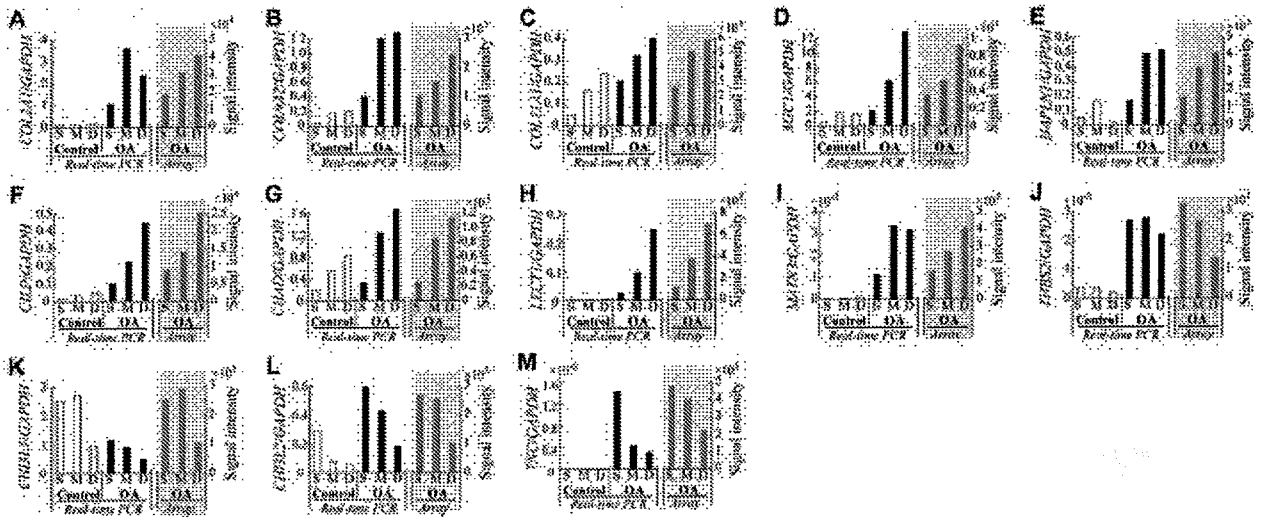

Figure 1. Comparison of cartilage matrix gene expression between microarray and real-time polymerase chain reaction (PCR) analyses. Expression of cartilage matrix genes was determined in superficial (S), middle (M), and deep (D) zones of osteoarthritic (OA) and control cartilage using laser capture microdissection and real-time PCR, and the result is shown together with that of the microarray analysis (Array). At least 8 samples were used for real-time PCR analysis. Results of real-time PCR analyses are shown as the ratio of the expression of the gene to that of *GAPDH*. Results of microarray analyses are shown as signal intensities. Expression of the following genes is shown: *COL2A1* (collagen, type II, α1) (A), *COL9A2* (collagen, type IX, α2) (B), *COL11A1* (collagen, type XI, α1) (C), *AGC1* (aggrecan 1) (D), *HAPLN1* (hyaluronan and proteoglycan link protein 1) (E), *CILP* (cartilage intermediate-layer protein, nucleotide pyrophosphohydrolase) (F), *CHAD* (chondroadherin) (G), *LECT1* (leukocyte cell–derived chemotaxin 1) (H), *MATN3* (matrilin 3) (I), *THBS2* (thrombospondin 2) (J), *CHI3L1* (chitinase 3–like 1 [cartilage glycoprotein 39]) (K), *CHI3L2* (chitinase 3–like 2) (L), *TNC* (tenascin C) (M).

seconds (see Supplementary Table 2, available in the online version of this article at http://www3.interscience.wiley.com/journal/76509746/home), and 72°C for 30 seconds. The amount of specific cDNA was quantified with a standard curve based on the known amounts of PCR product. The levels of cDNA among samples were normalized to the expression of GAPDH.

Statistical analysis. Statistically significant differences in gene expression between the zones were determined by 2-tailed *t*-test, and correlation of expression was evaluated by linear regression analysis. *P* values less than 0.05 were considered significant.

RESULTS

Microarray analysis findings. Significant levels of signals were obtained with all 538 probe sets (527 chondrocyte-related genes and 11 housekeeping genes) contained in the microarray. Among the 11 housekeeping genes, the expression of 4 genes (actin, beta; beta-2-microglobulin; phosphoglycerate kinase 1; transferrin receptor) was significantly different among cartilage zones. Thus, the normalization was performed on the expression of the remaining 7 housekeeping genes (glyceraldehyde-3-phosphate dehydrogenase; glucuronidase, beta; hypoxanthine phosphoribosyltransferase 1; peptidylprolyl isomerase A; ribosomal protein, large, P0; TATA box binding protein; 18S ribosomal RNA) (see Supplementary Table 1B, available in the online version of this article at http://www3.interscience.wiley.com/journal/76509746/home).

The expression of 198 of 527 chondrocyte-related genes was significantly different between any 2 cartilage zones. Between the superficial and middle zones, the expression of 93 genes was significantly different (see Supplementary Table 3, available in the online version of this article at http://www3.interscience.wiley.com/journal/76509746/home). Among them, 9 genes were expressed ≥2 times in the superficial zone, while another 7 genes were expressed ≥2 times in the middle zone (Table 1).

A total of 115 genes were expressed at significantly different levels between the middle and deep zones (see Supplementary Table 4, available in the online version of this article at http://www3.interscience.wiley.com/journal/76509746/home). The expression of 12 of them was increased ≥2-fold in the middle zone, while that of another 9 genes was increased ≥2-fold in the deep zone (Table 2). Among these genes, the expression of matrix metalloproteinase 13 (*MMP13*) was most increased in the middle zone, while the expres-

