J Hum Genet (2006) 51:151–154
DOI 10.1007/s10038-005-0337-6

SHORT COMMUNICATION

**Aritoshi Iida · Hideki Kizawa · Yusuke Nakamura
Shiro Ikegawa**

High-resolution SNP map of *ASPN*, a susceptibility gene for osteoarthritis

Received: 22 September 2005 / Accepted: 20 October 2005 / Published online: 26 November 2005
© The Japan Society of Human Genetics and Springer-Verlag 2005

Abstract Osteoarthritis (OA) is a very common bone and joint disease characterized by breakdown of cartilage in the joint. We recently found that an aspartic-acid repeat polymorphism of the asporin gene (*ASPN*) on chromosome 9 is associated with susceptibility to OA in Japanese. We provide here a high-resolution single nucleotide polymorphism (SNP) map within a 33.4-kb genomic region containing *ASPN*. A total of 19 SNPs were isolated from the region by systematic screening using 48 Japanese patients with OA: 7 SNPs in the 5′ flanking region, 8 in introns, and 4 in the 3′ untranslated region. Nine SNPs were novel. This high-resolution SNP map will be a useful resource for analyzing genes associated with OA and other bone and joint diseases.

Keywords Asporin · Osteoarthritis · Single nucleotide polymorphism · High-resolution SNP map · Association study

Introduction

Asporin was first identified as a new member of the class I small leucine-rich repeat proteoglycan (SLRP) family (Lorenzo et al. 2001; Henry et al. 2001) that is up-regulated in the early stages of osteoarthritis (OA) (Lorenzo et al. 2001). The cDNA of the human asporin gene (*ASPN*); GenBank accession number, NM_017680.2) is 2,466 bp long; it encodes a 382 amino-acid protein with polymorphic aspartic acid (D) repeat in the N-terminus. The predicted amino-acid sequence of *ASPN* is 91% identical to its mouse counterpart. The human gene was mapped to chromosome 9q22, as part of an SLRP gene cluster composed of the genes encoding extracellular matrix protein 2, osteomodulin, and osteoglycin (Henry et al. 2001).

Osteoarthritis is a common disease that is characterized by breakdown of cartilage in the joint. In a recent case-control association study, we identified *ASPN* as one of the susceptibility genes for OA (Kizawa et al. 2005). In two independent Japanese populations, a $(GAT)_{14}$ allele of *ASPN* encoding 14 aspartic acid repeats (D14) is associated with knee OA; the frequency of this allele increased with radiographic severity of OA. In addition, the D14 allele is also over-represented in hip OA. Subsequent functional analyses revealed that asporin acts a negative regulator of chondrogenesis—it inhibits function of TGF-β through direct physical interaction with TGF-β.

In this report, we provide a high-resolution single nucleotide polymorphism (SNP) map within the 33.4-kb genomic region containing *ASPN*, in which we detected a total of 19 SNPs, nine of which are novel.

Materials and methods

We screened for SNPs in the 33.4-kb region containing *ASPN* by polymerase chain reaction (PCR) followed by direct sequence analysis as described previously (Iida et al. 2005). We used genomic DNAs from 48 Japanese individuals with knee OA. We obtained written informed consent to participation in this study, which was approved by the ethical committee of the RIKEN SNP Research Center.

A. Iida (✉) · Y. Nakamura
Laboratory for Pharmacogenetics, RIKEN SNP Research Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato, Tokyo 108-8639, Japan
E-mail: iidaari@src.riken.go.jp
Tel.: +81-3-54495785
Fax: +81-3-54495785

H. Kizawa · S. Ikegawa (✉)
Laboratory for Bone and Joint Diseases, RIKEN SNP Research Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan
E-mail: sikegawa@ims.u-tokyo.ac.jp

