-wide population-based case–control studies and genome-wide linkage studies (GWLSs), with the latter representing a comprehensive

Received December 12, 2012; accepted September 26, 2013.
The online-only Data Supplement is available at http://circgenetics.ahajournals.org/lookup/suppl/doi:10.1161/CIRCGENETICS.111.000027/-/DC1.
The current address for Dr Nakatochi is Center for Advanced Medicine & Clinical Research, Nagoya University Hospital, Nagoya, Japan.
*Drs Ichihara and Yamamoto contributed equally to this work.
Correspondence to Mitsuhiro Yokota, MD, PhD Department of Genome Science, Aichi-Gakuin University School of Dentistry, 2-11 Suemori-dorim, Chikusa-ku, Nagoya 464–8651, Japan. E-mail myokota@dpc.aichi-gakuin.ac.jp

approach that is unbiased with respect to gene function.[12,13] In the present study, a GWLS performed with 221 Japanese families with CAD yielded a statistically significant logarithm of the odds (LOD) score of 3.44 for chromosome 2p13 and MI. Subsequent association analysis implicated Alström syndrome 1 gene (*ALMS1*) as a candidate gene within the linkage region. Validation association analysis revealed that representative SNPs of the *ALMS1* promoter region were significantly associated with early-onset MI in both Japanese and Korean populations. Moreover, direct sequencing of the *ALMS1* coding region identified a glutamic acid repeat polymorphism in exon 1, which was subsequently found to be associated with early-onset MI.

Methods

Study Subjects

The characteristics of the subjects analyzed at each stage of the study are described in Table 1. The 455 individuals from 221 families with CAD analyzed in the first step (Figure 1) were recruited through participating hospitals in Japan. Each family has ≥2 affected siblings with CAD. The population recruited for association analysis comprised 2186 unrelated subjects with MI, including 351 with early-onset MI and 135 with early-onset angina pectoris (AP), as well as 6026 unrelated population controls in Japan (Table 1). Early-onset MI or AP was defined as the occurrence of MI or AP at ≤50 or ≤55 years in men or women, respectively. The control subjects were recruited from consecutive individuals who visited outpatient clinics of the participating hospitals for an annual health checkup or who were community-dwelling individuals recruited to a prospective cohort study. In addition, 346 individuals with early-onset MI, 614 with early-onset AP, and 1819 unrelated population controls were recruited through participating hospitals in Korea (Table 1). The study protocol was approved by the committees on the ethics of human research of Aichi-Gakuin University, Nagoya University School of Medicine, Kyushu University School of Medicine, and participating hospitals in Japan and Korea. Written informed consent was obtained from each participant. Further details are provided in Methods in the online-only Data Supplement.

Genotyping

Genome-wide SNP genotyping was performed with the Illumina Human 660 W-quad chip (Illumina, San Diego, CA). We genotyped 455 samples from 221 CAD families using the BeadStudio genotype module (Illumina). Samples with a genotype call rate of <0.99 (n=2) were removed from the analysis. SNPs with either a call rate of <0.98 or a minor allele frequency of <0.05 were eliminated. Pairwise proportion of identity-by-descent estimation revealed that 1 and 4 pairs showed an identity-by-descent proportion (PI_HAT) of 1.0 and <0.3, respectively (Figure I in the online-only Data Supplement). We removed these 5 pairs (n=10) because DNA duplication or half-sibs were suspected. Finally, 443 samples (228 informative affected sib-pairs) were selected for linkage analysis (Table I in the online-only Data Supplement) with SNP markers (Table II in the online-only Data Supplement). To select SNPs for loading to the easyLINKAGE program, we randomly extracted 20% of quality-checked SNPs mentioned above with the use of the PLINK v1.07 program.[14] For linkage analysis on chromosome 2, we genotyped 37 microsatellite markers (*D2S319*, *D2S304*, *D2S2211*, *D2S359*, *D2S162*, *D2S423*, *D2S168*, *D2S2267*, *D2S305*, *D2S165*, *D2S367*, *D2S2259*, *D2S391*, *D2S2153*, *D2S337*, *D2S2320*, *D2S2368*, *D2S2110*, *D2S286*, *D2S438*, *D2S2333*, *D2S2216*, *D2S160*, *D2S347*, *D2S112*, *D2S151*, *D2S142*, *D2S2330*, *D2S335*, *D2S364*, *D2S117*, *D2S325*, *D2S164*, *D2S126*, *D2S206*, *D2S338*, and *D2S125*). The markers were amplified by the polymerase chain reaction (PCR) with fluorescently labeled primers; the PCR products were subjected to electrophoresis in an ABI3100 sequencer (Life Technologies, Carlsbad, CA), and analysis and assignment of marker alleles were performed with Genotyper software (Life Technologies).

For association analysis with the 10.5-Mb region on chromosome 2 corresponding to a LOD score of >2.6 for the MI subgroup, SNP genotyping was performed with the Illumina Golden Gate Assay for 112 cases (probands of MI sib-pairs) and 495 controls (second step in Figure 1). Samples with a genotype call rate of <0.95 (n=1) were removed from the analysis. SNPs with a call rate of <0.95, minor allele frequency of <0.05, or Hardy–Weinberg $P<0.0001$ were eliminated. The association analysis was thus performed with the remaining 111 probands of MI sib-pairs and 480 hospital-based controls and with 1248 SNPs located on 2p14-p12. To validate the association between SNPs in the promoter region of *ALMS1*, rs10191517 and rs6748040, and MI (third step in Figure 1), we performed SNP genotyping by the TaqMan method (Life Technologies) in 2186 patients with MI, including those with early-onset MI (n=351) or early-onset AP (n=135), and 6026 controls in Japan (Table 1). The validation of

Table 1. Characteristics of Case and Control Subjects in the Present Study

	First Step		Second Step			Third and Fifth Steps (Japanese)			Third Step (Korean)		
Characteristic	All Pairs	MI Pairs	Control	MI	Control	MI	Early-Onset MI	Early-Onset AP	Control	Early-Onset MI	Early-Onset AP
No. of individuals	455	226	495	112	6,026	2,186	351	135	1819	346	614
No. of pedigrees	221	112									
Sex (% male)	78.2	83.2	67.9	82.1	58.5	82.4	83.8	64.4	79.6	88.2	73.9
Age at recruitment, y	65.2±8.8	65.6±9.0	67.6±5.8	65.2±8.7	61.9±7.0	63.4±10.3	48.9±8.0	49.2±6.9	54.2±8.5	44.5±5.5	46.5±4.9
Age at first event, y	63.1±9.1	62.8±9.5		62.5±8.6		61.1±10.5	44.5±5.2	45.5±5.6		NA	NA
Body mass index, kg/m^2	24.0±2.9	24.0±3.0	23.3±2.8	23.8±2.9	23.0±3.0	23.7±3.0	24.5±3.2	24.2±3.2	24.1±2.7	25.2±3.3	25.6±2.9
Former or current smoker, %	63.0	69.3	45.5	64.8	44.7	69.8	73.8	54.1	56.6	77.1	57.1
MI, %	69.7	100		100		100	100	0		100	0
AP, %	35.2	2.8		3.4		18.1	12.3	100		0	100
Hypertension, %	56.6	51.7	68.5	53.4	48.9	50.4	43.9	53.3	29.2	53.5	38.0
Diabetes mellitus, %	34.7	37.5	20.4	34.1	16.0	36.5	29.9	36.3	NA	NA	NA
Hyperlipidemia, %	58.3	55.7	36.8	56.8	50.8	51.1	62.4	52.6	31.9	62.2	47.6

Data for continuous variables are mean±SD. AP indicates angina pectoris; MI, myocardial infarction; and NA, not available.

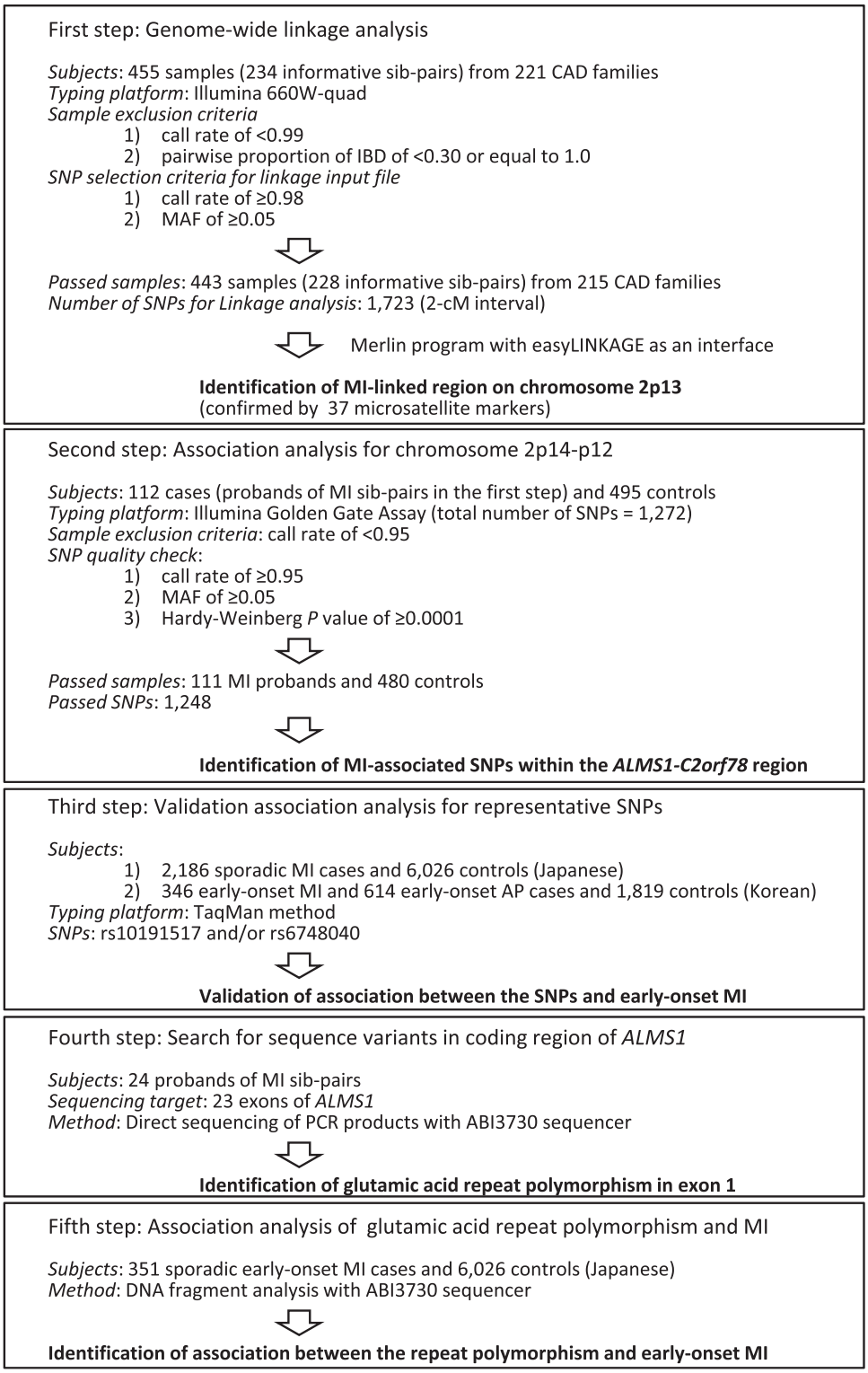


Figure 1. Study design. Genome-wide linkage analysis with single-nucleotide polymorphism (SNP) markers and further genotyping with microsatellite markers were performed in the first step. Association analysis for chromosome 2p14-p12 and validation association analysis for representative SNPs were performed in the second and third steps. *ALMS1*, which showed the strongest association with myocardial infarction (MI), was sequenced in the fourth step, and the association between a glutamic acid repeat polymorphism of *ALMS1* and early-onset MI was analyzed in the fifth step. AP indicates angina pectoris; CAD, coronary artery disease; IBD, identity-by-descent; MAF, minor allele frequency; and PCR, polymerase chain reaction.

the association between rs6748040 and MI was also performed with individuals with early-onset MI (n=346) or early-onset AP (n=614), as well as 1819 controls in Korea (Table 1).