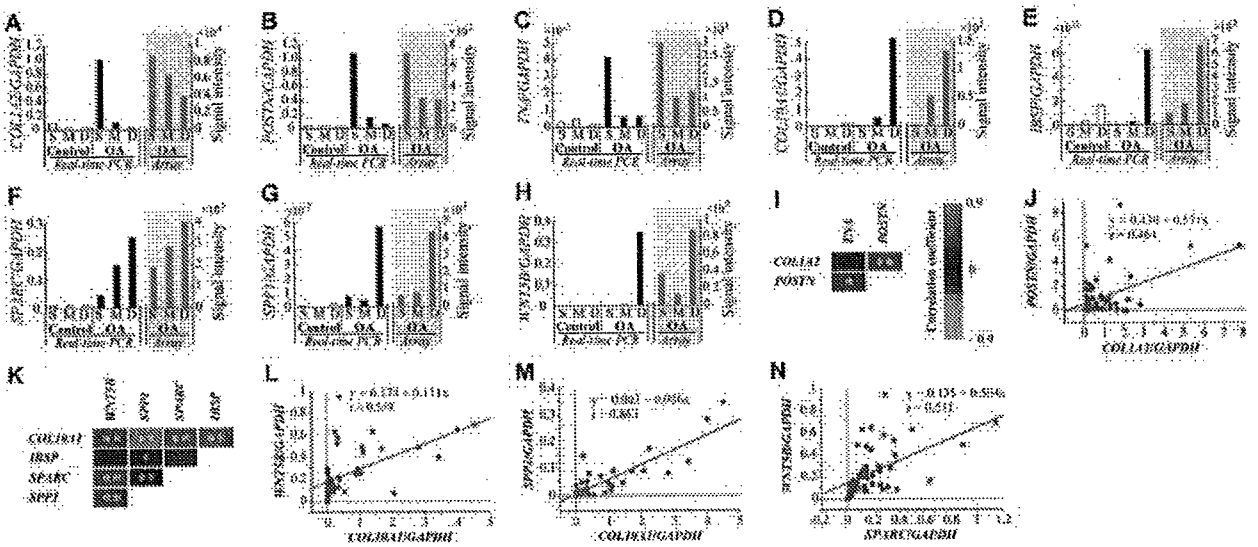

Figure 2. A–H, Comparison of bone-related gene expression between microarray and real-time PCR analyses. Expression of bone-related genes was determined in respective zones by laser capture microdissection and real-time PCR, and the result is shown together with that of the microarray analysis. Data are presented in the manner shown in Figure 1. At least 13 samples were used for real-time PCR analysis. I–N, Correlation of expression among the bone-related genes in OA cartilage. For genes primarily expressed in the superficial zone, correlation coefficients of expression are shown on a heat map (I), and the relationship of *POSTN* (periostin, osteoblast-specific factor) and *COL1A2* (collagen, type I, α2) expression is presented on a scattergram (J). For genes whose expression is most enhanced in the deep zone, correlation coefficients of expression are shown on a heat map (K), and relationships of expression between *COL10A1* (collagen, type X, α1) and *WNT5B* (wingless-type MMTV integration site family, member 5B, transcript variant 2), between *COL10A1* and *SPP1* (secreted phosphoprotein 1 [osteopontin, bone sialoprotein I]), and between *SPARC* (secreted protein, acidic, cysteine-rich) and *WNT5B* are shown on scattergrams (L–N, respectively). Results of 42 OA cartilage samples are shown. In I and K, red and green colors indicate positive and negative correlations, respectively. Expression of the following genes is also shown: *TNA* (tetranectin) (C), *IBSP* (integrin-binding sialoprotein [bone sialoprotein, bone sialoprotein II]) (E). * = $P < 0.05$; ** = $P < 0.01$. See Figure 1 for other definitions.

sion of wingless-type MMTV integration site family, member 5B, transcript variant 2 (*WNT5B*) was most enhanced in the deep zone.

The difference in gene expression was most obvious between the superficial and deep zones. The expression of 126 genes differed significantly between these zones (see Supplementary Table 5, available in the online version of this article at http://www3.interscience.wiley.com/journal/76509746/home). The expression of 25 of the 126 genes was increased ≥2 times in the superficial zone (Table 3). The expression of *MMP13* was most increased, followed by that of matrix metalloproteinase 2 (*MMP2*) and periostin, osteoblast-specific factor (*POSTN*). Meanwhile, the expression of another 23 genes was increased ≥2 times in the deep zone (Table 3). The expression of collagen, type X, α1 (*COL10A1*) was most increased (nearly 30 times), followed by that of leukocyte cell–derived chemotaxin 1 (*LECT1*), matrilin 3 (*MATN3*), integrin-binding sialoprotein (bone sialoprotein, bone sialoprotein II) (*IBSP*), and secreted phosphoprotein 1 (osteopontin, bone sialoprotein I) (*SPP1*).

Comparison of cartilage matrix gene expression between microarray analysis and real-time PCR. Real-time PCR analysis was performed on selected genes to validate the result of microarray analysis and to compare their expression in OA and control cartilage. Sixteen OA cartilage samples and 13 control cartilage samples were used for this analysis. Those cartilages were separated into 3 zones by laser capture microdissection for precise zone isolation based on their histologic features (2).

We chose a total of 26 genes for this analysis, considering the difference in expression levels among the zones (≥2-fold difference between any 2 zones), signal intensities, and relevance to OA pathology. These genes were categorized into 3 groups. The first group consisted of 13 genes encoding components of articular cartilage. The second group contained 8 genes that are expressed during hypertrophic change of the chondrocytes and those expressed by osteoblasts (bone-related

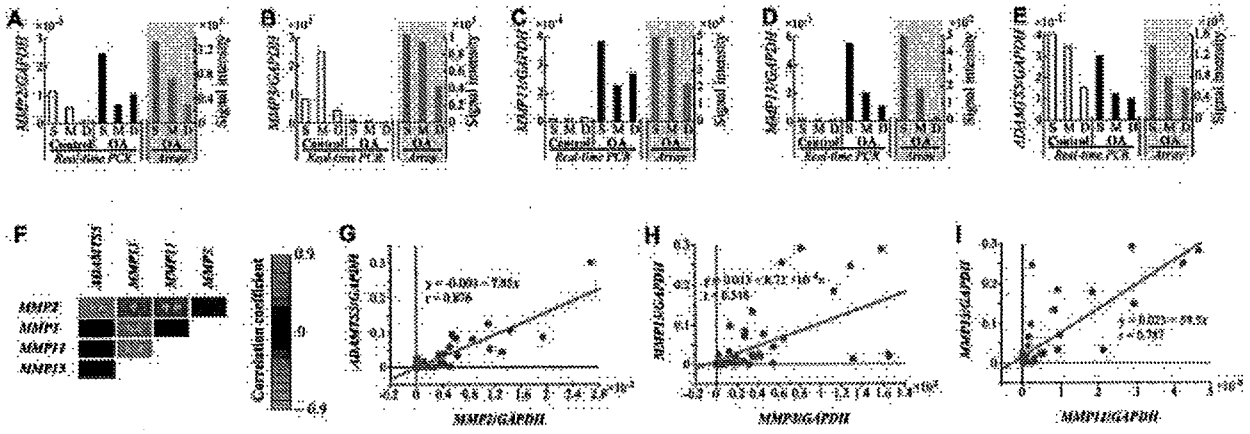

Figure 3. A–E, Comparison of metalloproteinase gene expression between microarray and real-time PCR analyses. Expression of bone-related genes was determined in respective zones by laser capture microdissection and real-time PCR, and the result is shown together with that of the microarray analysis. Data are presented in the manner shown in Figure 1. At least 13 samples were used for real-time PCR analysis. F–I, Correlation of expression of the proteinase genes in the superficial zone of OA cartilage. Correlation coefficients of expression among the genes are shown on a heat map (F), and relationships of expression between *MMP2* (matrix metalloproteinase 2) and *ADAMTS5* (a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 5 [aggrecanase 2]), between *MMP3* (matrix metalloproteinase 3) and *MMP13* (matrix metalloproteinase 13), and between *MMP11* (matrix metalloproteinase 11) and *MMP13* are shown on scattergrams (G–I, respectively). Results of 42 OA cartilage samples are shown. In F, red and green colors indicate positive and negative correlations, respectively. * = $P < 0.05$; ** = $P < 0.01$. See Figure 1 for other definitions.

genes). The genes encoding metalloproteinases that could promote cartilage degeneration were assigned to the third group. The results of the microarray analysis and real-time PCR were compared in those 3 groups, respectively.

For 13 cartilage matrix genes, the result of microarray analysis was generally well consistent with that of real-time PCR, confirming the validity of our microarray analysis (Figure 1). The following features were noticed with the expression of cartilage matrix genes in OA cartilage. First, except for chitinase 3–like 1 (*CHI3L1*) (Figure 1K), the expression of all cartilage matrix genes was enhanced in OA cartilage. Second, despite such increase in expression, the change of gene expression across the zones was similar between OA and control cartilage. In other words, within OA cartilage, the level of increase in expression was similar in all 3 zones. Third, however, the degree of increase in expression was considerably different among the genes. For example, while the expression of *LECT1* and tenascin C (*TNC*) was increased 40 times and 220 times, respectively, in OA cartilage (Figures 1H and M), the increase of chondroadherin (*CHAD*) expression was merely 2-fold (Figure 1G).

Expression of bone-related genes. For this group of genes, the results were also consistent between the microarray and real-time PCR analyses (Figures 2A–H). The genes in this group were further divided into 2 categories by the pattern of expression across the zones. The expression of collagen, type I, α2 (*COL1A2*), *POSTN*, and tetranectin (*TNA*) was most enhanced in the superficial zone of OA cartilage (Figures 2A–C). Meanwhile, the expression of *COL10A1*, *IBSP*, secreted protein, acidic, cysteine-rich (*SPARC*), *SPP1*, and *WNT5B* was highest in the deep zone (Figures 2D–H). Since the expression of those genes in the control cartilage was very low, their expression in the specific zones of OA cartilage indicated the phenotypic change of the chondrocytes at those sites.

With these genes, a large difference was observed in expression levels among OA cartilage samples. For example, the expression level of *COL10A1* in the deep zone differed more than 1×10^4–fold among samples. If the induction of such gene expression was in fact related to the phenotypic change of the cells, there should be some correlation in their expression. To examine this possibility, we evaluated the expression of those genes in the superficial or deep zone in a greater number of OA cartilage samples. Among the 3 genes expressed in the superficial zone, a positive correlation was observed between *COL1A2* and *POSTN* ($r = 0.464$, $P < 0.001$) and between *POSTN* and *TNA* ($r = 0.300$, $P = 0.036$) (Figures 2I and J). A significant correlation was also observed among the 5 genes expressed in the deep zone.