Y. Nakamura
Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

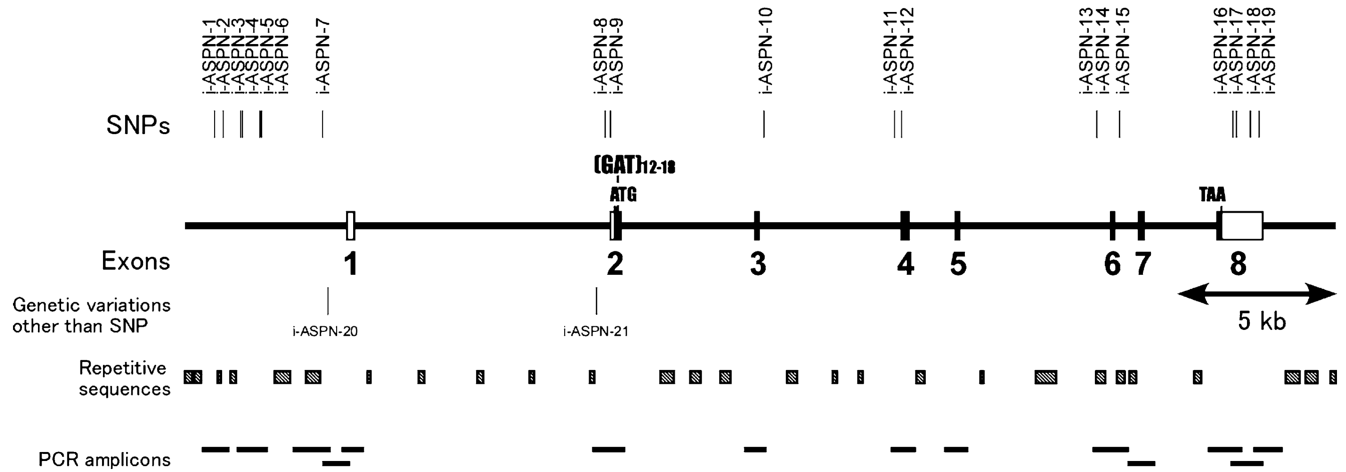

Fig. 1 Genomic organization of the asporin gene (*ASPN*) and locations of 21 sequence variations within the gene. The eight exons of *ASPN* are represented by *rectangles*. The PCR amplicons are drawn as *horizontal bars* below the map. Single nucleotide polymorphisms (SNPs—corresponding to the left-most column of Table 1) are indicated above the "chromosome" line. The locations of two sequence variations other than SNPs are indicated below the map

We defined the exon–intron structures of *ASPN* by comparing its cDNA and genomic sequences (accession numbers: AL157827.17 and AL137848.5, respectively). By direct sequencing of 14 PCR amplicons using a total of 48 sequence primers, we explored SNPs in the 5′-flanking region, all exons, intronic sequences flanking each exon, and the 3′-flanking region. We excluded repetitive sequences from the analysis using RepeatMasker (http://www.repeatmasker.org/). PCR conditions and primer sequences are freely available from the JSNP website (http://www.snp.ims.u-tokyo.ac.jp). For SNP discovery, we performed each PCR using a mixture of genomic DNAs from three individuals; 16 PCR products (corresponding to 48 individuals) amplified from the PCR experiments served as templates for direct sequencing. All SNPs detected by the Polyphred computer program were confirmed by sequencing both strands of each PCR product. Since the genotype of each individual cannot be obtained by this method, we subsequently performed genotyping of individual samples using another 96 Japanese samples. Calculation of Hardy–Weinberg equilibrium linkage disequilibrium (LD) coefficient (D') was as described (Yamada et al. 2001). Haplotype frequencies were estimated using SNPAlyze software (DYNACOM, Chiba, Japan).

Table 1 Characterization of 21 genetic variations in the asporin gene (*ASPN*) locus. Both 5′ and 3′ flanking sequences to each variation are denoted by lower case letters. Variant nucleotides are indicated by capital letters. *SNP* Single nucleotide polymorphism