Association Analysis of a Repeat Polymorphism With Early-Onset MI

To search for sequence variants of *ALMS1* (fourth step in Figure 1), we subjected PCR products derived from the coding region of the gene to direct sequencing with an ABI3730 sequencer (Life Technologies; Table III in the online-only Data Supplement). The samples were obtained from 24 probands of MI affected sib-pairs included in the initial linkage and association analysis. To validate the association between the detected glutamic acid repeat polymorphism of *ALMS1* and early-onset MI (fifth step in Figure 1), we performed DNA fragment analysis with the ABI3730 sequencer for 351 early-onset MI patients and 6026 controls (Table 1).

Analysis of ALMS1 Expression in Epstein-Barr Virus–Transformed Lymphoblastoid Cell Lines

Epstein-Barr virus–transformed B-cell lines derived from 92 healthy Japanese subjects were obtained from the Cell Bank at RIKEN BioResource Center (Tsukuba, Japan)[15] and genotyped by DNA fragment analysis. Total RNA was extracted from the cells with the use of an RNeasy Mini Kit (Qiagen, Hilden, Germany) and subjected to reverse transcription with the use of a first-strand cDNA Synthesis Kit (GE Healthcare, Amersham, UK). The resulting cDNA was then subjected to real-time PCR analysis for evaluation of *ALMS1* expression with the use of a TaqMan Gene Expression Assay (Life Technologies) and the On ECO Real-time PCR System (Illumina), with GAPDH cDNA serving as an endogenous control. The amount of ALMS1 mRNA was calculated by the comparative cycle threshold method,[16] and all data were normalized by the amount of GAPDH mRNA.

Statistical Analysis

Multipoint nonparametric LOD scores were calculated with the use of Merlin v1.0.1 (with easyLINKAGE-Plus v5.08).[17,18] A total of 1723 SNPs whose pairwise r^2 values were <0.1 were automatically selected by easyLINKAGE for linkage analysis to the whole genome (2.0-cM spacing). Nonparametric linkage analysis with 37 microsatellite markers spanning chromosome 2p16-q13 was also performed with Merlin v1.0.1 (with easyLINKAGE-Plus v5.08). The genetic position of markers was based on the Marshfield linkage map,[19] and the sex-averaged position was applied. Association analysis for allelic and genotypic tests was performed with PLINK. The results from Japanese and Korean subjects were then combined into a fixed-effects meta-analysis with inverse variance weighting with the use of the METAL package (www.sph.umich.edu/csg/abecasis/metal). Heterogeneity between Japanese and Korean populations was tested with the χ^2-based Cochran Q statistic. Pairwise linkage disequilibrium (LD) between the glutamic acid repeat polymorphism and SNPs was calculated with Haploview[20] in control samples. The base position of SNPs was based on the database of NCBI36/hg18. Relations between nongenetic factors and genotype information for early-onset MI cases were examined by the χ^2 test for sex and classical risk factors and by 1-way ANOVA for age at recruitment, age at first event, and body mass index. To assess the independent association of early-onset MI with the glutamic acid repeat polymorphism in exon 1 of *ALMS1*, we specified 3 models: in model 1, we performed logistic regression analysis for the additive test with adjustment for sex; in model 2, we adjusted for sex and the absence or presence of diabetes mellitus; and in model 3, we adjusted for body mass index, smoking, hypertension, and hyperlipidemia, as well as for the variables in model 2. Differences in gene expression between 2 groups were evaluated by the *U* test. A $P<0.05$ was considered statistically significant.

Results

Mapping of MI Susceptibility Loci by Linkage Analysis

Although each conventional risk factor, including diabetes mellitus, hypertension, and hypercholesterolemia, is in part under genetic control, a family history of CAD is an independent predictor of this condition, suggesting the existence of additional susceptibility genes.[21] We designed a 5-step protocol, consisting of a GWLS followed by association analysis, to identify novel genetic variants that confer susceptibility to CAD and MI (Figure 1). In the first step, we performed a 2-cM genome scan with SNP markers for Japanese sib-pairs with CAD (n=228 informative sib-pairs) and for the subgroup of sib-pairs with MI (n=114 informative sib-pairs) to identify chromosomal regions linked to CAD or MI. We detected signals showing a multiple nonparametric LOD score of >1.0 for both phenotypes on chromosomes 2 and 12 (Figure 2). The LOD score for the MI subgroup at the peak marker rs1406105 on chromosome 2 was higher than that for the CAD pairs (3.44 versus 1.66; Figure 2 and Table IV in the online-only Data Supplement), whereas the peak scores on chromosome 12 were similar for the 2 phenotypes (1.79 versus 1.35, respectively). Further linkage analysis with microsatellite markers spanning chromosome 2p16-q13 revealed a strong linkage signal (multipoint LOD score=2.21) for the MI subgroup and the marker *D2S2110* at chromosome 2p13 (Figure 3A; Table V in the online-only Data Supplement).

Association Analysis With SNPs at Chromosome 2p14-p12

To identify the gene responsible for the observed linkage with MI, we further examined the 10.5-Mb region on chromosome 2 corresponding to a LOD score of >2.6 for the MI subgroup in the SNP scan by performing association analysis with SNPs (second step in Figure 1). Initial screening was performed for 112 probands of MI sib-pairs and 495 hospital-based controls with 1272 SNPs located at 2p14-p12. A group of SNPs spanning a 0.6-Mb region at 2p13.1 showed association with MI (*P* value for the additive model of between 1×10^{-3} and 1×10^{-4}; Figure 3B). The strongest association was observed with rs10191517 (*P* value for the additive model=1.8×10^{-4}, odds ratio [OR]=1.81), and the SNPs highly correlated with this polymorphism also showed an association signal (Figure 3C and Table VI in the online-only Data Supplement). Consistent with the r^2 values for rs10191517, the association signals for the other SNPs were almost nil after conditioning with this landmark SNP (Figure 3C). Among the 11 known genes in this region, the SNPs within the LD block located in the promoter region of *ALMS1* showed the strongest association, implicating *ALMS1*, the gene responsible for all cases of Alström syndrome,[22,23] as a strong candidate gene within the linkage region (Figure 3C). A quantile–quantile plot revealed that the distribution of *P* values for the association of all 1248 examined SNPs diverged markedly from the null expectation,

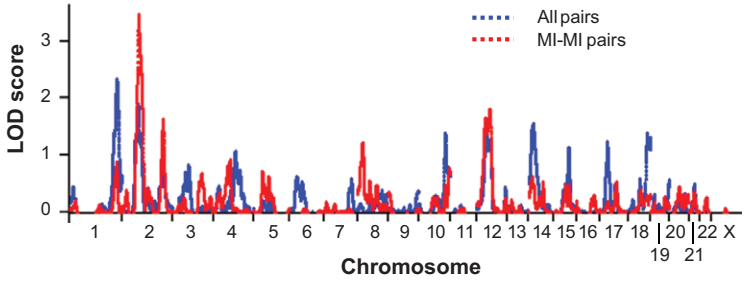


Figure 2. Linkage scan for sib-pairs affected by coronary artery disease (CAD) or myocardial infarction (MI). Blue dots represent the distribution of multipoint logarithm of the odds (LOD) scores from the Merlin program for all pairs affected by CAD (228 pairs), and the red dots represent that for the MI subgroup (114 pairs). The marker interval is 2.0 cM.

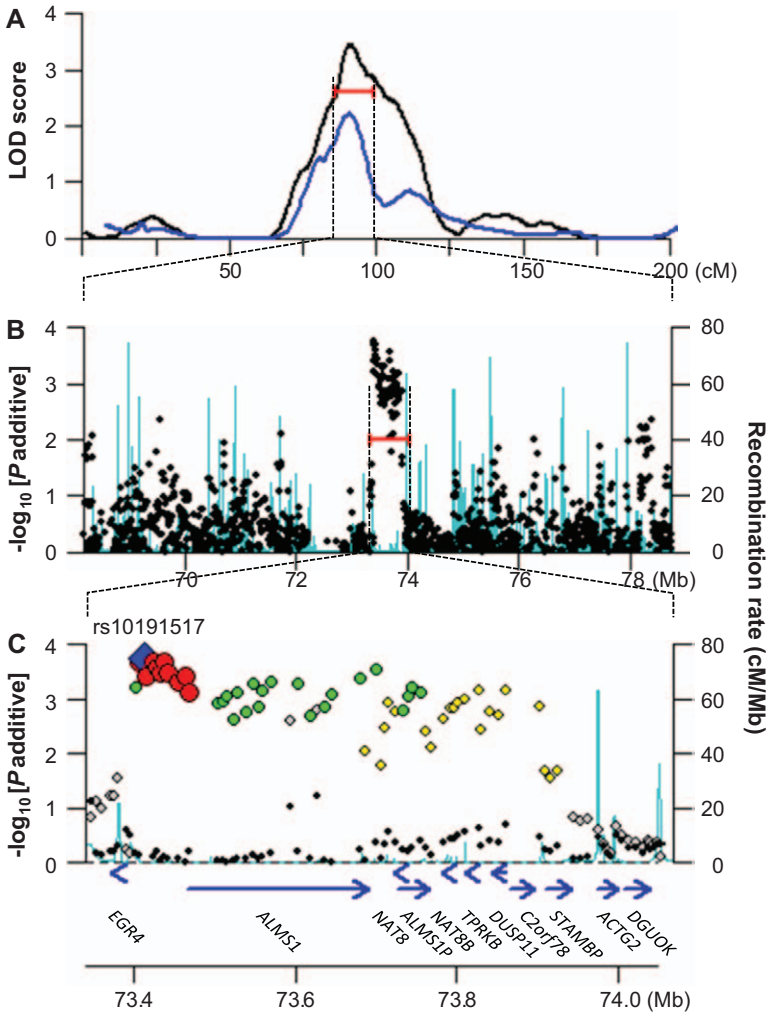


Figure 3. Mapping of susceptibility to myocardial infarction (MI) on chromosome 2p13 and in the *ALMS1-C2orf78* region in particular. **A**, Linkage analysis results for chromosome 2 and MI pairs with single-nucleotide polymorphisms (SNPs; black line) or microsatellite markers (blue line). **B**, Regional association analysis for the probands of MI pairs. The $-\log_{10}$ (*P* values) for the additive test are shown for SNPs distributed in a 10.5-Mb genomic region including a peak of linkage analysis. **C**, Linkage disequilibrium structure of the associated genomic region (from 73.34–74.05 Mb) and gene locations. The landmark SNP (rs10191517) is depicted as a blue diamond. SNPs showing r^2 values with rs10191517 of >0.8, 0.8–0.5, 0.5–0.2, and <0.2 are shown in red, green, yellow, and gray, respectively. The association signals conditioned with rs10191517 are depicted as small black dots. The r^2 values were calculated in control samples. The genetic recombination rate (cM/Mb; estimated from the HapMap CHB+JPT samples) is shown as a light blue line, and the genes within the region are annotated and shown as blue arrows. Genetic information is from the Human Genome build hg18.

whereas the quantile–quantile plot obtained after exclusion of the 52 SNPs within the *ALMS1-C2orf78* region followed the 45º line (Figure II in the online-only Data Supplement). These results suggested that the ≈0.5-Mb *ALMS1-C2orf78* region might contain the candidate susceptibility gene for MI. Cases (n=111) and controls (n=480) examined in the second step of the present study clustered together with samples from the CHB and JPT HapMap populations in a principal component analysis plot (Figure III in the online-only Data Supplement), further indicative of nonstratification of the study population.

Association Analysis in the Validation Data Set

To validate the association between the SNPs in the *ALMS1* promoter region and MI, we prepared another sample set consisting of 2186 Japanese patients with MI and 6026 controls (third step in Figure 1). We found that 2 representative SNPs, rs10191517 (*P* value for additive model=4.23×10^{-5}; OR=1.40 [95% confidence interval=1.19–1.65]) and rs6748040 (*P* value for additive model=7.18×10^{-6}; OR=1.45 [1.23–1.70]), showed a significant association with early-onset MI, whereas they were not associated with MI in the entire sample group (Table 2). This association was specific for early-onset MI because no positive relation was observed between the SNPs and early-onset AP (Table 2). These results suggested that genetic factors influence MI occurrence, particularly in individuals with a family history or with early-onset disease. We therefore further assessed the association with the use of a Korean sample set containing early-onset MI (n=346) and early-onset AP (n=614) cases, as well as 1819 controls. Consistent with the findings with the Japanese samples, rs6748040, which is in strong LD with rs10191517 (r^2=0.91), showed a significant association with early-onset MI (*P* value for additive model=2.98×10^{-4}; OR=1.38 [1.16–1.64]) but not with early-onset AP (*P* value for additive model=0.20; Table 2). Combined analysis with the Japanese and Korean samples revealed an even more significant association (*P* value for additive model=8.98×10^{-9}; OR=1.41 [1.26–1.59]; Table 2). Together, our results suggested that SNPs in the promoter region of *ALMS1* are associated with susceptibility to early-onset MI in East Asian populations. All *P* values based on the Q statistic were >0.05, indicative of the absence of heterogeneity for rs6748040 between the Japanese and Korean populations.