Except for correlations between *IBSP* and *SPARC* and between *IBSP* and *WNT5B*, the expression of those 5 genes was significantly correlated in any other possible combination (Figure 2K). In particular, a close correlation was found between *COL10A1* and *WNT5B* ($r = 0.559$, $P < 0.001$), between *COL10A1* and *SPP1* ($r = 0.863$, $P < 0.001$), and between *SPARC* and *WNT5B* ($r = 0.511$, $P < 0.001$) (Figures 2L–N). This result implies that the expression of those bone-related genes could be induced in association with the phenotypic shift of the chondrocytes in the superficial and deep zones, respectively.

Expression of metalloproteinase genes. We also performed real-time PCR analysis on the expression of 5 metalloproteinase genes. Again, for these genes, the result of real-time PCR was almost consistent with that of the microarray analysis (Figures 3A–E). Interestingly, the expression of 2 proteinase genes was either not altered (a disintegrin-like and metalloproteinase with thrombospondin type 1 motif, 5 [aggrecanase 2]; termed *ADAMTS5*) (Figure 3E) or rather reduced (matrix metalloproteinase 3 [*MMP3*]) (Figure 3B) in OA cartilage compared with control cartilage. The expression of the other 3 genes was increased in OA cartilage. In particular, the expression of matrix metalloproteinase 11 (*MMP11*) and *MMP13* was highly enhanced in OA cartilage and reached >40-fold that in control cartilage.

The expression of these 5 proteinase genes, including 2 genes whose expression was not enhanced in OA cartilage, was most enhanced in the superficial zone. We then investigated the relationship of their expression in that zone in an increased number of OA cartilage samples. The results indicated that the expression of all 5 proteinase genes was mutually correlated (Figure 3F). A close correlation was observed between *MMP2* and *ADAMTS5* ($r = 0.876$, $P < 0.001$), between *MMP3* and *MMP13* ($r = 0.548$, $P < 0.001$), and between *MMP11* and *MMP13* ($r = 0.787$, $P < 0.001$) (Figures 3G–I), while a significant correlation was also seen between *MMP2* and *MMP11* ($r = 0.373$, $P = 0.009$) and between *MMP2* and *MMP13* ($r = 0.329$, $P = 0.023$).

DISCUSSION

In this study, the result of microarray analysis was in good agreement with that of real-time PCR (Figures 1–3) (see Supplementary Figure 1, available in the online version of this article at http://www3.interscience.wiley.com/journal/76509746/home). This corroborates the validity of our current analyses. Although some discrepancies were observed, they might not have been related to possible technical problems but more likely stemmed from individual differences in expression levels, given that a large individual difference is an inevitable problem in the analysis of human cartilage samples (2,20,21).

Our microarray analysis showed that ~40% of the investigated genes were expressed at significantly different levels among the zones. This zonal variation in expression would reflect both the physiologic difference in cell metabolism among the zones and the changes caused by the disease. This was illustrated by the expression of cartilage matrix genes (Figure 1). Our current analyses indicated that more than a dozen cartilage matrix genes were expressed at significantly different levels among the zones within OA cartilage. Despite this difference in gene expression levels, a similar zonal change of expression was observed in the control cartilage. Thus, it was assumed that in OA cartilage, the expression of those genes was amplified (or reduced) equally across the 3 zones by a certain mechanism(s) involved in the disease. This finding could be a clue for elucidating the mechanism(s) for enhanced matrix synthesis in OA, which remains entirely unknown.

Current analysis also revealed that the magnitude of the increase in expression was considerably different among the cartilage matrix genes. This finding may imply an unrecognized but potentially important mechanism for the progression of OA. Our current and previous evaluations showed that the 3 constitutive collagens of cartilage matrix (types II, IX, and XI collagen) could be synthesized at an altered ratio by OA chondrocytes (22). While the expression of *COL2A1* in OA cartilage was increased 10–24-fold that in control cartilage, the expression of collagen, type IX, $\alpha 2$ gene (*COL9A2*) and collagen, type XI, $\alpha 1$ gene (*COL11A1*) was enhanced 7–11-fold and 2–4-fold, respectively. In accordance with this, our analysis of cartilage collagens revealed that the amounts of pepsin-extractable type IX and type XI collagens are in fact reduced in OA cartilage relative to that of type II collagen (Fukui N: unpublished observations).

Type IX and type XI collagens form a collagen fibril network within cartilage, together with type II collagen. As known from observations of human hereditary diseases and gene-manipulated mice, the proper level of expression of these minor collagens is indispensable to maintain the normal properties of cartilage matrix (23–27). Therefore, the relative reduction of *COL9A2* and *COL11A1* expression in OA cartilage may lead to the fragility of newly synthesized cartilage, which might rather facilitate loss of cartilage matrix. Although

not yet demonstrated, disproportionate expression of other matrix components could have a similar significance in disease progression (28). We now assume that such imbalance in cartilage matrix gene expression could be significantly involved in the pathology of OA.

The expression of a series of bone-related genes in OA cartilage is another novel finding of this study. While the induction of *COL1A2* and *COL10A1* expression in OA cartilage has been known for more than a decade (2–4,21,29–36), enhanced expression of *TNA*, *POSTN*, *IBSP*, and *WNT5B* in human OA cartilage has not been reported previously. Our results suggested that the expression of these genes in OA cartilage could be related to the phenotypic change of the chondrocytes. The fact that their expression was induced either in the superficial zone or in the deep zone, and not in both, may reflect the occurrence of distinctive phenotypic changes in those zones.

Three among these bone-related genes expressed in the superficial zone (*COL1A2*, *POSTN*, and *TNA*) are known to be expressed by the osteoblasts (37,38). Meanwhile, 2 genes expressed in the deep zone (*COL10A1* and *WNT5B*) are expressed characteristically in the chondrocytes undergoing hypertrophic change (39). Thus, the phenotypic change in the superficial zone might have an aspect of osteoblastic differentiation, whereas that in the deep zone could have a trait of chondrocyte hypertrophy. In light of the developmental process of articular cartilage, this notion might not be unreasonable. At present, it is not known why these genes are expressed in OA cartilage. In the future, a more comprehensive analysis of gene expression profiles in the respective zones may clarify the molecular mechanism(s) involved in the phenotypic changes.

Meanwhile, the expression of 5 proteinase genes was most enhanced in the superficial zone of OA cartilage. In OA, the cartilage surface is the region where cartilage degeneration begins (10–12). Therefore, those proteinases could be responsible for the initiation of matrix degeneration at that site. Furthermore, the finding that their expression was mutually correlated at the superficial zone implies that those proteinases could work synergistically in certain cases to cause cartilage degeneration. That correlation also suggests the presence of a common regulatory mechanism(s) for their expression. Regulation of proteinase expression in OA cartilage is critically important to inhibit the progression of the disease. Thus, elucidation of such a mechanism(s) may be useful to develop new therapeutic strategies for OA.

Although it is important, the change of chondrocyte metabolism in OA has been understood only partly. This incomplete understanding could be ascribed, at least in part, to the regional difference in cellular metabolism within cartilage. As demonstrated in this study, consideration of the regional difference within cartilage could provide further insights into the metabolic change of the chondrocytes with the disease. Conventionally, such regional differences have been studied by histologic evaluations of a limited number of genes. Compared with those techniques, the experimental methods we employed here could be advantageous in that they allow a more comprehensive evaluation of the cellular metabolism in specific sites. Obviously, this study has several limitations. The number of samples used for the microarray analysis was rather small, and the number of genes contained in the array was limited to 527. Although confirmed for selected genes, the reliability of the microarray analysis has not been fully validated. Therefore, it is very likely that some genes are left unnoticed while they are expressed at different intensities across the zones. For all its limitations, the present study promises to provide greater understanding of the pathology of OA.

AUTHOR CONTRIBUTIONS

Drs. Fukui and Ikegawa had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study design. Fukui, Ikegawa.