ID	Location	Exon	Position[a]	5′ Flanking sequences	Variation	3′ Flanking sequences	dbSNP
i-ASPN-1	5′ Flanking region		−3,800	aactcatgaaatcttaatca	C/T	ttatcccttcacagaagctt	rs1980848
i-ASPN-2	5′ Flanking region		−3,549	gaatgaatggattggtggac	G/C	gcctttcttgtagctactac	
i-ASPN-3	5′ Flanking region		−3,076	aaagggtgtgagaaatcata	C/T	gagagcttagattaggcaat	
i-ASPN-4	5′ Flanking region		−3,039	caatttatttttattttcac	G/A	atggttagcatatataaggc	
i-ASPN-5	5′ Flanking region		−2,517	ttaccacattatctgcagaa	C/T	agagtgaataagccaaatat	
i-ASPN-6	5′ Flanking region		−2,451	tccagggcctatttctttcc	C/T	taggccctctgaccaaagat	
i-ASPN-7	5′ Flanking region		−702	gtcttagttgatgtagaatt	G/A	ggcagtctcttggggatgag	rs7860786
i-ASPN-8	Intron 1		7,183	tagaattactatctgaaaaa	G/A	agatatttttagagttccca	
i-ASPN-9	Intron 1		7,353	tccatttaaatttaatatcc	C/A	atgtctgtttagtcctagac	rs3739606
i-ASPN-10	Intron 3		137	aaaatccctgcattcaacag	C/T	gttctagtcaattaaaaatt	
i-ASPN-11	Intron 3		3,904	tggcacaacattaattatta	T/C	agctctgcacttttctaact	rs13301537
i-ASPN-12	Intron 3		4,071	atcaaatccacttcctaata	T/C	tgtgtaactttctatgcttg	
i-ASPN-13	Intron 5		3,971	ggtggctcacgcctgtaatc	T/C	cagcacttt(A/G)ggaggccgag	rs2181712
i-ASPN-14	Intron 5		3,981	gcctgtaatc(T/C)cagcacttt	A/G	ggaggccgaggggggtggat	rs7860774
i-ASPN-15	Intron 6		144	tagaaagaaccacaattggc	C/T	gggtgtggtggctcacgcct	
i-ASPN-16	3′ Untranslated region	8	1,394	aaatgatcttacataaatct	C/G	atgcttgaccattcctttct	rs17591776
i-ASPN-17	3′ Untranslated region	8	1,465	taacactttgttatcaagca	C/T	attttaaaaagaactgtact	rs3174352
i-ASPN-18	3′ Untranslated region	8	1,884	taactcgcattttaatgatc	C/A	gctattataagcttttaata	rs8067
i-ASPN-19	3′ Untranslated region	8	2,114	caaattaaacagacagaaac	C/T	gaaagctctatataaatgct	rs17519719
i-ASPN-20	5′ Flanking region		−572	atattctcttgaaaagggaa	A/del	gttgtatgcatcctaagagg	
i-ASPN-21	Intron 1		6,953	tttttatatttgcttttttt	T/del	atttcacaaaagataatatc	

[a]Nucleotide numbering is according to the mutation nomenclature (den Dunnen and Antonarakis 2000)

Table 2 Genotype and minor-allele frequencies of 19 SNPs in the *ASPN* locus (96 chromosomes were genotyped). *MAF* Minor allele frequency, *HWE* Hardy–Weinberg equilibrium

ID	Genotype distribution			MAF	HWE test (*P*-value)
	11	12	22		
i-ASPN-1	69	23	4	0.16	0.26
i-ASPN-2	92	4	0	0.02	0.84
i-ASPN-3	94	2	0	0.01	0.92
i-ASPN-4	94	2	0	0.01	0.92
i-ASPN-5	69	23	4	0.16	0.26
i-ASPN-6	93	3	0	0.02	0.88
i-ASPN-7	62	27	7	0.21	0.11
i-ASPN-8	94	2	0	0.01	0.92
i-ASPN-9	63	30	3	0.19	0.8
i-ASPN-10	93	3	0	0.02	0.88
i-ASPN-11	69	23	4	0.16	0.21
i-ASPN-12	96	0	0	0	–
i-ASPN-13	58	34	4	0.22	0.72
i-ASPN-14	63	30	3	0.19	0.8
i-ASPN-15	94	2	0	0.01	0.92
i-ASPN-16	93	3	0	0.02	0.88
i-ASPN-17	56	32	8	0.25	0.28
i-ASPN-18	62	28	6	0.2	0.26
i-ASPN-19	93	3	0	0.02	0.92

Results and discussion

We sequenced a total of 12-kb of genomic DNA from the *ASPN* region. We identified 19 SNPs, which are distributed an average of 632 nucleotide apart (Fig. 1, Table 1). We found seven SNPs in the 5′-flanking region, eight in introns, and four in the 3′-untranslated region. After comparing our data with the NCBI dbSNP database (late August 2005), we found 9 of 19 SNPs to be novel. The overall frequencies of nucleotide substitutions were 11% for A/C, 21% for A/G, 11% for C/G, and 58% for C/T. Transitions occurred 3.8 times more frequently than transversions. The substitution rate deviated from 2:1 transition-to-transversion theory and previous results in a gene-based SNP study (Haga et al. 2002). In addition, we identified two insertion–deletion polymorphisms in the 5′-flanking region and intron 1.