Association Analysis for a Glutamic Acid Repeat Polymorphism of ALMS1 and MI

ALMS1 comprises 23 exons and encodes a protein of 4169 amino acids.[24] To search for sequence variants in the coding

Table 2. SNPs Located in the 5' Region of *ALMS1* Are Associated With Early-Onset MI

SNP	RAF				Allele Test						Additive Model			
					All MI		Early-Onset MI		Early-Onset AP		All MI		Early-Onset MI		Early-Onset AP	
	All MI	Early-Onset MI	Early-Onset AP	Control	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value	OR (95% CI)*	*P* Value	OR (95% CI)*	*P* Value	OR (95% CI)*	*P* Value
Japanese	(n=2186)	(n=351)	(n=135)	(n=6026)												
rs10191517	0.278	0.337	0.285	0.267	1.06 (0.98–1.14)	0.15	1.40 (1.19–1.64)	4.61×10^{-5}	1.10 (0.84–1.43)	0.50	1.06 (0.98–1.15)	0.13	1.40 (1.19–1.65)	4.23×10^{-5}	1.10 (0.84–1.43)	0.50
rs6748040	0.276	0.344	0.278	0.266	1.05 (0.97–1.14)	0.19	1.45 (1.24–1.70)	5.42×10^{-6}	1.06 (0.81–1.39)	0.66	1.05 (0.97–1.14)	0.19	1.45 (1.23–1.70)	7.18×10^{-6}	1.06 (0.81–1.39)	0.66
Korean	(n=346)	(n=614)	(n=1819)													
rs6748040	0.327	0.279	0.259				1.39 (1.16–1.65)	2.37×10^{-4}	1.11 (0.96–1.28)	0.18			1.38 (1.16–1.64)	2.98×10^{-4}	1.10 (0.95–1.27)	0.20
Combined†	(n=697)	(n=749)	(n=7845)													
rs6748040	0.336	0.279	0.263				1.42 (1.26–1.60)	6.03×10^{-9}	1.10 (0.96–1.24)	0.16			1.41 (1.26–1.59)	8.98×10^{-9}	1.09 (0.96–1.24)	0.18

AP indicates angina pectoris; CI, confidence interval; MI, myocardial infarction; OR, odds ratio; RAF, risk allele frequency; and SNP, single-nucleotide polymorphism.

*The odds ratio (OR) was adjusted for sex.

†Combined analysis with the fixed-effect inverse variance model; the means of the RAFs are presented. All *P* values based on the Q statistic were >0.05, indicating the absence of heterogeneity for rs6748040 between Japanese and Korean populations.

region of *ALMS1*, we sequenced all exons of the gene in 24 probands of MI sib-pairs that were studied in the initial linkage and association analysis (fourth step in Figure 1). We identified 3 nonsynonymous SNPs (rs2037814 in exon 8, rs3820700 in exon 10, and rs1052161 in exon 19) and 1 repeat polymorphism (glutamic acid repeat polymorphism in exon 1; Figure 4). Given that there is a complete correlation among the 3 nonsynonymous SNPs and that rs3820700 showed a weaker association with MI than did rs10191517 (Table VI in the online-only Data Supplement), we focused on the repeat polymorphism for further analysis. In the fifth step of the study (Figure 1), among 14 alleles (9–22 repeats, designated A9–A22) observed in Japanese individuals with early-onset MI (n=351) and controls (n=6026), we tested the 7 alleles with a frequency of >0.01 for LD with representative SNPs (rs10191517 and rs6748040). We found that 2 frequent alleles, A14 and A17, showed weak to moderate LD with these SNPs (r^2=0.78 and 0.29, respectively, for rs10191517; r^2=0.84 and 0.32, respectively, for rs6748040; Table VII in the online-only Data Supplement). This result prompted us to examine whether an allele of the glutamic acid repeat polymorphism might exhibit a stronger association with early-onset MI. The allele frequency for A14 was increased in cases compared with controls (*P* value for additive model=9.50×10^{-6}; Table 3). Furthermore, the allele frequency for short (A9 to A16) repeats was significantly associated with early-onset MI (Table 3), and the association signal was stronger than that for the SNP rs6748040 (*P* value for additive model=1.17×10^{-6} and 7.18×10^{-6} for the short repeat and rs6748040, respectively). This result suggested that the repeat polymorphism might be primarily associated with early-onset MI, and that a short repeat of glutamic acid residues might affect the molecular function of ALMS1 with respect to the pathogenesis of this disease. Moreover, the relative expression level of *ALMS1* in Epstein-Barr virus–transformed B-cell lines with the rs6748040 A/A or A/G genotypes, which are a risk factor for early-onset MI, was significantly higher than that in those with the G/G genotype (Figure 5). This result might reflect a difference in the binding of transcription factors between alleles of SNPs located in the *ALMS1* promoter region including rs6748040 (Table VIII in the online-only Data Supplement). The observation that the risk allele of rs6748040 was in strong LD with A14 (Table VII in the online-only Data Supplement) further suggests that the short-repeat form of *ALMS1* might be expressed at a higher level than the long-repeat form. We examined the possibility of confounding effects of age, sex, and classical risk factors for early-onset MI cases but did not find any clear relations between genotype for the repeat polymorphism and these factors (Table IX in the online-only Data Supplement). Furthermore, logistic regression analysis for the additive test with adjustment for sex (model 1), for sex and the absence or presence of diabetes mellitus (model 2), or for body mass index, smoking, hypertension, and hyperlipidemia, as well as for the variables in model 2 (model 3), revealed that the glutamic acid repeat polymorphism was significantly associated with early-onset MI in each model (Table X in the online-only Data Supplement), suggesting

Figure 4. Glutamic acid repeat polymorphism located in exon 1 of *ALMS1*. The DNA sequence from positions 73 466 485 to 73 466 720 of chromosome 2 (hg18 database) is shown together with the encoded amino acid sequence. The exonic sequence is presented in uppercase, and the locations of primers for polymerase chain reaction–based genotyping of the glutamic acid repeat polymorphism are underlined.

that the repeat polymorphism of *ALMS1* is an independent risk factor for early-onset MI.

Discussion

We have identified a susceptibility locus for MI on chromosome 2p13 with genome-wide significance for linkage. Given that several novel loci associated with the risk for CAD or MI have been identified by GWASs performed mostly in populations of European descent, we aimed to identify novel loci associated with these conditions in Asians. Whereas some previous studies succeeded in identifying chromosomal regions associated with susceptibility to CAD, they failed to identify the responsible gene variants. In the present study, however, further association analysis identified *ALMS1* as a candidate gene within the linkage region on 2p13. Moreover, we found that a glutamic acid repeat polymorphism in exon 1 of the gene is associated with early-onset MI.

In the present study, the multipoint 2-cM genome scan performed with SNP markers yielded the highest LOD score (3.44) for the MI subgroup of subjects at the marker rs1406105 on chromosome 2. Given that a LOD score threshold of 2.2 has been shown to provide a rigorous standard for suggestive linkage in genome-wide scans for complex trait loci,[25] our results implicated the 2p13 region as a susceptibility locus for MI. We adopted SNPs instead of microsatellites as the genetic markers for linkage analysis because they are more abundant and their genotypes are more amenable to automatic calling.[26] Analysis of 37 microsatellite markers yielded a linkage signal for the MI subgroup in the same chromosomal region (2p13) at the marker *D2S2110* (multipoint LOD score=2.21), supporting the effectiveness of GWLSs with SNP markers for the identification of genomic variations associated with complex or multigenic diseases.[27,28] In our linkage analysis, the peak marker (rs1406105) showed a *Z*-score of 3.05 with an asymptotic *P* value of 0.0011, corresponding to a Kong and Cox LOD score of 3.44. However, the second step of our study yielded association signals with *P* values in the range of 0.01 to 0.0002 for the SNPs in the peak region identified by linkage analysis. Similar levels of significance were thus observed for both the linkage analysis and second-step association analysis with the probands of MI sib-pairs.

In the present study, we identified a susceptibility locus on chromosome 2p13 for MI with a GWLS in the Japanese population. Previous GWLSs have identified loci associated with MI on chromosome 2q21.1-q22,[29] 2p12-2q23.3,[30] and 2q36-q37.3[31] for Europeans or Australians. Although 1 previous linkage study identified an MI-linked region in the vicinity of 2p11,[13] the loci on chromosome 2 and other chromosomes showing suggestive linkage to CAD or MI in all previous genome scans do not overlap with that identified in the present study. Each previous study that identified a susceptibility locus for MI or CAD with a GWLS also showed a different linkage map.[12,13,29,31] It is difficult to cover the whole genome with only ≈500 microsatellite markers. Moreover, given the relatively small sample sizes, the results of such studies would be expected to be affected by a variety of environmental factors, including racial differences. Such factors might explain the differences in linkage maps obtained for Japanese in the present study and for other ethnic groups in previous GWLSs. However, previous GWASs have implicated loci on chromosomes 2q36.3 (rs2942634) and 2q33 (rs6725887) in *WDR12* (WD repeat domain 12) as being significantly associated with CAD or MI.[4,7] However, the loci showing suggestive linkage

Table 3. Association of Early-Onset MI With the Glutamic Acid Repeat Polymorphism in Exon 1 of *ALMS1*

	Frequency		Allele Test		Additive Test	
Allele	Early-Onset MI	Control	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
A14	0.340	0.264	1.44 (1.22–1.69)	1.12×10^{-5}	1.46 (1.24–1.73)	9.50×10^{-6}
A15	0.025	0.034	0.72 (0.44–1.17)	0.19	0.70 (0.43–1.15)	0.16
A16	0.098	0.075	1.34 (1.03–1.74)	0.027	1.36 (1.05–1.76)	0.022
A17	0.393	0.487	0.68 (0.58–0.80)	1.70×10^{-6}	0.67 (0.57–0.79)	1.25×10^{-6}
A18	0.077	0.079	0.97 (0.73–1.30)	0.86	0.99 (0.74–1.33)	0.95
A19	0.020	0.017	1.21 (0.70–2.08)	0.50	1.15 (0.67–1.98)	0.62
A20	0.036	0.036	1.00 (0.66–1.50)	0.99	0.99 (0.66–1.49)	0.95
Short*	0.471	0.379	1.46 (1.25–1.71)	1.15×10^{-6}	1.48 (1.26–1.73)	1.17×10^{-6}

Results for alleles with a frequency of <0.01 (A9 to A13, A21, and A22) are not shown. CI indicates confidence interval; MI, myocardial infarction; and OR, odds ratio.

*Alleles were classified into short (A9 to A16) or long (A17 to A22) repeat groups, with the results for the short-repeat group being shown.