Acquisition of data. Miyamoto, Nakajima, Ikeda, Hikita, Furukawa, Mitomi, Tanaka, Katsuragawa, Yamamoto, Sawabe, Juji, Mori, Suzuki.

Analysis and interpretation of data. Fukui, Miyamoto, Nakajima.

Manuscript preparation. Fukui, Ikegawa.

Statistical analysis. Fukui, Miyamoto, Nakajima.

REFERENCES

1. Pelletier JP, Martel-Pelletier J, Abramson SB. Osteoarthritis, an inflammatory disease: potential implication for the selection of new therapeutic targets. Arthritis Rheum 2001;44:1237–47.
2. Fukui N, Ikeda Y, Ohnuki T, Tanaka N, Hikita A, Mitomi H, et al. Regional differences in chondrocyte metabolism in osteoarthritis: a detailed analysis by laser capture microdissection. Arthritis Rheum 2008;58:154–63.
3. Von der Mark K, Kirsch T, Nerlich A, Kuss A, Weseloh G, Gluckert K, et al. Type X collagen synthesis in human osteoarthritic cartilage: indication of chondrocyte hypertrophy. Arthritis Rheum 1992;35:806–11.
4. Aigner T, Bertling W, Stoss H, Weseloh G, von der Mark K. Independent expression of fibril-forming collagens I, II, and III in chondrocytes of human osteoarthritic cartilage. J Clin Invest 1993;91:829–37.
5. Young RD, Lawrence PA, Duance VC, Aigner T, Monaghan P. Immunolocalization of collagen types II and III in single fibrils of human articular cartilage. J Histochem Cytochem 2000;48:423–32.
6. Buckwalter JA, Mankin HJ, Grodzinsky AJ. Articular cartilage

and osteoarthritis. In: Pellegrini VD Jr, editor. AAOS instructional course lectures. Vol 54. Rosemont (IL): American Academy of Orthopaedic Surgeons; 2005. p. 465–80.
7. Poole RA. Cartilage in health and disease. In: Koopman WJ, Moreland LW, editors. Arthritis and allied conditions: a textbook of rheumatology. 15th ed. Philadelphia: Lippincott Williams & Wilkins; 2004. p. 223–69.
8. Sandell LJ, Heinegard D, Hering TM. Cell biology, biochemistry, and molecular biology of articular cartilage in osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, editors. Osteoarthritis: diagnosis and medical/surgical management. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 73–106.
9. Mow VC, Hung CT. Mechanical properties of normal and osteoarthritic articular cartilage, and the mechanobiology of chondrocytes. New York: Oxford University Press; 2003.
10. Hollander AP, Pidoux I, Reiner A, Rorabeck C, Bourne R, Poole AR. Damage to type II collagen in aging and osteoarthritis starts at the articular surface, originates around chondrocytes, and extends into the cartilage with progressive degeneration. J Clin Invest 1995;96:2859–69.
11. Lark MW, Bayne EK, Flanagan J, Harper CF, Hoerrner LA, Hutchinson NI, et al. Aggrecan degradation in human cartilage: evidence for both matrix metalloproteinase and aggrecanase activity in normal, osteoarthritic, and rheumatoid joints. J Clin Invest 1997;100:93–106.
12. Poole RA, Guilak F, Abramson SB. Etiopathogenesis of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, editors. Osteoarthritis: diagnosis and medical/surgical management. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 27–49.
13. Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, et al. Laser capture microdissection. Science 1996;274:998–1001.
14. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039–49.
15. Baelde HJ, Cleton-Jansen AM, van Beerendonk H, Namba M, Bovee JV, Hogendoorn PC. High quality RNA isolation from tumours with low cellularity and high extracellular matrix component for cDNA microarrays: application to chondrosarcoma. J Clin Pathol 2001;54:778–82.
16. Zhang H, Liew CC, Marshall KW. Microarray analysis reveals the involvement of β-2 microglobulin (B2M) in human osteoarthritis. Osteoarthritis Cartilage 2002;10:950–60.
17. Zhang H, Marshall KW, Tang H, Hwang DM, Lee M, Liew CC. Profiling genes expressed in human fetal cartilage using 13,155 expressed sequence tags. Osteoarthritis Cartilage 2003;11:309–19.
18. Matsumura H, Bin Nasir KH, Yoshida K, Ito A, Kahl G, Kruger DH, et al. SuperSAGE array: the direct use of 26-base-pair transcript tags in oligonucleotide arrays. Nat Methods 2006;3: 469–74.
19. Nuwaysir EF, Huang W, Albert TJ, Singh J, Nuwaysir K, Pitas A, et al. Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. Genome Res 2002;12: 1749–55.
20. Yagi R, McBurney D, Laverty D, Weiner S, Horton WE Jr. Intrajoint comparisons of gene expression patterns in human osteoarthritis suggest a change in chondrocyte phenotype. J Orthop Res 2005;23:1128–38.
21. Aigner T, Fundel K, Saas J, Gebhard PM, Haag J, Weiss T, et al. Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. Arthritis Rheum 2006;54:3533–44.
22. Fukui N, Ikeda Y, Ohnuki T, Yamane S, Suzuki R, Yamamoto S, et al. Integrin $\alpha5\beta1$ may be responsible for the induction of type I and type III collagen during dedifferentiation in the monolayer cultured adult human articular chondrocytes. Osteoarthritis Cartilage 2005;13 Suppl A:S20–1.
23. Briggs MD, Chapman KL. Pseudoachondroplasia and multiple epiphyseal dysplasia: mutation review, molecular interactions, and genotype to phenotype correlations. Hum Mutat 2002;19:465–78.
24. Fassler R, Schnegelsberg PN, Dausman J, Shinya T, Muragaki Y, McCarthy MT, et al. Mice lacking $\alpha1$(IX) collagen develop noninflammatory degenerative joint disease. Proc Natl Acad Sci U S A 1994;91:5070–4.
25. Hu K, Xu L, Cao L, Flahiff CM, Brussiau J, Ho K, et al. Pathogenesis of osteoarthritis-like changes in the joints of mice deficient in type IX collagen. Arthritis Rheum 2006;54:2891–900.
26. Spranger J. The type XI collagenopathies. Pediatr Radiol 1998; 28:745–50.
27. Xu L, Flahiff CM, Waldman BA, Wu D, Olsen BR, Setton LA, et al. Osteoarthritis-like changes and decreased mechanical function of articular cartilage in the joints of mice with the chondrodysplasia gene (cho). Arthritis Rheum 2003;48:2509–18.
28. Cs-Szabo G, Melching LI, Roughley PJ, Glant TT. Changes in messenger RNA and protein levels of proteoglycans and link protein in human osteoarthritic cartilage samples. Arthritis Rheum 1997;40:1037–45.
29. Girkontaite I, Frischholz S, Lammi P, Wagner K, Swoboda B, Aigner T, et al. Immunolocalization of type X collagen in normal fetal and adult osteoarthritic cartilage with monoclonal antibodies. Matrix Biol 1996;15:231–8.
30. Boos N, Nerlich AG, Wiest I, von der Mark K, Ganz R, Aebi M. Immunohistochemical analysis of type-X-collagen expression in osteoarthritis of the hip joint. J Orthop Res 1999;17:495–502.
31. Adam M, Deyl Z. Altered expression of collagen phenotype in osteoarthrosis. Clin Chim Acta 1983;133:25–32.
32. Aigner T, Gluckert K, von der Mark K. Activation of fibrillar collagen synthesis and phenotypic modulation of chondrocytes in early human osteoarthritic cartilage lesions. Osteoarthritis Cartilage 1997;5:183–9.
33. Aigner T, Reichenberger E, Bertling W, Kirsch T, Stoss H, von der Mark K. Type X collagen expression in osteoarthritic and rheumatoid articular cartilage. Virchows Arch B Cell Pathol Incl Mol Pathol 1993;63:205–11.
34. Aigner T, Zien A, Gehrsitz A, Gebhard PM, McKenna L. Anabolic and catabolic gene expression pattern analysis in normal versus osteoarthritic cartilage using complementary DNA–array technology. Arthritis Rheum 2001;44:2777–89.
35. Goldwasser M, Astley T, van der Rest M, Glorieux FH. Analysis of the type of collagen present in osteoarthritic human cartilage. Clin Orthop Relat Res 1982;(167):296–302.
36. Miosge N, Hartmann M, Maelicke C, Herken R. Expression of collagen type I and type II in consecutive stages of human osteoarthritis. Histochem Cell Biol 2004;122:229–36.
37. Wewer UM, Ibaraki K, Schjorring P, Durkin ME, Young MF, Albrechtsen R. A potential role for tetranectin in mineralization during osteogenesis. J Cell Biol 1994;127:1767–75.
38. Horiuchi K, Amizuka N, Takeshita S, Takamatsu H, Katsuura M, Ozawa H, et al. Identification and characterization of a novel protein, periostin, with restricted expression to periosteum and periodontal ligament and increased expression by transforming growth factor β. J Bone Miner Res 1999;14:1239–49.
39. Church V, Nohno T, Linker C, Marcelle C, Francis-West P. Wnt regulation of chondrocyte differentiation. J Cell Sci 2002;115: 4809–18.