To examine the frequencies of the SNPs discovered in *ASPN*, we genotyped another 96 Japanese samples. The distributions of genotypes and minor-allele frequencies for the 19 SNPs are summarized in Table 2. All genotype distributions were in Hardy–Weinberg equilibrium ($P > 0.05$). The SNP i-ASPN-12 (4071T > C) was not identified in the cohort, probably because of the rarity of this SNP. In earlier report, we used seven genetic variations plus a D-repeat polymorphism to perform an association study of knee OA, and found that all seven variations were in strong LD with each other (Kizawa et al. 2005). Using the same seven genetic variations, we constructed haplotypes around *ASPN* and found that five haplotypes covered more than 90% of those in the population (Table 3). These haplotypes can be represented by three SNPs: i-ASPN-5, i-ASPN-21, and i-ASPN-17. This information on haplotype structure and "tag" SNPs will be a useful resource with which to examine *ASPN* associations.

Asporin is an extracellular matrix (ECM) protein that works as a principal negative regulator of TGF-β signaling in chondrocytes (Kizawa et al. 2005). TGF-β is a cytokine playing critical roles in the regulation of proliferation and differentiation of various types of cells and tissues. Disruption of the functional relationship between TGF-β activity and ECM proteins is known to contribute to several disease phenotypes (Border et al. 1992; Isaka et al. 1996; Neptune et al. 2003). Here, we report a high-resolution SNP map of the 33.4-kb region containing *ASPN*, showing locations of 21 sequence variations found among a Japanese cohort. The sequence variations identified here should provide a useful resource, not only for further examination of the rela-

Table 3 Haplotype structures of *ASPN* in Japanese samples

Haplotype ID	Frequency	SNP						
		i-ASPN-1	i-ASPN-5	i-ASPN-7	i-ASPN-21	i-ASPN-11	i-ASPN-17	i-ASPN-18
Haplotype 1	0.7	C	T	A	T	T	C	C
Haplotype 2	0.07	T	C	G	T	C	T	A
Haplotype 3	0.05	T	C	G	T/del	C	T	A
Haplotype 4	0.05	C	C	A	T	C	C	A
Haplotype 5	0.04	C	T	A	T	T	T	C

tionships between genotypes and susceptibility to OA, but also for screening of genes associated with other complex diseases involved in TGF-β signaling cascades, in particular bone and joint diseases.

References

Border WA, Noble NA, Yamamoto T, Harper JR, Yamaguchi Y, Pierschbacher MD, Ruoslahti E (1992) Natural inhibitor of transforming growth factor-beta protects against scarring in experimental kidney disease. Nature 360:361–364

Den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mut 15:7–12

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 variations in the human genome. J Hum Genet 47:605–610

Henry SP, Takanosu M, Boyd TC, Mayne PM, Eberspaecher H, Zhou W, de Crombrugghe B, Hook M, Mayne R (2001) Expression pattern and gene characterization of asporin. A newly discovered member of the leucine-rich repeat protein family. J Biol Chem 276:12212–12221

Iida A, Ozaki K, Tanaka T, Nakamura Y (2005) Fine-scale SNP map of an 11-kb genomic region at 22q13.1 containing the galectin-1 gene. J Hum Genet 50:42–45

Isaka Y, Brees DK, Ikegaya K, Kaneda Y, Imai E, Noble NA, Border WA (1996) Gene therapy by skeletal muscle expression of decorin prevents fibrotic disease in rat kidney. Nat Med 2:418–423

Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, Fukuda A, Mabuchi A, Kotani A, Kawakami A, Yamamoto S, Uchida A, Nakamura K, Notoya K, Nakamura Y, Ikegawa S (2005) An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. Nat Genet 37:138–144

Lorenzo P, Aspberg A, Onnerfjord P, Bayliss MT, Neame PJ, Heinegard D (2001) Identification and characterization of asporin. a novel member of the leucine-rich repeat protein family closely related to decorin and biglycan. J Biol Chem 276:12201–12211

Neptune ER, Frischmeyer PA, Arking DE, Myers L, Bunton TE, Gayraud B, Ramirez F, Sakai LY, Dietz HC (2003) Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. Nat Genet 33:407–411

Yamada R, Tanaka T, Unoki M, Nagai T, Sawasa 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