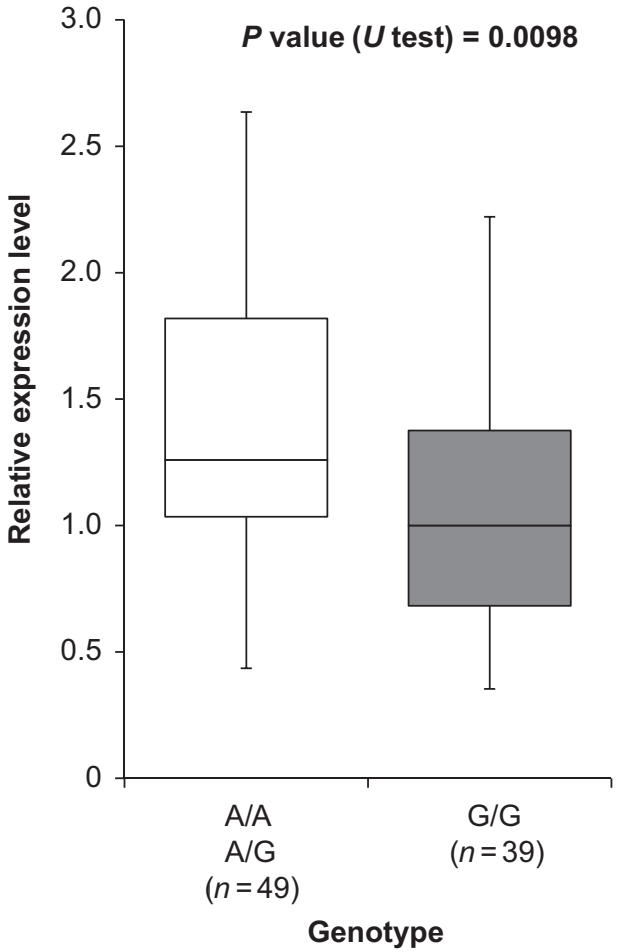


Figure 5. Expression of *ALMS1* in Epstein-Barr virus–transformed B-cell lines (EB-LCLs) according to genotype for rs6748040 located in the promoter region of *ALMS1*. The amount of ALMS1 mRNA in EB-LCLs with the A/A or A/G genotype (n=49) or those with the G/G genotype (n=39) was determined by reverse transcription and real-time polymerase chain reaction analysis. Data were normalized by the abundance of GAPDH mRNA, are expressed relative to the value for the G/G genotype, and are presented as a box-and-whisker plot.

to CAD or MI in GWASs also do not overlap with that identified in the present study. GWLSs and GWASs have different characteristics with regard to localization of trait loci.[32] The differences between the 2 approaches often lead to the identification of different disease-related genetic variants, possibly explaining why *ALMS1* has not yet been implicated in early-onset MI by a well-powered GWAS.

Alström syndrome is a homogenous autosomal recessive disorder characterized by childhood obesity, insulin resistance, and type 2 diabetes mellitus, as well as retinal dystrophy, sensorineural hearing loss, cardiomyopathy, and kidney and liver dysfunction.[24] Direct sequencing of *ALMS1* has identified 79 different mutations, clustering in exons 8, 10, and 16, in patients with Alström syndrome.[33,34] Although *ALMS1* polymorphisms are common in the general population, no evidence was found of an association between common variants of the gene and type 2 diabetes mellitus.[35,36] A previous GWAS suggested that variants in or near *NAT8/ALMS1* are associated with differences in glomerular filtration in the general population, however.[37] The ALMS1 protein is widely expressed in human tissues, but its function is unknown.[33] The pathophysiology of Alström syndrome is thought to be related to defects in microtubule organization and intracellular transport.[38] The development of CAD has not been found to be affected in individuals with Alström syndrome although a case of premature-onset CAD was recently described in the longest surviving patient with this syndrome.[39] Alström syndrome is also associated with renal failure secondary to glomerulofibrosis and with fibrotic lung disease.[24] Moreover, MRI has revealed fatty infiltration of the myocardium and patchy left and right ventricular fibrosis,[40] suggesting that multiorgan fibrotic infiltration is a common feature of this syndrome. Indeed, fibroblasts harboring *ALMS1* mutations proliferate continuously and synthesize high levels of extracellular matrix, which likely contribute to excessive remodeling and destruction of normal tissue architecture, resulting in fibrosis.[41] Given that atherosclerosis is associated with the development of intravascular fibrous plaque,[42] mutated *ALMS1* may directly result in multiorgan fibrotic changes, subsequently leading to CAD.

Direct sequencing of PCR products resulted in the identification of a glutamic acid repeat polymorphism in exon 1 of *ALMS1*. Two frequent alleles, A14 and A17, of this polymorphism showed weak to moderate LD with rs6748040, an SNP in the promoter region of *ALMS1* found to be associated with early-onset MI. The allele frequency for A14 was similar to the minor allele frequency for rs6748040, suggesting that A14 and the minor allele of rs6748040 form a haplotype. Similarly, A17 and the major allele of rs6748040 might constitute a haplotype. Association analysis showed that A14 and A17 were significantly related to the risk for and protection from early-onset MI, respectively. Trinucleotide repeat polymorphisms have been associated with several genetic diseases, including Huntington disease and myotonic dystrophy,[43,44] although such polymorphisms manifest a high level of variability among normal individuals. The long form of a glutamic acid (GAG trinucleotide) repeat polymorphism in the α2C2-adrenergic receptor is required for agonist-promoted receptor phosphorylation and desensitization.[45] A GAG trinucleotide repeat polymorphism in the gene for the catalytic subunit of γ-glutamylcysteine ligase was found to be associated with altered activity and expression (at the translational level) of this glutathione biosynthetic enzyme and with altered glutathione levels.[46] Furthermore, continuous amino acid repeats comprising ≥6 residues are important determinants of the 3-dimensional structure of proteins.[47] In the present study, we detected a small but significant difference in the expression of *ALMS1* in Epstein-Barr virus–transformed B-cell lines between genotypes for rs6748040, suggesting that the short-repeat form of *ALMS1* might be expressed at a higher level than the long-repeat form. Further investigation is required, however, to determine the precise effect of the glutamic acid repeat polymorphism on the molecular function of ALMS1.

In conclusion, a GWLS and SNP fine mapping on chromosome 2p13 have identified *ALMS1* as a new MI susceptibility gene. We found that a glutamic acid repeat polymorphism of *ALMS1* was significantly associated with early-onset MI, suggesting that this polymorphism may contribute to the development of MI, including that in individuals with a family history

or that with an early-onset. This polymorphism is also a potential new target for prevention of early-onset MI.

Acknowledgments

We thank all patients and family members who participated in this study. In addition to the authors, the following investigators at Japanese institutions and hospitals contributed to the study: H. Horibe, M. Watarai, K. Takemoto, and S. Shimizu (Anjo Kosei Hospital, Anjo); Y. Murase and Y. Suzuki (Okazaki City Hospital, Okazaki); Y. Yoshida (Nagoya Daini Red Cross Hospital, Nagoya); F. Somura (Nagoya University Hospital, Nagoya); K. Takagi (Ogaki Municipal Hospital, Ogaki); C. Takanaka (Hamamatsu Medical Center, Hamamatsu); M. Maeda and Y. Nishinaka (Chita City Hospital, Chita); S. Ishikawa and F. Saito (Showa Hospital, Konan); R. Ishiki (Toyota Memorial Hospital, Toyota); and S. Ogawa and T. Fujimura (Tokai Central Hospital, Kagamihara). We also thank T. Akinaga, M. Sonoda, and K. Fukuyama for technical assistance.

Sources of Funding

This study was supported, in part, by Grants-in-Aid for Scientific Research, including those of Category (A) and (B) and the NEXT Program of the Japan Society for the Promotion of Science and on Priority Areas Comprehensive Genomics and Applied Genomics, from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by the Joint Usage/Research Program of the Medical Research Institute at Tokyo Medical and Dental University.

Disclosures

None.

Appendix

From the Graduate School of Regional Innovation Studies, Mie University, Tsu, Japan (S.I., S.T.-O.); Division of Genome Analysis, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan (K.Y., H.K.); Chiousha Iwakura Hospital, Iwakura, Japan (H.A.); Theranostic Research Institute Inc, Nagoya, Japan (M.N., M.S.); Departments of Occupational and Environmental Health (G.I.) and Cardiology (A.H., T.M.), Nagoya University Graduate School of Medicine, Nagoya, Japan; Fujita Health University, Banbuntane Hotokukai Hospital, Nagoya, Japan (H.I.); Anjo Kosei Hospital, Anjo, Japan (F.T.); Toyota Memorial Hospital, Toyota, Japan (H.U., M.I., H.I.); Nagoya Daini Red Cross Hospital, Nagoya, Japan (H.H.); Ogaki Municipal Hospital, Ogaki, Japan (T.S.); Department of Cardiology, Gifu University Graduate School of Medicine, Gifu, Japan (K.N., S.M.); Center for Genome Science, Korea National Institute of Health, Chungcheongbuk-do, South Korea (M.-C.C., J.-Y.L.); Department of Internal Medicine, Cardiovascular Genome Center, Yonsei University, Seoul, South Korea (Y.J.); Department of Internal Medicine, Cardiovascular Center, Seoul National University Hospital, Seoul, South Korea (H.-S.K.); Division of Cardiology, Samsung Medical Center, Seoul, Korea (J.E.P.); Departments of Internal Medicine (T.M.) and Genome Science (M.Y.), Aichi-Gakuin University School of Dentistry, Nagoya, Japan; Department of Cardiovascular Medicine, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan (K.S.); Department of Cardiovascular Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan (H.S.); Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan (A.K.); and Division of Human Genetics, National Institute of Genetics, Mishima, Japan (I.I.).

References

1. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet*. 1997;349:1498–1504.
2. Shea S, Ottman R, Gabrieli C, Stein Z, Nichols A. Family history as an independent risk factor for coronary artery disease. *J Am Coll Cardiol*. 1984;4:793–801.
3. Yamada Y, Izawa H, Ichihara S, Takatsu F, Ishihara H, Hirayama H, et al. Prediction of the risk of myocardial infarction from polymorphisms in candidate genes. *N Engl J Med*. 2002;347:1916–1923.
4. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–678.
5. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007;316:1488–1491.
6. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007;316:1491–1493.
7. Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet*. 2009;41:334–341.
8. Takeuchi F, Yokota M, Yamamoto K, Nakashima E, Katsuya T, Asano H, et al. Genome-wide association study of coronary artery disease in the Japanese. *Eur J Hum Genet*. 2012;20:333–340.
9. CARDIoGRAM Consortium. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43:333–338.
10. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet*. 2011;43:339–344.
11. IBC 50K CAD Consortium. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet*. 2011;7:e1002260.
12. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, Comuzzie AG, et al. A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet*. 2002;30:210–214.
13. Wang Q, Rao S, Shen GQ, Li L, Moliterno DJ, Newby LK, et al. Premature myocardial infarction novel susceptibility locus on chromosome 1P34-36 identified by genomewide linkage analysis. *Am J Hum Genet*. 2004;74:262–271.
14. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
15. Iwakawa M, Goto M, Noda S, Sagara M, Yamada S, Yamamoto N, et al. DNA repair capacity measured by high throughput alkaline comet assays in EBV-transformed cell lines and peripheral blood cells from cancer patients and healthy volunteers. *Mutat Res*. 2005;588:1–6.
16. Yuan JS, Reed A, Chen F, Stewart CN Jr. Statistical analysis of real-time PCR data. *BMC Bioinformatics*. 2006;7:85.
17. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet*. 2002;30:97–101.
18. Hoffmann K, Lindner TH. easyLINKAGE-Plus–automated linkage analyses using large-scale SNP data. *Bioinformatics*. 2005;21:3565–3567.
19. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet*. 1998;63:861–869.
20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21:263–265.
21. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med*. 1994;330:1041–1046.
22. Collin GB, Marshall JD, Ikeda A, So WV, Russell-Eggitt I, Maffei P, et al. Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alström syndrome. *Nat Genet*. 2002;31:74–78.
23. Hearn T, Renforth GL, Spalluto C, Hanley NA, Piper K, Brickwood S, et al. Mutation of ALMS1, a large gene with a tandem repeat encoding 47 amino acids, causes Alström syndrome. *Nat Genet*. 2002;31:79–83.
24. Alstrom CH, Hallgren B, Nilsson LB, Asander H. Retinal degeneration combined with obesity, diabetes mellitus and neurogenous deafness: a specific syndrome (not hitherto described) distinct from the Laurence-Moon-Bardet-Biedl syndrome: a clinical, endocrinological and genetic examination based on a large pedigree. *Acta Psychiatr Neurol Scand Suppl*. 1959;129:1–35.

25. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 1995;11:241–247.
26. Selmer KK, Brandal K, Olstad OK, Birkenes B, Undlien DE, Egeland T. Genome-wide linkage analysis with clustered SNP markers. *J Biomol Screen*. 2009;14:92–96.
27. Middleton FA, Pato MT, Gentile KL, Morley CP, Zhao X, Eisener AF, et al. Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *Am J Hum Genet*. 2004;74:886–897.
28. John S, Shephard N, Liu G, Zeggini E, Cao M, Chen W, et al. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *Am J Hum Genet*. 2004;75:54–64.
29. Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, Perola M, et al. Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. *Am J Hum Genet*. 2000;67:1481–1493.
30. Samani NJ, Burton P, Mangino M, Ball SG, Balmforth AJ, Barrett J, et al; BHF Family Heart Study Research Group. A genomewide linkage study of 1,933 families affected by premature coronary artery disease: The British Heart Foundation (BHF) Family Heart Study. *Am J Hum Genet*. 2005;77:1011–1020.
31. Harrap SB, Zammit KS, Wong ZY, Williams FM, Bahlo M, Tonkin AM, et al. Genome-wide linkage analysis of the acute coronary syndrome suggests a locus on chromosome 2. *Arterioscler Thromb Vasc Biol*. 2002;22:874–878.
32. Ott J, Kamatani Y, Lathrop M. Family-based designs for genome-wide association studies. *Nat Rev Genet*. 2011;12:465–474.
33. Titomanlio L, De Brasi D, Buoninconti A, Sperandeo MP, Pepe A, Andria G, et al. Alström syndrome: intrafamilial phenotypic variability in sibs with a novel nonsense mutation of the ALMS1 gene. *Clin Genet*. 2004;65:156–157.
34. Marshall JD, Hinman EG, Collin GB, Beck S, Cerqueira R, Maffei P, et al. Spectrum of ALMS1 variants and evaluation of genotype-phenotype correlations in Alström syndrome. *Hum Mutat*. 2007;28:1114–1123.
35. 't Hart LM, Maassen JA, Dekker JM, Heine RJ, Maassen JA. Lack of association between gene variants in the ALMS1 gene and type 2 diabetes mellitus. *Diabetologia*. 2003;46:1023–1024.
36. Patel S, Minton JA, Weedon MN, Frayling TM, Ricketts C, Hitman GA, et al. Common variations in the ALMS1 gene do not contribute to susceptibility to type 2 diabetes in a large white UK population. *Diabetologia*. 2006;49:1209–1213.
37. Köttgen A. Genome-wide association studies in nephrology research. *Am J Kidney Dis*. 2010;56:743–758.
38. Collin GB, Cyr E, Bronson R, Marshall JD, Gifford EJ, Hicks W, et al. Alms1-disrupted mice recapitulate human Alström syndrome. *Hum Mol Genet*. 2005;14:2323–2333.
39. Jatti K, Paisey R, More R. Coronary artery disease in Alström syndrome. *Eur J Hum Genet*. 2012;20:117–118.
40. Loudon MA, Bellenger NG, Carey CM, Paisey RB. Cardiac magnetic resonance imaging in Alström syndrome. *Orphanet J Rare Dis*. 2009;4:14.
41. Zulato E, Favaretto F, Veronese C, Campanaro S, Marshall JD, Romano S, et al. ALMS1-deficient fibroblasts over-express extra-cellular matrix components, display cell cycle delay and are resistant to apoptosis. *PLoS One*. 2011;6:e19081.
42. Ross R. The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature*. 1993;362:801–809.
43. Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*. 1992;68:799–808.
44. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*. 1993;72:971–983.
45. Jewell-Motz EA, Liggett SB. An acidic motif within the third intracellular loop of the alpha2C2 adrenergic receptor is required for agonist-promoted phosphorylation and desensitization. *Biochemistry*. 1995;34:11946–11953.
46. Nichenametla SN, Ellison I, Calcagnotto A, Lazarus P, Muscat JE, Richie JP Jr. Functional significance of the GAG trinucleotide-repeat polymorphism in the gene for the catalytic subunit of gamma-glutamylcysteine ligase. *Free Radic Biol Med*. 2008;45:645–650.
47. Gayatri M, Guruprasad K. Analysis of the conformations corresponding to hexapeptide and large sequences characterized by continuous single amino acid repeats in proteins. *Protein Pept Lett*. 2010;17:1459–1465.

CLINICAL PERSPECTIVE

Myocardial infarction (MI) is a leading cause of death worldwide. Given that a family history is an independent risk factor for coronary artery disease, genetic variants are thought to contribute directly to the development of this condition. The identification of susceptibility genes for coronary artery disease or MI may thus help to identify high-risk individuals and offer the opportunity for disease prevention. Here, we identify a glutamic acid repeat polymorphism of *ALMS1* (the gene responsible for Alström syndrome) as a novel genetic risk factor for early-onset MI. A genome-wide–affected sib-pair linkage study with 221 Japanese families yielded a statistically significant logarithm of the odds score of 3.44 for chromosome 2p13 and MI. Further association analysis implicated *ALMS1* as a candidate gene within the linkage region. Validation association analysis revealed that representative single-nucleotide polymorphisms of the *ALMS1* promoter region were significantly associated with early-onset MI in both Japanese and Korean populations. Moreover, direct sequencing of the *ALMS1* coding region identified a glutamic acid repeat polymorphism in exon 1, which was subsequently found to be associated with early-onset MI. This polymorphism of *ALMS1* may thus provide insight into the pathogenesis of early-onset MI as well as contribute to the personalized prevention of this condition.

Circulation
Cardiovascular Genetics

Identification of a Glutamic Acid Repeat Polymorphism of *ALMS1* as a Novel Genetic Risk Marker for Early-Onset Myocardial Infarction by Genome-Wide Linkage Analysis

Sahoko Ichihara, Ken Yamamoto, Hiroyuki Asano, Masahiro Nakatochi, Mayo Sukegawa, Gaku Ichihara, Hideo Izawa, Akihiro Hirashiki, Fumimaro Takatsu, Hisashi Umeda, Mitsunori Iwase, Haruo Inagaki, Haruo Hirayama, Takahito Sone, Kazuhiko Nishigaki, Shinya Minatoguchi, Myeong-Chan Cho, Yangsoo Jang, Hyo-Soo Kim, Jeong E. Park, Saeko Tada-Oikawa, Hidetoshi Kitajima, Tatsuaki Matsubara, Kenji Sunagawa, Hiroaki Shimokawa, Akinori Kimura, Jong-Young Lee, Toyoaki Murohara, Ituro Inoue and Mitsuhiro Yokota

Circ Cardiovasc Genet. 2013;6:569-578; originally published online October 11, 2013;
doi: 10.1161/CIRCGENETICS.111.000027
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1942-325X. Online ISSN: 1942-3268

SUPPLEMENTAL MATERIAL

Supplemental Methods

Study subjects

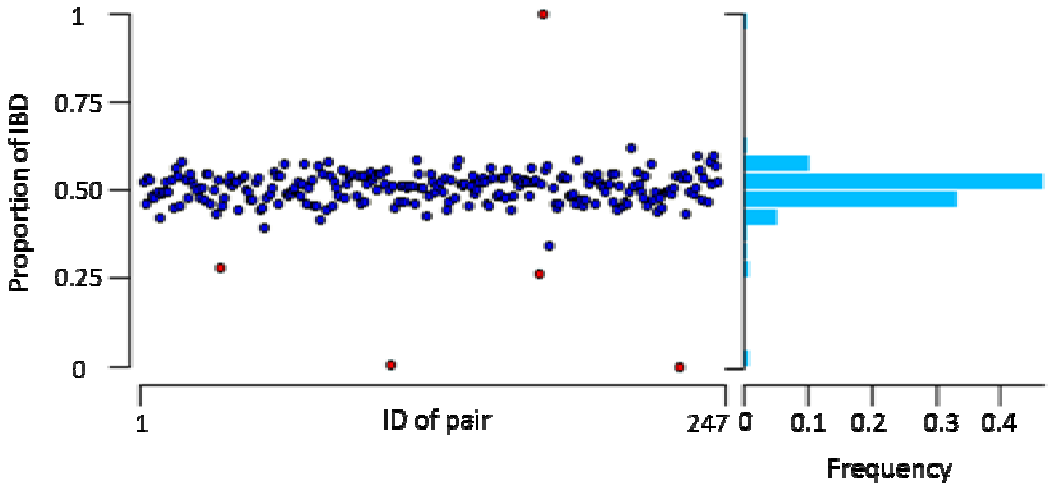
The characteristics of the subjects analyzed at each stage of the study are described in Table 1. The 455 individuals from 221 families with CAD analyzed in the first step (Figure 1) were recruited through participating hospitals in Japan. Each family has at least two affected siblings with CAD. The diagnosis of CAD was defined as the occurrence of MI or AP verified by coronary angiography. The diagnosis of MI was based on typical electrocardiographic changes and increased serum activities of enzymes including creatine kinase, aspartate aminotransferase, and lactate dehydrogenase; it was confirmed by the presence of wall motion abnormality on left ventriculography and attendant stenosis in any of the major coronary arteries. Given that the incidence of MI is low in Japan and that MI is a fatal disease, it is difficult to collect families with MI, especially early-onset MI, in the Japanese population. We therefore recruited families with CAD through participating hospitals in Japan for analysis in the first step of the study. In the present study, six families and 11 subjects with early-onset MI were included in the first and second steps, respectively. The population recruited for association analysis comprised 2,186 unrelated subjects with MI, including 351 with early-onset MI and 135 with early-onset AP, as well as 6,026 unrelated population controls in Japan (Table 1). Early-onest MI or AP was defined as the occurrence of MI or AP at ≤50 or ≤55 years of age in men or women, respectively. The control subjects were recruited from consecutive individuals who visited outpatient clinics of the participating hospitals for an annual health checkup or who were community-dwelling individuals recruited to a prospective cohort study. They had no history of ischemic heart disease or other cardiovascular diseases. In addition, 346 individuals with early-onset MI, 614 with early-onset AP, and 1,819 unrelated population controls were recruited through participating

hospitals in Korea (Table 1). Subjects with a systolic blood pressure of ≥140 mmHg or a diastolic blood pressure of ≥90 mmHg, or those who were currently taking antihypertensive medication, were categorized as having hypertension. Hyperlipidemia was defined as a serum concentration of total cholesterol of ≥220 mg/dl or the taking of lipid-lowering drugs. Type 2 diabetes was defined as a fasting plasma glucose concentration of ≥126 mg/dl, a hemoglobin A_{1c} level of ≥6.5%, or current treatment for diabetes. The study protocol was approved by the committees on the ethics of human research of Aichi-Gakuin University, Nagoya University School of Medicine, Kyushu University School of Medicine, and participating hospitals in Japan and Korea. Written informed consent was obtained from each participant.

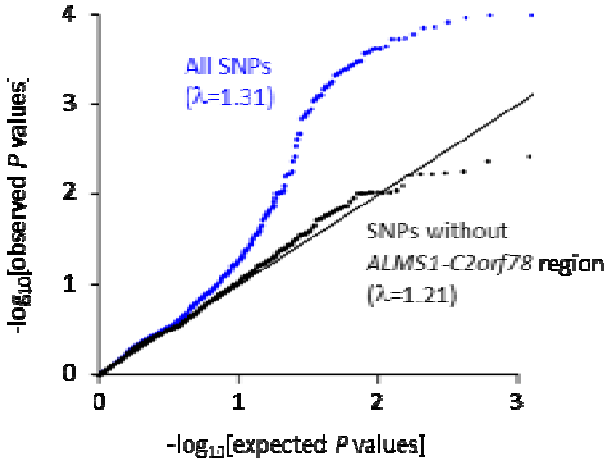
Analysis of transcription factor binding motifs at SNP sites of the *ALMS1* promoter region

To assess whether SNP sites in the *ALMS1* promoter region might contribute to transcriptional regulation, we examined the sites associated with early-onset MI for transcription factor binding motifs with the use of the TFSEARCH program (http://www.cbrc.jp/research/db/TFSEARCH.html). On the basis of the results of the association analysis shown in Figure 3 and Table VI, we selected 11 SNPs (rs6546820, rs10191517, rs7560272, rs7604588, rs6718864, rs6706179, rs7573719, rs6720094, rs6740173, rs6748040, rs6546829) that are in strong LD with each other and are located in the promoter region of *ALMS1*. The plus-strand DNA sequences (33 bases centered on the SNP site) corresponding to the major or minor allele were analyzed with the TFSEARCH program. The threshold score was set to 85.0 (default score).