 doi: 10.1111/j.1442-2042.2009.02263.x

Original Article

Polymorphism of cytochrome P450 2B6 and prostate cancer risk: A significant association in a Japanese population

Takayuki Kurosaki,[1] Motofumi Suzuki,[1] Yutaka Enomoto,[1] Tomio Arai,[2] Motoji Sawabe,[2] Takayuki Hosoi,[3] Yukio Homma[1] and Tadaichi Kitamura[1,4]

[1]Department of Urology, Graduate School of Medicine, The University of Tokyo, Tokyo, [2]Department of Pathology, Tokyo Metropolitan Geriatric Hospital, Tokyo, [4]Department of Urology, Asoka Hospital, Tokyo, and [3]Department of Advanced Medicine, National Center for Geriatrics and Gerontology, Ohbu, Aichi, Japan

Objectives: To explore whether Lys262Arg polymorphism of the Cytochrome P450 2B6 (CYP2B6) gene could act as a genetic marker for prostate cancer risk among Japanese men.

Methods: A total of 350 patients with sporadic prostate cancer and 328 controls were examined. A single nucleotide polymorphism with non-synonymous amino acid change located at Lys262Arg of the CYP2B6 gene was genotyped using a TaqMan assay.

Results: The frequency of the Arg/Arg genotype among prostate cancer patients was significantly higher than that among the controls ($P = 0.027$). The frequency of the G allele of the Lys262Arg polymorphism was also significantly higher in prostate cancer patients than in the controls (30.4% *vs* 24.8%, $P = 0.025$). Patients with the Lys/Arg plus Arg/Arg genotypes carried a low Gleason score more frequently than those with the Lys/Lys genotype ($P = 0.042$). The frequency of the G allele of the Lys262Arg polymorphism was significantly higher in the low Gleason score group than that in the high Gleason score group (34.3% *vs* 26.8%, $P = 0.038$).

Conclusions: Lys262Arg polymorphism of the CYP2B6 gene may be a genetic marker for evaluating the risk of sporadic prostate cancer in native Japanese men.

Key words: CYP2B6, prostate cancer, single nucleotide polymorphism, TaqMan assay, testosterone.

Introduction

In Japan, the morbidity rate for prostate cancer is in sixth place and the mortality rate is in ninth place among all cancers in men.[1] Both rates have been increasing constantly, and prostate cancer is drawing attention as a public health problem even in Japan. Some evidence suggests that androgens influence the development and progression of prostate cancer.[2] Abundant biological data have suggested that androgens play an important role in the development of prostate cancer. For instance, the growth and maintenance of the prostate are dependent on androgens, prostate cancer regresses after androgen ablation or anti-androgen therapy, and the administration of testosterone induces prostate carcinoma in laboratory animals.[3] Consequently, the most important treatment method for patients with advanced prostate cancer is androgen deprivation therapy (ADT), yet the efficacy of ADT is usually attenuated in patients with advanced prostate cancer. Androgen and androgen receptor (AR) appear to play an important role in such situations.[4,5]

The cytochrome P450 enzymes (CYP) play an important role in the metabolism of a wide variety of endogenous and foreign compounds, including steroid hormones.[6] The CYP2B6 gene is located in the chromosome 19q13.2, and CYP2B6 is involved in the metabolism of several therapeutically significant drugs and environmental toxicants. In addition, CYP2B6 hydroxylates testosterone in the liver, deactivating testosterone's hormonal function, and also inactivates dihydrotestosterone (DHT) in the prostate.[7–12] CYP2B6 is mainly expressed in liver, but it is also detected at lower levels in extrahepatic tissues, including the small intestine, kidney, lung, skin, brain, and prostate.[13]

The role of the genetic variants of *CYP2B6* in drug metabolism remains largely unknown. However, CYP2B6*5 (Arg487Cys), CYP2B6*6 (Lys262Arg and Gln172His), and CYP2B6*7 (Lys262Arg and Gln172His and Arg487Cys) variants are thought to reduce protein expression levels of the CYP2B6 in human liver.[14–16] If the decline in protein expression levels of the CYP2B6 caused by the genetic variants of *CYP2B6* were to hinder the hydroxylation of testosterone, this process could affect the development and progression of prostate cancer. Therefore, we herein focused on the relations among the genetic polymorphisms of the CYP2B6 gene, prostate cancer susceptibility, and clinical characteristics (Gleason sum and clinical T stage) in a Japanese population.

Methods

Study design

We examined 350 native Japanese patients with sporadic prostate cancer; the patients were between the ages of 48 and 89 years (mean ± standard deviation [SD], 71.0 ± 7.9 years) and were recruited at the University of Tokyo Hospital and its affiliated hospitals between January 1999 and April 2007. Patients with a family history of prostate cancer were carefully excluded from this study. One hundred and thirty two patients were diagnosed pathologically by total prostatectomy, and 218 patients were diagnosed by prostate needle biopsy. The tumor status of the patients was evaluated using a digital rectal examination, transrectal ultrasonography and pelvic computed tomography and was classified according to the 2002 Tumor, Node, Metastasis

Correspondence: Senior Assistant Professor Motofumi Suzuki MD PhD, Department of Urology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. Email: suzukim-uro@h.u-tokyo.ac.jp

Disclosure Statement: This study was supported by a grant from the Japanese Foundation for Prostate Research. The sponsor of the study had no role in the study design, data collection, data analysis, data interpretation, writing of the report, or in the decision to submit the manuscript for publication.

Received 20 June 2008; accepted 14 January 2009.

(TNM) staging system for cancer.[17] One hundred and forty seven patients were classified as localized stage (T1 and T2), and 203 patients were classified as advanced stage (T3 and T4). The pathological grade was divided into two groups: low grade (Gleason sum 2 to 7A, 169 patients) and high grade (Gleason sum 7B to 10, 181 patients). The study was conducted with the approval of the Ethics Committee of the University of Tokyo after obtaining written informed consent from each patient prior to entry in the study.

As a control group, we also examined 328 residence-matched Japanese men who had died at Tokyo Metropolitan Geriatric Hospital. All of these men were consecutively autopsied and pathologically confirmed to have no malignancies. The causes of death among the control subjects were described previously by Suzuki *et al.*[18] Their ages ranged from 49 to 100 years (mean ± SD, 79.2 ± 9.2 years). All of these patients had been registered in a database of Japanese single nucleotide polymorphisms for geriatric research (JG-SNP).[19] Written informed consent was obtained from the family members under the Act of Postmortem Examination. The study was also reviewed and approved by the Ethics Committee of Tokyo Metropolitan Geriatric Hospital.

Genotyping assay

Genomic DNA was extracted from the peripheral blood lymphocytes in the treated patients and from autopsied frozen kidney samples in the controls. The Lys262Arg (A-to-G transition) polymorphism was genotyped using a TaqMan assay method. The basis of the TaqMan assay was an allele-specific oligonucleotide probe labeled with a fluorescent reporter dye and a quencher dye, which was cleaved during the polymerase chain reaction (PCR) amplification process by the 5′ nuclease activity of TaqDNA polymerase and generated an increase in the intensity of the fluorescence in relation to the accumulation of the PCR product; this PCR product was measured directly in the reaction well.[20] A TaqMan SNP Genotyping Assays mix (Applied Biosystems, Foster City, CA, USA) was used. The primers used for the assay were a forward 5′-TGGAGAAGCACCGTGAAACC-3′ and a reverse 5′-TGGAGC AGGTAGGTGTCGAT-3′. Real-time PCR was carried out in a total volume of 25 mL containing 50 ng of genomic DNA, 12.5 mL 2 × TaqMan universal PCR master mix and 0.3 mL 20 × TaqMan SNP Genotyping Assays mix. The amplification reaction was carried out with one cycle at 50°C for 2 min, onecycle at 95°C for 10 min, and 49 cycles at 92°C for 30 s and 60°C for 1 min. Sequence detection software was used to assign the generated fluorescence signals of the samples automatically using a highly accurate statistical method. The fluorescence of each sample was measured and then analyzed after thermal cycling using an ABI PRISM 7000 or 7300 sequence detection system (Applied Biosystems), according to the manufacturer's instructions. In addition, we carried out a PCR-based direct sequencing method to validate our results. For use in the PCR solution, the DNA concentration was adjusted to 100 ng/mL and stored at –20°C. The primers used for DNA amplification were as follows: forward, 5′-AAGGATGAGGGAGGAAGA TGC-3′; reverse, 5′-GGCTGAATTCACTGTGTGCG-3′. The cyclic thermal conditions were 95°C for 10 min for one cycle; 95°C for 30 s, 60°C for 30 s, and 72°C for 3 min for 37 cycles; followed by an elongation cycle of 72°C for 10 min. The PCR reactions were carried out in a total 25-mL reaction volume containing 5 mL of 10×PCR Gold Buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP (Applied Biosystems), 0.5 mM of each specific primer (synthesized by Fasmac, Atsugi, Kanagawa, Japan), 1.25 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, Branchburg, NJ, USA), and 100 ng of genomic DNA. GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) was used to carry out PCR. All PCR products were purified with the Montage PCRm96 Plate (Millipore Corporation, Bedford, MA, USA) to remove deoxynucleotide triphosphates and excess primers. All sequencing reactions were carried out using dye terminator chemistry (ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Kit; Applied Biosystems, Warrington, UK) with each sequencing primer, and the products were purified using the MultiScreen filter plates (Millipore Corporation, Bedford, MA, USA) with Sephadex G-50 Superfine (Amersham Biosciences, Uppsala, Sweden) to remove contaminating salts and unincorporated dye terminators. Purified samples were applied to an ABI Prism 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA), and single nucleotide polymorphisms (SNPs) were identified using Sequencher software (version 4.1.2; Gene Codes Corporation, Ann Arbor, MI, USA). The results of the two methods were completely consistent.

Statistical analysis

The χ^2-test was conducted to examine the Hardy–Weinberg equilibrium and to compare the distribution of the tumor status and the Gleason sum among the genotypes. Odds ratios (ORs) with 95% confidence intervals (CIs) for prostate cancer susceptibility were estimated using a logistic regression analysis. These statistical analyses were conducted using the JMP software, version 6.0.3 (SAS, Cary, NC, USA). Results were considered significant when $P < 0.05$.

Results

The genotypes of the polymorphism were successfully determined at position Lys262Arg in all of the subjects. The distributions of the genotypes in all of the groups did not deviate from the Hardy–Weinberg equilibrium (Table 1). Table 1 shows the proportions of the SNP and allele frequencies in the clinical case and control groups. The rates of the Arg/Arg genotype of the Lys262Arg polymorphism in the clinical case group were significantly higher than that in the controls (11.1% *vs* 6.7%). The age-adjusted ORs for the Arg/Arg genotype and the Lys/Arg plus Arg/Arg genotypes for the risk of prostate cancer were 2.202 (95% CI, 1.165 to 4.284; $P = 0.017$) and 1.547 (95% CI, 1.103 to 2.179; $P = 0.012$), respectively. The G allele frequency of the Lys262Arg polymorphism was significantly higher in the cases than that in the controls (30.4% *vs* 24.8%). The age-adjusted OR of the G allele for the risk of prostate cancer was 1.485 (95% CI, 1.136 to 1.946; $P = 0.004$).

We next compared the genotypes of the polymorphism and the clinical characteristics (clinical T stage and Gleason sum) among the cases (Tables 2 and 3). Patients with the Lys/Arg plus Arg/Arg genotypes of the Lys262Arg polymorphism carried a significantly lower Gleason sum than those with the Lys/Lys genotype. The OR of the Lys/Arg plus Arg/Arg genotypes for the risk of a higher Gleason sum were 0.632 (95% CI, 0.414 to 0.964; $P = 0.042$). The G allele frequency of the Lys262Arg polymorphism was significantly higher in the low Gleason sum group than that in the high Gleason sum group (34.3% *vs* 26.8%). The OR of the G allele was 0.701 (95% CI, 0.507 to 0.968; $P = 0.038$). However, no significant difference was observed between the clinical T stage and the Lys262Arg polymorphism.

Discussion

In the present study, we examined the association between the Lys262Arg polymorphism of the CYP2B6 gene and the risk of prostate cancer, the clinical T stage, and the Gleason sum. We derived some significant results. First, the frequency of the Arg/Arg genotype of the

Table 1 The Lys262Arg genotype in patients with clinical cancer and in controls

		No. (%)		Unadjusted		Age adjusted	
		Case	Control	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Genotype							
Lys262Arg							
Lys/Lys		176 (50.3)	187 (57.0)	1.0 (Reference)		1.0 (Reference)	
Lys/Arg		135 (38.6)	119 (36.3)	1.205 (0.874 to 1.662)	0.288	1.421 (0.995 to 2.036)	0.054
Arg/Arg		39 (11.1)	22 (6.7)	1.884 (1.074 to 3.303)	0.027	2.202 (1.165 to 4.284)	0.017
Lys/Arg + Arg/Arg		174 (49.7)	141 (43.0)	1.311 (0.969 to 1.775)	0.090	1.547 (1.103 to 2.179)	0.012
Allele frequency	A allele (Lys)	0.696	0.752	1.0 (Reference)			
	G allele (Arg)	0.304	0.248	1.323 (1.041 to 1.681)	0.025	1.485 (1.136 to 1.946)	0.004
HWE	*P* value	0.250	0.875				

CI, confidence interval; HWE, Hardy–Weinberg equilibrium; OR, odds ratio

Table 2 Relation between the Lys262Arg genotype and clinical T stage

		No. (%) Clinical stage		OR (95% CI)	*P*-value
		Localized (T1 and T2)	Advanced (T3 and T4)		
Genotype					
Lys262Arg					
Lys/Lys		75 (51.0)	101 (49.8)	1.0 (Reference)	
Lys/Arg		55 (37.4)	80 (39.4)	1.080 (0.685 to 1.703)	0.817
Arg/Arg		17 (11.6)	22 (10.8)	0.961 (0.477 to 1.935)	1.000
Lys/Arg + Arg/Arg		72 (49.0)	102 (50.2)	1.052 (0.688 to 1.608)	0.829
Allele frequency	A allele (Lys)	0.697	0.695	1.0 (Reference)	
	G allele (Arg)	0.303	0.305	1.013 (0.731 to 1.404)	1.000

CI, confidence interval; OR, odds ratio

Table 3 Relation between the Lys262Arg genotype and Gleason sum

		No. (%) Gleason sum		OR (95% CI)	*P*-value
		Low (2 to 7A)	High (7B to 10)		
Genotype					
Lys262Arg					
Lys/Lys		75 (44.4)	101 (55.8)	1.0 (Reference)	
Lys/Arg		72 (42.6)	63 (34.8)	0.650 (0.414 to 1.020)	0.067
Arg/Arg		22 (13.0)	17 (9.4)	0.574 (0.285 to 1.155)	0.154
Lys/Arg + Arg/Arg		94 (55.6)	80 (44.2)	0.632 (0.414 to 0.964)	0.042
Allele frequency	A allele (Lys)	0.657	0.732	1.0 (Reference)	
	G allele (Arg)	0.343	0.268	0.701 (0.507 to 0.968)	0.038

CI, confidence interval; OR, odds ratio

Lys262Arg polymorphism in the clinical case group was significantly higher than that among the controls. Second, a statistical difference in the Lys262Arg polymorphisms was observed between the genotype distributions and the Gleason sums. Our results suggest that the Lys262Arg polymorphism of the CYP2B6 gene is significantly associated with the risk of sporadic prostate cancer and may serve as a candidate genetic marker for prostate cancer in the Japanese population. As far as we know, our study is the first report to reveal a significant association between the CYP2B6 polymorphism and the risk of prostate cancer.