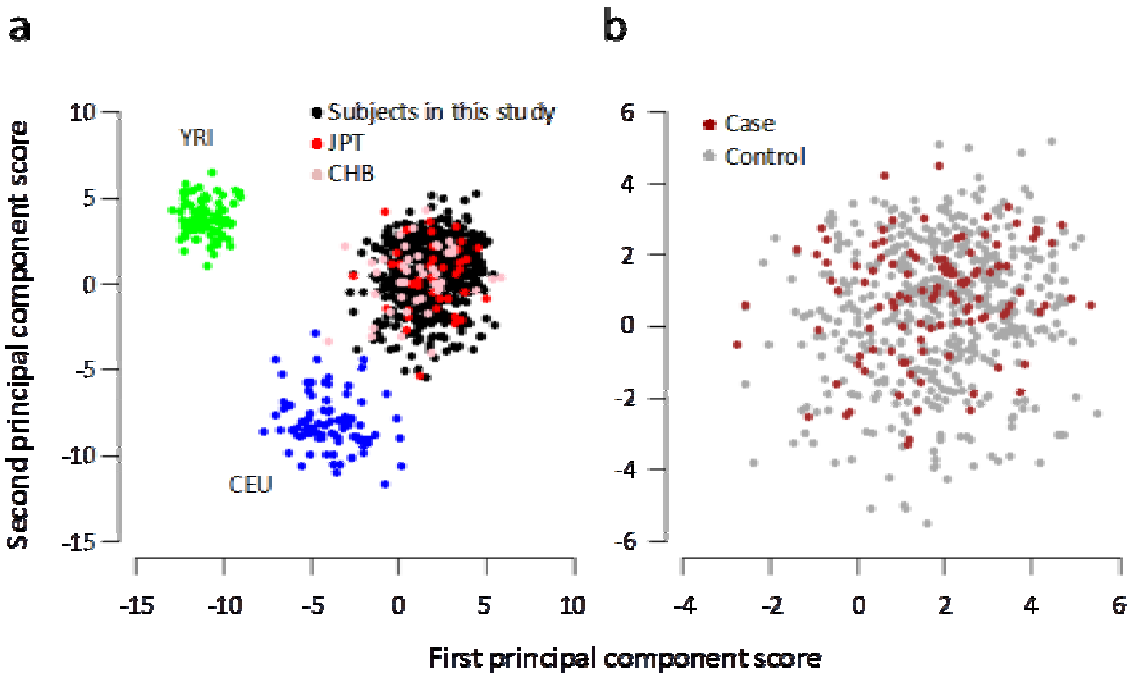
Supplemental Figure 1. Distribution of the proportion of pairwise identity-by-descent. The pairwise proportion of identity-by-descent (IBD) was calculated for 247 sib-pairs by the PLINK program with genome-wide SNP genotypes. Pairs showing a proportion of IBD of <0.3 or equal to 1 are shown as red dots (left panel) and were excluded from further linkage analysis. The distribution is summarized in a histogram (right panel), with each class spanning an interval of 0.05.



Supplemental Figure 2. Quantile-quantile plot of observed *P* values for the association of MI with SNPs in the candidate region of chromosome 2. The blue dots represent the distribution of *P* values for the association of all 1,248 SNPs in 111 probands of MI sib-pairs and 480 controls, and the black dots represent that for the association of the SNPs remaining after exclusion of the 52 SNPs within the *ALMS1-C2orf78* region (chromosome 2: 73,402,295–73,902,334).



Supplemental Figure 3. Principal component analysis (PCA) plots for subjects in the second-step association analysis focusing on chromosome 2p14-p12. a, Genotypes of 1,248 SNPs for YRI (n = 819), CEU (n = 81), CHB (n = 43), and JPT (n = 42) populations were obtained from the HapMap database and were merged with those for cases (n = 111) and controls (n = 480) in the present study. We excluded 52 SNPs located in the *ALMS1-C2orf78* region (chromosome 2: 73,402,295–73,902,334) as well as 681 SNPs whose genotypes were not fully obtained for all samples. The remaining 515 SNPs were subjected to PCA analysis. **b**, The region of the plot in **a** including the subjects of the present study is shown enlarged. The plot is suggestive of nonstratification among the samples.



Supplemental Table 1. Subjects included in the affected sib-pair (ASP) linkage analysis.

	CAD	MI subgroup
Families with two affected sibs	203	110
Families with three affected sibs	11	2
Families with four affected sibs	1	0
No. of samples	443	226
No. of families	215	112
No. of informative ASPs	228	114

Supplemental Table 2. Summary of SNPs used for linkage analysis.

Chromosome	No. of SNPs	Mean (SD) interval (cM)	Call rate		Minor allele frequency		
			Mean	Minimum	Mean (SD)	Maximum	Minimum
1	137	2.10 (0.16)	0.998	0.988	0.267 (0.132)	0.498	0.050
2	133	2.09 (0.19)	0.998	0.995	0.259 (0.128)	0.497	0.047
3	109	2.08 (0.11)	0.998	0.993	0.243 (0.131)	0.498	0.054
4	101	2.09 (0.12)	0.998	0.993	0.264 (0.134)	0.497	0.049
5	96	2.08 (0.09)	0.998	0.990	0.277 (0.132)	0.495	0.049
6	93	2.08 (0.22)	0.998	0.995	0.255 (0.129)	0.491	0.048
7	88	2.08 (0.12)	0.998	0.988	0.257 (0.117)	0.488	0.056
8	80	2.10 (0.27)	0.999	0.993	0.259 (0.134)	0.482	0.048
9	80	2.10 (0.35)	0.999	0.993	0.274 (0.133)	0.494	0.056
10	85	2.07 (0.16)	0.998	0.993	0.271 (0.126)	0.498	0.055
11	72	2.07 (0.17)	0.998	0.993	0.282 (0.121)	0.486	0.058
12	82	2.09 (0.16)	0.998	0.984	0.269 (0.135)	0.492	0.050
13	53	2.06 (0.11)	0.998	0.992	0.271 (0.125)	0.488	0.059
14	65	2.10 (0.16)	0.998	0.993	0.294 (0.131)	0.499	0.058
15	57	2.10 (0.13)	0.998	0.982	0.267 (0.127)	0.497	0.053
16	62	2.16 (0.45)	0.998	0.990	0.274 (0.127)	0.489	0.056
17	69	2.22 (0.46)	0.998	0.985	0.289 (0.136)	0.496	0.057
18	59	2.17 (0.42)	0.998	0.992	0.269 (0.124)	0.496	0.046
19	49	2.11 (0.14)	0.998	0.993	0.267 (0.126)	0.499	0.058
20	51	2.11 (0.20)	0.999	0.995	0.288 (0.129)	0.492	0.062
21	27	2.06 (0.08)	0.999	0.996	0.284 (0.140)	0.497	0.056
22	31	2.10 (0.13)	0.999	0.995	0.243 (0.122)	0.499	0.049
X	44	2.07 (0.08)	0.998	0.995	0.239 (0.124)	0.499	0.050
All	1723	2.10 (0.22)	0.998	0.982	0.267 (0.130)	0.499	0.046

Supplemental Table 3. Primers for direct sequencing of the coding region of *ALMS1*.

Primer	Sequence (5′→3′)	Target exon	Expected product size (bp)
ex01-1F	CAACGTCGCCTGTAGCAA	1	303
ex01-1R	GCAGCCTCCTCCTCTTCC		
ex01-2F	CTAAGCTGGGCCACAACC	1	320
ex01-2R	TCGTCCTCCTCGTCGTCTAT		
ex01-3F	AACATGGAGCCCGAGGAT	1	380
ex01-3R	AGCCTCCACCCCCAACTC		
ex02-F	TGTGAAAGGGCTTTATAAACTGG	2	296
ex02-R	TCAGAATTTCTTGTGCTGCTT		
ex03-F	GATGAGAATGTCATTAAATAGGAAGC	3	375
ex03-R	GTCCCCGCAAAACTTACAAA		
ex04-F	GCTGCTACTGCTGTTGCTTTT	4	295
ex04-R	AAGGTTACATGCAAATAAATCATCC		
ex05-1F	TCTGAAATTAGGAGAGCTGTGTTTT	5	393
ex05-1R	AACGATCACAGTCTTTGGGAAT		
ex05-2F	TGTATCTCAGCACCCGCTTA	5	385
ex05-2R	TGAATCTGAAAAAGAATCAGGGTA		
ex06-F	AGAGACATTGCTGAAAGCTGA	6	250
ex06-R	GGCCACGTTTCTTACAACTATC		
ex07-F	TTTTTAGCAGATGAGGTTTGAAAT	7	247
ex07-R	GGACATTTACTATCTCCTTTTTCTACG		
ex08-1F	CCTTTGATGGCTGTTTCCTT	8	400
ex08-1R	TGGTGTTGCAGTCTTCTGGT		
ex08-2F	AAACTACTACTGGTCAACACACTGA	8	400
ex08-2R	CCCGTCTTCTGCTCCACTG		
ex08-3F	TGACCAGACAACTGGCATGT	8	382
ex08-3R	TTGATGACTGTCTGCGAACTC		
ex08-4F	GACAGACCTTGCCAGATGGT	8	436
ex08-4R	AAGGCCTGTTGGTAGAAAACAA		
ex08-5F	CCCACAGGACTTAGCAGACA	8	388
ex08-5R	TGCTGAGGGTATCCCAGTCT		
ex08-6F	CAGCAGACCTTACCCAATAGTCA	8	370
ex08-6R	TTGGGACAGTCTTCTGGTCA		

ex08-7F	AGACTGGGACACCAACAGTG	8	396
ex08-7R	CTTTCAGACTCTCTTCAGGTAGATG		
ex08-8F	TGCAACTGAAAAGGCTCTGA	8	399
ex08-8R	AAAATGCTGGGCTTTTCTCT		
ex08-9F	GCCAGAGAGTCATCTGCCTAA	8	381
ex08-9R	GAAGCAGAGGTTGGAGTTGG		
ex08-10F	TTGGATACCAAGAGTACTTTCTACCTT	8	407
ex08-10R	CAGGAACCGCTGAAACATTC		
ex08-11F	GGCACACCAGCTGTAACCTC	8	223
ex08-11R	TATGTGGCAAGACCTGTTGG		
ex08-12F	GAATGTTTCAGCGGTTCCTG	8	233
ex08-12R	TCGGCTTCTCTGTATGTTGTG		
ex08-13F	TACCAACAGGTCTTGCCACA	8	228
ex08-13R	AGTCTTCCGGTCACCTGGTC		
ex08-14F	CCAACTGAAGAGGCTCTGAAAA	8	374
ex08-14R	GAGCTGGAAGGGGAGGTTAC		
ex08-15F	GTGACCGGAAGACTGGGATA	8	421
ex08-15R	CCAGGAGCAGAAGAAACTCTG		
ex08-16F	CAACTGGCGCACCAACTA	8	390
ex08-16R	CAACCTCTTTATTTAGAGGACTGTCT		
ex08-17F	AGGCTCTGAAAGTTTCCATTG	8	399
ex08-17R	GGGCCTGTTGGTAGAAGACA		
ex08-18F	CCAGCAGACCCTATCAGACA	8	391
ex08-18R	GGAAGTAGAGGTTACTGTTGATACCC		
ex08-19F	TTTCAGCTGTTCCTCAACCA	8	464
ex08-19R	GGATGCTATTTGTATCCCAGTCTT		
ex08-20F	TCCTGCCCTCTAATTCCTACC	8	393
ex08-20R	GCTCTGCTGGTATAGAAACAGGT		
ex08-21F	ACAAGGTGACCGGAAGACTG	8	390
ex08-21R	TTGGGCTTTACTGTTTGAGAA		
ex08-22F	CCAGCAGCTACCCTTAGTTCC	8	383
ex08-22R	GGAAGAGCTGTTGGTAATACTGTTT		
ex08-23F	TTATTTTCACAGAGAGAAATCGAA	8	374
ex08-23R	TTGAGCTAACACTGGAGTCAGAA		
ex08-24F	CTTGGGCAAGCTGATCAAA	8	399
ex08-24R	TTTGAAGCAAACCACCTTCTT		
ex08-25F	TTCGGACACTTTTGATGGAG	8	369