The Lys262Arg polymorphism was related to the low Gleason sum, but no significant difference between the polymorphisms and the clinical T stage was observed in our studies. Since cancer is typically diagnosed at various phases, not all cancers with an advanced clinical T stage may be high-grade malignancies. Consequently, patients with the G allele of the Lys262Arg polymorphism may have a better prognosis than those with the A allele because of a lower Gleason sum even if they have a higher prevalence of prostate cancer. However, it is unclear how the Lys262Arg polymorphism, which may hinder the hydroxylation of testosterone, affects pathological characteristics of prostate cancer. Recently several studies have revealed that a low pretreatment serum total testosterone level was associated with a high Gleason sum and an advanced pathological stage.[21–24] Since it is likely that the decrease in hydroxylation of testosterone may elevate concentration of testosterone in blood, our results about the relation between the Lys262Arg polymorphism and the Gleason sum are consistent with these studies. However, we could not clarify a relation between the pretreatment serum total testosterone levels and the Lys262Arg polymorphism, because we could not collect adequate data on the pretreatment serum total testosterone levels to carry out statistical analyses in both groups. We intend to carry out a further study using a larger sample to confirm this relation. Meanwhile, it was reported that men on finasteride, 5-alpha-reductase inhibitor, which inhibits the conversion of testosterone to the androgen DHT, had a lower overall rate of prostate cancer but were at an increased risk for high grade tumors.[25] By decreasing intraprostatic testosterone, finasteride might have created an environment for high grade cancers that were less dependent on androgens for growth.[26,27] As CYP2B6 is known to function in DHT inactivation in the prostate,[8–12] the Lys262Arg polymorphism may increase intraprostatic DHT. If so, the action of finasteride on DHT is opposite to that of the Lys262Arg polymorphism. Therefore our results are compatible with the previous reports on the relation between finasteride and prostate cancer. Heracek *et al.* investigated the correlation of androgen concentration between intraprostatic and serum levels in benign prostatic hyperplasia and prostate cancer. Significantly higher intraprostatic androgen concentrations were found in patients with prostate cancer than in benign prostatic hyperplasia. However, no differences were found in serum levels.[28] Accordingly, it may be hypothesized that intraprostatic androgen and serum testosterone have different roles for prostate cancer. From this point of view, the absence of data regarding the intraprostatic androgen level is a limitation of this study. Further studies are needed to reveal the relation between the circulating testosterone level and the intraprostatic androgen concentrations.

Moreover, other polymorphisms may influence the serum and intraprostatic testosterone levels. For instance, steroid 5-alpha-reductase 2 (SRD5A2) is responsible for the biosynthesis of DHT in the prostate.[29] Söderström *et al.* reported an association with the Leu/Leu genotype of the Val89Leu polymorphism of the SRD5A2 gene and metastases at the time of diagnosis, but no association with an altered risk of prostate cancer was seen.[30] It may be interesting to analyze the effects of several polymorphisms together.

If the influence of the Lys262Arg polymorphism on the risk of prostate cancer was to be characterized in additional studies, patients with the G allele of the Lys262Arg polymorphism might exhibit an improved prognosis if they were to be diagnosed as having prostate cancer during an early stage. In addition, since individuals with the G allele have a relatively higher susceptibility to prostate cancer and a lower Gleason sum than those with the A allele, it may be useful for future studies to compare prognosis according to the Lys262Arg polymorphism.

Conclusions

Our results indicate that the Lys262Arg polymorphism of the CYP2B6 gene is associated with susceptibility to prostate cancer in native Japanese men. In addition, individuals with the G allele of the Lys262Arg polymorphism had a lower Gleason sum than those with the A allele.

Acknowledgments

We thank the following investigators for their help and cooperation in collecting the patients' data: Drs Takeshi Ozeki, Yoshitsugu Yanagihara, Takeo Murayama, Makoto Suzuki, Masaharu Niizuma, Hiromu Tokuda, Shigeharu Kurimoto, Shinsuke Hayami, Hisashi Matsushima, Takamitsu Tachikawa, Yasunori Ishii, Atsushi Kato, Toru Hara, Yasuyuki Asakage, Junzo Tomoishi, Yasuo Hosoi, Takeshi Kawamura, and Yoshikazu Hirano. We also thank Ms Akiko Hirose and Ms Etsuko Tanaka for their excellent technical assistance and Ms Chise Kato for her data collection.

References

1 Ajiki W, Tsukuma H, Oshima A. Cancer incidence and incidence rates in Japan in 1999: estimates based on data from 11 population-based cancer registries. *Jpn. J. Clin. Oncol.* 2004; **34**: 352–6.
2 Isaacs JT. Role of androgens in prostatic cancer. *Vitam. Horm.* 1994; **49**: 433–502.
3 Hsing AW, Reichardt JK, Stanczyk FZ. Hormones and prostate cancer: current perspectives and future directions. *Prostate* 2002; **52**: 213–35.
4 Agoulnik IU, Weigel NL. Androgen receptor action in hormone-dependent and recurrent prostate cancer. *J. Cell. Biochem.* 2006; **99**: 362–72.
5 Ross JS. The androgen receptor in prostate cancer: therapy target in search of an integrated diagnostic test. *Adv. Anat. Pathol.* 2007; **14**: 353–7.
6 Jacob RM, Johnstone EC, Neville MJ *et al.* Identification of CYP2B6 sequence variants by use of multiplex PCR with allele-specific genotyping. *Clin. Chem.* 2004; **50**: 1372–7.
7 Waxman DJ, Attisano C, Guengerich FP *et al.* Human liver microsomal steroid metabolism: identification of the major microsomal steroid hormone 6 beta-hydroxylase cytochrome P450 enzyme. *Arch. Biochem. Biophys.* 1988; 26_3: 424–36.
8 Isaacs JT, McDermott IR, Coffey DS. The identification and characterization of a new C19O3 steroid metabolite in the rat ventral prostate: 5 alpha-androstane-3 beta, 6 alpha, 17 beta-triol. *Steroids* 1979; **33**: 639–57.
9 Isaacs JT, McDermott IR, Coffey DS. Characterization of two new enzymatic activities of the rat ventral prostate: 5 alpha-androstane-3 beta, 17 beta-diol 6 alpha-hydroxylase and 5 alpha-androstane-3 beta, 17 beta-diol 7 alpha-hydroxylase. *Steroids* 1979; **33**: 675–92.
10 Sundin M, Warner M, Haaparanta T *et al.* Isolation and catalytic activity of cytochrome P-450 from ventral prostate of control rats. *J. Biol. Chem.* 1987; **262**: 12293–7.
11 Parkinson A, Mudra DR, Johnson C *et al.* The effects of gender, age, ethnicity, and liver cirrhosis on cytochrome P450 enzyme activity in human liver microsomes and inducibility in cultured human hepatocytes. *Toxicol. Appl. Pharmacol.* 2004; **199**: 193–209.
12 Kittles RA, Chen W, Panguluri RK *et al.* CYP3A4-V and prostate cancer in African Americans: causal or confounding association because of population stratification? *Hum. Genet.* 2002; **110**: 553–60.
13 Antona CR, Sundberg MI. Cytochrome P450 pharmacogenetics and cancer. *Oncogene* 2006; **25**: 1679–91.
14 Lang TA, Klein K, Fisher J *et al.* Extensive genetic polymorphism in the human CYP2B6 gene with impact on expression and function in human liver. *Pharmacogenetics* 2001; **11**: 399–415.