ex08-25R	TGGCTGCTTGATATTGCCTA		
ex08-26F	GGAGAGGCATTGAAAATTGG	8	389
ex08-26R	GCTACCACCACCATCCTCAC		
ex08-27F	CCTGCTCGTGGGACAGTAAT	8	296
ex08-27R	CGTTTATAAGCCTGAAACTTTCTTC		
ex09-F	TTGCAGTGGGTATTAAATTGC	9	295
ex09-R	TTCCATCACCCATTCTTTCA		
ex10-1F	TGCTGTGCTCTTTGTCCTGT	10	400
ex10-1R	TGAAATGGGAGTTGTTCTGG		
ex10-2F	AGACACCCTTGTGCTTTCAGA	10	400
ex10-2R	CAACAGAGGAATTGGAAATGC		
ex10-3F	AAAGATGGACCCTTGGCTGT	10	400
ex10-3R	TTGAAAAGAACGCTCAAAATCA		
ex10-4F	AAAACCTCTTCCCCTTCATCA	10	385
ex10-4R	GTGGGGCTTTGCTTTGTTC		
ex10-5F	CCCTTATTAGCATGGGCAGA	10	391
ex10-5R	GCTTTGCTTTGGCATTGTT		
ex10-6F	GAACAGTGCAAAGCCCCATA	10	382
ex10-6R	TGGAAGAAGTCTTGGGACAAA		
ex10-7F	ACTGTGGTCTCCCAGTCAGC	10	389
ex10-7R	GAAGGTATGGTTTTCAAGTCCTG		
ex10-8F	ATCACTGGAACCAACCTCCA	10	299
ex10-8R	AGTCCATCCAGCTTGCTTGT		
ex11-F	CATTGATGTGTCCACAATATATTCCTA	11	397
ex11-R	CAATCTAAAATGCAATCACATTCTCT		
ex12-F	TCATTCTTCTTGAAGGCAGAGA	12	337
ex12-R	GCCAATACACAAATAAAATATCAAACA		
ex13-F	TCCAAAGTGCTGGGATTACA	13	371
ex13-R	GATGAAAAGGGGAGAAAAAGC		
ex14-F	CACAGGTGGAGCCTAAGAGG	14	298
ex14-R	CGCAGACAATCCAGTTTAGC		
ex15-F	TCAATATCAGTAACAAAGCCTTTCA	15	300
ex15-R	CATACACAACCCAATCCCATT		
ex16-1F	TTCTTTCCTTTAGTCTTTGTTCTGT	16	387
ex16-1R	TCCCAAATTCACATTGATGC		
ex16-2F	TGTTTTCATGAGACATTCTTGGA	16	390
ex16-2R	TTAGCCAAACGATCAAGTCG		

ex16-3F	TCCACAAAGGGATCAGAAGG	16	445
ex16-3R	AAAACCTGGCTGTTTACTCAGAA		
ex16-4F	GCGGTTTAAAAGCCTAGAGAAA	16	388
ex16-4R	AATTCGTCTGGGTGACAAGC		
ex16-5F	CTTTCCGGCACCACTTCTAC	16	369
ex16-5R	TTTTCACCTGTGTGCAAAGC		
ex17-F	CCTCCAAATGCATCTCAACA	17	295
ex17-R	AAGAGCCATTTCAATGTATTTCA		
ex18-F	ATCCCACACAAAGGGATTGT	18	385
ex18-R	ATCGCAGGGGACTTGAAAT		
ex19-F	CTGCAGCAAAACCAGACTCA	19	445
ex19-R	TTGCCCTTCTTCTCCCTACA		
ex20-F	TCAGACTTCCCCAAACCTCTT	20	348
ex20-R	CAGCAGACCTTGCCTTCTCT		
ex21-F	GTCTTGCCAGCTCAGGGTAG	21	230
ex21-R	TGGTGGCCTGGTTTACACTT		
ex22-F	GGATGAGCTCCTGGAGAGTG	22	250
ex22-R	CTGATCCATCCCACTCTAGACA		
ex23-1F	GACCACACTGATTCTCCTTGG	23	359
ex23-1R	ATTGGACAAGGCCACTATGC		
ex23-2F	TGAACCTAGAGAAGCAGAATCC	23	296
ex23-2R	TCTCCAGATGGGAAAGAATTG		

Supplemental Table 4. Multipoint LOD scores from linkage analysis for CAD (all) and MI sib-pairs with SNP markers on chromosome 2p16-q14.

SNP	Genetic location (cM)	Base position	Call rate	MAF	Multipoint LOD by linear model		Multipoint LOD by exponential model	
					All pairs	MI pairs	All pairs	MI pairs
rs805341	76.29	54,045,661	1	0.40	1.32	1.23	1.22	1.00
rs973966	78.32	59,199,119	1	0.22	1.37	1.43	1.23	1.12
rs842639	80.32	60,948,749	1	0.20	1.62	1.78	1.39	1.36
rs1545196	82.40	64,135,814	1	0.08	1.75	2.12	1.51	1.67
rs6706382	84.43	66,668,779	1	0.48	1.73	2.38	1.50	1.97
rs751759	86.46	68,204,119	1	0.47	1.76	2.52	1.51	2.11
rs7562668	88.46	71,051,131	1	0.33	1.90	3.10	1.61	2.62
rs1406105	90.50	73,527,769	1	0.26	1.66	3.44	1.38	2.91
rs4853033	92.95	74,681,207	1	0.14	1.30	3.33	1.07	2.87
rs2422388	95.05	75,998,457	1	0.15	1.21	3.08	1.02	2.70
rs6547142	97.06	77,358,673	1	0.16	1.32	2.90	1.14	2.47
rs13422559	99.08	78,594,435	1	0.15	1.31	2.86	1.17	2.48
rs10177878	101.12	79,305,167	1	0.40	1.36	2.70	1.24	2.36
rs11678886	103.13	80,951,713	1	0.47	1.25	2.52	1.17	2.26
rs1019837	105.15	83,268,591	1	0.29	1.18	2.43	1.11	2.14
rs2028898	107.20	85,630,781	1	0.30	1.11	2.37	1.05	2.06
rs12714206	109.62	88,080,959	1	0.19	0.94	2.11	0.89	1.81
rs11692699	111.62	99,097,089	1	0.10	0.68	1.86	0.65	1.58
rs2056400	113.75	101,913,233	1	0.31	0.49	1.65	0.47	1.38
rs6743006	115.77	104,644,590	1	0.26	0.32	1.36	0.31	1.15
rs7419384	117.81	106,254,302	1	0.47	0.18	0.90	0.17	0.80
rs260642	119.82	108,903,216	1	0.16	0.07	0.52	0.07	0.46
rs11890171	121.82	111,331,237	1	0.09	0.03	0.31	0.03	0.28
rs1549775	123.85	114,684,697	1	0.08	0.01	0.16	0.01	0.15

MAF, minor allele frequency.

Supplemental Table 5. Multipoint LOD scores from linkage analysis for CAD (all) and MI sib-pairs with microsatellite markers on chromosome 2p16-q13.

Marker	Genetic location (cM)	Call rate	No. of observed alleles	Heterozygosity	Multipoint LOD by linear model		Multipoint LOD by exponential model	
					All pairs	MI pairs	All pairs	MI pairs
D2S2153	76.88	0.998	14	0.768	0.87	0.98	0.79	0.82
D2S337	80.69	0.968	16	0.823	0.99	1.42	0.87	1.19
D2S2320	82.29	0.995	17	0.871	0.95	1.37	0.83	1.11
D2S2368	85.48	0.986	16	0.837	0.66	1.63	0.63	1.42
D2S2110	90.82	0.998	9	0.743	0.63	2.21	0.57	1.79
D2S286	94.05	0.968	10	0.762	0.47	1.92	0.42	1.64
D2S438	99.41	0.995	7	0.696	0.36	0.77	0.35	0.68
D2S2333	103.16	0.975	10	0.833	0.19	0.56	0.19	0.51
D2S2216	111.21	0.986	9	0.709	0.33	0.84	0.35	0.83
D2S160	122.96	0.957	9	0.699	0.13	0.39	0.14	0.41

Supplemental Table 6. Results of the second-step association analysis between 52 SNPs located in the *ALMS1-C2orf78* region and probands of MI sib-pairs.

SNP	Position	MAF		Allele test		Additive model			r^2	Dominant model		Recessive model	
		Case	Control	OR (95% CI)	*P*	OR (95% CI)	*P*	*P* (conditioned)*	value†	OR (95% CI)	*P*	OR (95% CI)	*P*
rs6706562	73,402,295	0.30	0.19	1.79 (1.29-2.49)	0.00040	1.77 (1.28-2.46)	0.00060	0.70	0.75	1.82 (1.20-2.76)	0.0047	3.11 (1.45-6.66)	0.0035
rs6546820	73,406,905	0.37	0.24	1.81 (1.33-2.47)	0.00014	1.81 (1.32-2.48)	0.00020	0.75	0.93	2.14 (1.40-3.26)	0.00041	2.15 (1.10-4.21)	0.025
rs10191517	73,411,911	0.36	0.23	1.84 (1.34-2.51)	0.00011	1.81 (1.32-2.48)	0.00018			2.12 (1.39-3.23)	0.00044	2.24 (1.14-4.40)	0.019
rs7560272	73,414,993	0.36	0.24	1.77 (1.30-2.43)	0.00027	1.76 (1.29-2.42)	0.00038		0.99	2.08 (1.36-3.16)	0.00063	2.05 (1.03-4.10)	0.040
rs7604588	73,424,925	0.37	0.24	1.83 (1.34-2.50)	0.00010	1.78 (1.31-2.41)	0.00021	0.77	0.95	2.16 (1.42-3.30)	0.00032	2.05 (1.07-3.91)	0.030
rs6718864	73,428,099	0.38	0.26	1.80 (1.32-2.45)	0.00014	1.76 (1.30-2.39)	0.00026	0.65	0.88	2.12 (1.38-3.24)	0.00050	2.07 (1.10-3.88)	0.023
rs6706179	73,432,085	0.37	0.24	1.80 (1.32-2.46)	0.00017	1.75 (1.29-2.38)	0.00032	0.92	0.94	2.10 (1.37-3.20)	0.00058	2.07 (1.08-3.95)	0.027
rs7573719	73,436,494	0.37	0.24	1.83 (1.34-2.50)	0.00010	1.78 (1.31-2.41)	0.00021	0.77	0.95	2.16 (1.42-3.30)	0.00032	2.05 (1.07-3.91)	0.030
rs6720094	73,441,584	0.38	0.25	1.79 (1.31-2.43)	0.00018	1.75 (1.29-2.37)	0.00032	0.83	0.89	2.15 (1.41-3.29)	0.00038	1.92 (1.01-3.65)	0.046
rs6740173	73,455,774	0.36	0.24	1.78 (1.30-2.43)	0.00025	1.72 (1.27-2.34)	0.00047	0.87	0.90	1.99 (1.31-3.02)	0.0013	2.18 (1.14-4.19)	0.019
rs6748040	73,463,436	0.36	0.23	1.80 (1.32-2.46)	0.00019	1.74 (1.28-2.36)	0.00038	0.98	0.91	2.02 (1.33-3.07)	0.00095	2.18 (1.14-4.19)	0.019
rs6546829	73,467,297	0.35	0.24	1.74 (1.27-2.39)	0.00042	1.70 (1.25-2.31)	0.00073	0.64	0.90	1.99 (1.31-3.02)	0.0013	2.02 (1.03-3.93)	0.038
rs1881246	73,503,630	0.36	0.25	1.70 (1.25-2.32)	0.00068	1.65 (1.22-2.23)	0.0012	0.91	0.77	1.96 (1.29-2.99)	0.0016	1.87 (0.98-3.54)	0.055
rs12996463	73,509,004	0.35	0.24	1.72 (1.26-2.35)	0.00060	1.66 (1.22-2.26)	0.0011	0.95	0.78	1.99 (1.31-3.04)	0.0012	1.85 (0.96-3.58)	0.066
rs2178154	73,515,246	0.36	0.24	1.73 (1.27-2.37)	0.00047	1.67 (1.23-2.26)	0.00090	0.97	0.79	1.96 (1.29-2.98)	0.0015	1.98 (1.04-3.77)	0.037
rs1528169	73,522,694	0.34	0.24	1.66 (1.21-2.27)	0.0014	1.60 (1.18-2.18)	0.0024	0.94	0.65	1.93 (1.27-2.93)	0.0019	1.68 (0.86-3.30)	0.13
rs1406105	73,527,769	0.35	0.24	1.75 (1.28-2.40)	0.00037	1.68 (1.24-2.28)	0.00076	0.86	0.75	1.96 (1.29-2.98)	0.0015	2.05 (1.07-3.91)	0.030
rs6749680	73,539,360	0.36	0.25	1.67 (1.22-2.28)	0.0011	1.62 (1.19-2.20)	0.0018	0.86	0.76	1.90 (1.24-2.89)	0.0028	1.87 (0.99-3.56)	0.054
rs1246105	73,547,668	0.35	0.23	1.79 (1.30-2.45)	0.00024	1.71 (1.26-2.32)	0.00052	0.73	0.77	2.00 (1.31-3.04)	0.0011	2.11 (1.10-4.05)	0.024
rs780395	73,554,361	0.36	0.25	1.68 (1.23-2.29)	0.00087	1.63 (1.21-2.22)	0.0014	0.83	0.76	1.93 (1.27-2.94)	0.0021	1.87 (0.98-3.54)	0.055
rs1246096	73,558,596	0.35	0.23	1.76 (1.29-2.42)	0.00034	1.69 (1.24-2.30)	0.00070	0.78	0.76	1.95 (1.28-2.97)	0.0017	2.13 (1.11-4.09)	0.022

rs3820700	73,570,318	0.35	0.23	1.79 (1.31-2.45)	0.00023	1.71 (1.26-2.32)	0.00050	0.68	0.76	2.01 (1.32-3.05)	0.0010	2.11 (1.10-4.04)	0.024
rs7598660	73,593,933	0.26	0.37	0.59 (0.43-0.83)	0.0021	0.60 (0.43-0.83)	0.0026	0.10	0.15	0.51 (0.33-0.77)	0.0017	0.57 (0.27-1.19)	0.14
rs11685372	73,602,969	0.35	0.23	1.79 (1.30-2.45)	0.00024	1.72 (1.26-2.33)	0.00050	0.64	0.73	1.98 (1.31-3.02)	0.0013	2.17 (1.13-4.17)	0.019
rs6744697	73,618,946	0.36	0.25	1.65 (1.21-2.26)	0.0013	1.62 (1.19-2.20)	0.0020	0.83	0.74	1.92 (1.26-2.93)	0.0022	1.76 (0.91-3.40)	0.088
rs13008860	73,627,527	0.25	0.37	0.58 (0.41-0.81)	0.0012	0.58 (0.42-0.81)	0.0016	0.060	0.14	0.49 (0.32-0.75)	0.00095	0.56 (0.27-1.16)	0.12
rs17349804	73,636,668	0.36	0.25	1.69 (1.24-2.31)	0.00080	1.64 (1.21-2.22)	0.0014	0.90	0.67	1.92 (1.26-2.92)	0.0022	1.92 (1.01-3.65)	0.046
rs7566315	73,645,316	0.36	0.24	1.74 (1.27-2.38)	0.00044	1.68 (1.24-2.28)	0.00081	0.98	0.80	1.95 (1.28-2.97)	0.0017	2.05 (1.07-3.91)	0.030
rs2948441	73,679,278	0.36	0.23	1.80 (1.32-2.46)	0.00019	1.73 (1.27-2.35)	0.00040	0.58	0.75	2.04 (1.34-3.10)	0.00082	2.11 (1.10-4.05)	0.024
rs4852937	73,687,592	0.45	0.35	1.50 (1.12-2.02)	0.0063	1.46 (1.09-1.94)	0.0094	0.93	0.41	1.62 (1.05-2.51)	0.029	1.74 (1.04-2.92)	0.033
rs7570014	73,699,311	0.36	0.23	1.83 (1.34-2.51)	0.00012	1.76 (1.30-2.39)	0.00027	0.44	0.74	2.07 (1.36-3.15)	0.00061	2.18 (1.14-4.19)	0.019
rs4241256	73,707,017	0.44	0.35	1.45 (1.07-1.96)	0.015	1.44 (1.06-1.94)	0.017	0.84	0.34	1.57 (1.00-2.46)	0.046	1.70 (0.98-2.95)	0.058
rs4852940	73,710,510	0.44	0.33	1.59 (1.18-2.14)	0.0021	1.54 (1.15-2.05)	0.0033	0.46	0.34	1.89 (1.22-2.93)	0.0044	1.66 (0.97-2.84)	0.064
rs2421588	73,715,732	0.43	0.31	1.67 (1.24-2.25)	0.00065	1.61 (1.21-2.16)	0.0012	0.28	0.38	2.01 (1.30-3.11)	0.0016	1.76 (1.01-3.06)	0.043
rs4852951	73,723,304	0.44	0.32	1.63 (1.21-2.20)	0.0011	1.59 (1.19-2.14)	0.0017	0.42	0.43	1.94 (1.25-3.02)	0.0029	1.76 (1.01-3.06)	0.043
rs4852959	73,734,109	0.39	0.27	1.67 (1.23-2.26)	0.00090	1.61 (1.20-2.17)	0.0016	0.58	0.54	1.77 (1.16-2.70)	0.0073	2.14 (1.19-3.86)	0.011
rs2421669	73,740,138	0.33	0.22	1.77 (1.28-2.43)	0.00040	1.68 (1.23-2.29)	0.00088	0.65	0.65	1.92 (1.27-2.91)	0.0020	2.16 (1.10-4.23)	0.024
rs2421581	73,743,245	0.33	0.22	1.79 (1.30-2.47)	0.00027	1.71 (1.25-2.32)	0.00061	0.54	0.66	1.97 (1.30-3.00)	0.0013	2.16 (1.10-4.23)	0.024
rs4511748	73,755,999	0.37	0.25	1.74 (1.28-2.38)	0.00036	1.67 (1.23-2.25)	0.00076	0.41	0.54	1.92 (1.26-2.93)	0.0021	2.10 (1.13-3.88)	0.018
rs2421574	73,762,275	0.37	0.27	1.60 (1.18-2.18)	0.0024	1.54 (1.14-2.07)	0.0041	0.54	0.37	1.84 (1.21-2.81)	0.0042	1.66 (0.90-3.07)	0.10
rs7608328	73,767,527	0.36	0.27	1.54 (1.13-2.09)	0.0059	1.50 (1.11-2.03)	0.0081	0.75	0.38	1.80 (1.18-2.75)	0.0057	1.50 (0.79-2.86)	0.21
rs2947845	73,783,864	0.43	0.32	1.62 (1.20-2.18)	0.0014	1.56 (1.17-2.09)	0.0024	0.42	0.38	1.99 (1.29-3.09)	0.0019	1.59 (0.91-2.79)	0.098
rs1815028	73,792,434	0.41	0.30	1.66 (1.23-2.25)	0.00080	1.60 (1.19-2.14)	0.0015	0.38	0.44	2.03 (1.32-3.13)	0.0012	1.66 (0.94-2.94)	0.080
rs4852978	73,796,618	0.42	0.30	1.66 (1.23-2.24)	0.00080	1.59 (1.19-2.13)	0.0015	0.36	0.42	1.98 (1.29-3.05)	0.0018	1.73 (0.98-3.03)	0.054
rs2006997	73,801,354	0.42	0.30	1.68 (1.24-2.27)	0.00064	1.61 (1.20-2.16)	0.0012	0.31	0.43	2.02 (1.31-3.12)	0.0015	1.75 (1.00-3.07)	0.050
rs7210	73,810,632	0.43	0.31	1.68 (1.25-2.27)	0.00057	1.63 (1.21-2.18)	0.0010	0.28	0.42	2.06 (1.33-3.19)	0.0011	1.73 (0.98-3.03)	0.054
rs11126416	73,827,385	0.45	0.32	1.69 (1.26-2.28)	0.00046	1.66 (1.24-2.24)	0.00069	0.23	0.44	2.10 (1.34-3.28)	0.0011	1.80 (1.03-3.13)	0.036

rs2272178	73,831,117	0.43	0.33	1.56 (1.16-2.11)	0.0030	1.55 (1.15-2.09)	0.0036	0.51	0.39	1.83 (1.18-2.84)	0.0066	1.73 (0.98-3.03)	0.054
rs12624267	73,840,439	0.45	0.33	1.62 (1.20-2.19)	0.0012	1.59 (1.19-2.13)	0.0018	0.41	0.44	2.02 (1.29-3.15)	0.0019	1.66 (0.96-2.87)	0.069
rs2421559	73,853,162	0.44	0.33	1.62 (1.20-2.18)	0.0014	1.58 (1.18-2.12)	0.0020	0.44	0.43	2.07 (1.33-3.23)	0.0013	1.56 (0.89-2.72)	0.11
rs2272051	73,860,644	0.44	0.31	1.73 (1.28-2.35)	0.00033	1.66 (1.24-2.23)	0.00068	0.20	0.42	1.99 (1.28-3.10)	0.0022	1.99 (1.15-3.45)	0.013
rs2462127	73,902,334	0.43	0.31	1.65 (1.22-2.23)	0.00090	1.61 (1.20-2.17)	0.0014	0.36	0.41	2.00 (1.29-3.11)	0.0019	1.75 (1.00-3.07)	0.050

[*]Association analysis was performed by conditioning with rs10191517. [†]The r^2 values with rs10191517 are shown.

Supplemental Table 7. Linkage disequilibrium (LD) between two representative SNPs (rs10191517 and rs6748040) and the glutamic acid repeat polymorphism in exon 1 of *ALMS1*.

	Glutamic acid repeat allele						
	A14 (0.264)*	A15 (0.034)	A16 (0.075)	A17 (0.487)	A18 (0.079)	A19 (0.017)	A20 (0.036)
With rs10191517							
r^2	0.779	0.002	0.028	0.294	0.031	0.006	0.013
Hf†	0.242	0.013	0.000	0.010	0.000	0.000	0.000
With rs6748040							
r^2	0.839	0.003	0.028	0.325	0.031	0.006	0.013
Hf†	0.248	0.013	0.001	0.001	0.000	0.000	0.000

*Frequency of each allele is shown in parentheses. †Frequency of the haplotype (Hf) constructed with the risk allele of the indicated SNP and each allele of the repeat polymorphism.

Supplemental Table 8. Transcription factors predicted to bind to SNP sites in the *ALMS1* promoter region that are associated with early-onset MI.

SNP	Minor allele*	Major allele
rs6546820	None	None
rs10191517	SOX5	SOX5
rs7560272	SOX5	None
rs7604588	None	None
rs6718864	None	None
rs6706179	NKX-2, CHOP-C	None
rs7573719	HFH-1, HFH-2, HNF-3b, S8, CdxA	HFH-1, HFH-2, HNF-3b, S8, CdxA
rs6720094	None	None
rs6740173	EVI1	EVI1, CP2
rs6748040	GATA-1, GATA-2, EVI1	GATA-1, GATA-2
rs6546829	None	None

A 33-nucleotide sequence centered on each SNP was analyzed with the TFSEARCH program (http://www.cbrc.jp/research/db/TFSEARCH.html). *The minor allele of each SNP was a risk factor for early-onset MI.

Supplemental Table 9. Relation between nongenetic factors and the glutamic acid repeat polymorphism in exon 1 of *ALMS1* for early-onset MI cases.

Variable	Long/long (n = 93)	Short/long (n = 180)	Short/short (n = 73)	P
Sex (% male)	84.9	82.2	84.9	0.793
Age at recruitment (years)	49.8 ± 8.4	48.7 ± 7.9	48.3 ± 7.8	0.412
Age at first event (years)	45.4 ± 4.4	44.2 ± 5.4	44.6 ± 5.3	0.204
Body mass index (kg/m^2)	24.7 ± 3.0	24.4 ± 3.1	24.7 ± 3.5	0.656
Former or current smoker (%)	75.3	76.1	67.1	0.318
Angina pectoris (%)	10.8	12.8	13.7	0.832
Hypertension (%)	43.0	45.6	42.5	0.871
Diabetes mellitus (%)	30.1	32.8	21.9	0.230
Hyperlipidemia (%)	59.1	63.9	60.3	0.711

Data for continuous variables are means ± SD. Differences in variables among genotypes were evaluated by one-way ANOVA or the chi-squared test. Alleles were categorized as short (A9 to A16) or long (A17 to A22).

Supplemental Table 10. Association analysis for early-onset MI and the glutamic acid repeat polymorphism in exon 1 of *ALMS1* conditional on nongenetic factors.

	OR (95% CI)	*P*
Model 1	1.48 (1.26–1.73)	1.17×10^{-6}
Model 2	1.48 (1.26–1.73)	1.54×10^{-6}
Model 3	1.50 (1.28–1.76)	7.63×10^{-7}

Alleles were classified as short (A9 to A16) or long (A17 to A22), with the results for the short-repeat group being shown. Model 1 was adjusted for sex. Model 2 was adjusted for sex and diabetes mellitus. Model 3 was adjusted for sex, diabetes mellitus, body mass index, smoking, hypertension, and hyperlipidemia.