15 Xie HJ, Yasar U, Lundgren S *et al.* Role of polymorphic human CYP2B6 in cyclophosphamide bioactivation. *Pharmacogenomics. J.* 2003; **3**: 53–61.
16 Hesse LM, He P, Krishnaswamy S *et al.* Pharmacogenetic determinants of interindividual variability in bupropion hydroxylation by cytochrome P450 2B6 in human liver microsomes. *Pharmacogenetics* 2004; **14**: 225–38.
17 Sabin LH, Wittekind C. (eds). *TNM Classification of Malignant Tumors*, 6th edn. Wiley-Liss, New York, 2002.
18 Suzuki M, Kurosaki T, Arai T *et al.* The Val158Met polymorphism of the catechol-O-methyltransferase gene is not associated with the risk of sporadic or latent prostate cancer in Japanese men. *Int. J. Urol.* 2007; **14**: 800–4.
19 Sawabe M, Arai T, Kasahara I *et al.* Developments of geriatric autopsy database and Internet-based database of Japanese single nucleotide polymorphisms for geriatric research (JG-SNP). *Mech. Ageing Dev.* 2004; **125**: 547–52.
20 Livak KJ. Allelic discrimination using fluorogenic probes and the 5′ nuclease assay. *Genet. Anal.* 1999; **14**: 143–9.
21 Massengill JC, Sun L, Moul JW *et al.* Pretreatment total testosterone level predicts pathological stage in patients with localized prostate cancer treated with radical prostatectomy. *J. Urol.* 2003; **169**: 1670–5.
22 Schatzl G, Madersbacher S, Thurridl T *et al.* High-grade prostate cancer is associated with low serum testosterone levels. *Prostate* 2001; **47**: 52–8.
23 Isom-Batz G, Bianco FJ Jr, Kattan MW *et al.* Testosterone as a predictor of pathological stage in clinically localized prostate cancer. *J. Urol.* 2005; **173**: 1935–7.
24 Yamamoto S, Yonese J, Kawakami S *et al.* Preoperative serum testosterone level as an independent predictor of treatment failure following radical prostatectomy. *Eur. Urol.* 2007; **52**: 696–701.
25 Thompson IM, Goodman PJ, Tangen CM *et al.* The influence of finasteride on the development of prostate cancer. *N. Engl. J. Med.* 2003; **349**: 215–24.
26 Scardino PT. The prevention of prostate cancer – the dilemma continues. *N. Engl. J. Med.* 2003; **349**: 297–9.
27 Morgentaler A, Bruning CO 3rd, DeWolf WC. Occult prostate cancer in men with low serum testosterone levels. *JAMA* 1996; **276**: 1904–6.
28 Heracek J, Hampl R, Hill M *et al.* Tissue and serum levels of principal androgens in benign prostatic hyperplasia and prostate cancer. *Steroids* 2007; **72**: 375–80.
29 Neslund-Dudas C, Bock CH, Monaghan K *et al.* SRD5A2 and HSD3B2 polymorphisms are associated with prostate cancer risk and aggressiveness. *Prostate* 2007; **67**: 1654–63.
30 Söderström T, Wadelius M, Andersson SO *et al.* 5alpha-reductase 2 polymorphisms as risk factors in prostate cancer. *Pharmacogenetics* 2002; **12**: 307–12.

Hum Genet (2009) 126:539–547
DOI 10.1007/s00439-009-0698-6

ORIGINAL INVESTIGATION

Megakaryoblastic leukemia factor-1 gene in the susceptibility to coronary artery disease

Kunihiko Hinohara · Toshiaki Nakajima · Michio Yasunami · Shigeru Houda · Taishi Sasaoka · Ken Yamamoto · Bok-Soo Lee · Hiroki Shibata · Yumiko Tanaka-Takahashi · Megumi Takahashi · Takuro Arimura · Akinori Sato · Taeko Naruse · Jimin Ban · Hidetoshi Inoko · Yoshiji Yamada · Motoji Sawabe · Jeong-Euy Park · Toru Izumi · Akinori Kimura

Received: 30 April 2009 / Accepted: 30 May 2009 / Published online: 10 June 2009

Abstract Coronary artery disease (CAD) is based on the atherosclerosis of coronary artery and may manifest with myocardial infarction or angina pectoris. Although it is widely accepted that genetic factors are linked to CAD and several disease-related genes have been reported, only a few could be replicated suggesting that there might be some other CAD-related genes. To identify novel susceptibility loci for CAD, we used microsatellite markers in the screening and found six different candidate CAD loci. Subsequent single nucleotide polymorphism (SNP) association studies revealed an association between CAD and megakaryoblastic leukemia factor-1 gene (*MKL1*). The association with a promoter SNP of *MKL1*, −184C > T, was found in a Japanese population and the association was replicated in another Japanese population and a Korean population. Functional analysis of the *MKL1* promoter SNP suggested that the higher MKL1 expression was associated with CAD. These findings suggest that *MKL1* is involved in the pathogenesis of CAD.

Introduction

Coronary artery disease (CAD) is a major health problem in developed countries because of its increasing prevalence and high mortality (Lopez et al. 2006). CAD is

Electronic supplementary material The online version of this article (doi:10.1007/s00439-009-0698-6) contains supplementary material, which is available to authorized users.

K. Hinohara · T. Nakajima · M. Yasunami · H. Shibata · Y. Tanaka-Takahashi · M. Takahashi · T. Arimura · A. Sato · T. Naruse · A. Kimura (✉)
Department of Molecular Pathogenesis,
Medical Research Institute, Tokyo Medical and Dental University,
1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan
e-mail: akitis@mri.tmd.ac.jp

K. Hinohara · T. Nakajima · M. Yasunami · A. Kimura
Laboratory of Genome Diversity, School of Biomedical Science,
Tokyo Medical and Dental University, Tokyo, Japan

S. Houda · T. Sasaoka · T. Izumi
Department of Internal Medicine and Cardiology,
Kitasato University School of Medicine, Sagamihara, Japan

K. Yamamoto
Division of Molecular Population Genetics,
Department of Molecular Genetics,
Medical Institute of Bioregulation,
Kyushu University, Fukuoka, Japan

B.-S. Lee · J. Ban · J.-E. Park
Division of Cardiology, Samsung Medical Center,
Sungkyunkwan University School of Medicine, Seoul, Korea

H. Inoko
Department of Molecular Life Science,
Tokai University School of Medicine, Isehara, Japan

Y. Yamada
Department of Human Functional Genomics,
Mie University Graduate School of Medicine, Tsu, Japan

M. Sawabe
Department of Pathology,
Tokyo Metropolitan Geriatric Medical Center, Tokyo, Japan

based on the atherosclerosis of coronary artery and often manifests with sudden chest pain due to reversible (angina pectoris, AP) or irreversible (myocardial infarction, MI) ischemia in the heart. Smoking, hypertension, hypercholesterolemia and diabetes mellitus significantly contribute to the development of CAD (Wang 2005) but several lines of evidence indicate that genetic factors also play important roles in the pathogenesis of CAD (Ciruzzi et al. 1997).

Recent genome-wide association studies using single nucleotide polymorphism (SNP) have accumulated tremendous information about the susceptibility genes for common diseases including CAD. For instance, SNPs on the chromosome 9p21 were repeatedly reported to be associated with CAD (Wellcome Trust Case Control Consortium 2007; McPherson et al. 2007; Helgadottir et al. 2007; Samani et al. 2007; Hinohara et al. 2008). In contrast, a number of genetic variants that have been reported to confer a risk of CAD failed to be replicated in multiple case–control studies (Morgan et al. 2007), suggesting the importance of replication/validation studies in identification of the CAD-associated genes. Another strategy to identify the disease-associated loci was a multistep screening using microsatellite (MS) markers. Because MS markers display higher degree of heterozygosity as compared with SNPs, a smaller number of MS markers may provide a reasonable statistical power in the association analyses (Tamiya et al. 2005; Kawashima et al. 2006; Ohashi and Tokunaga 2003). It is well known that one of the major concerns of the association studies is the risk of false-positive association. There has been much debate concerning the reduction in the risk of false-positive association in large scale associations (Cardon and Bell 2001). The adjustment for multiple testing is a way to reduce the risk, and the multi-stage testing is an alternative.

We report here the identification of candidate CAD loci by the multistep screening using MS markers. Subsequent association studies revealed a promoter SNP of megakaryoblastic leukemia factor-1 gene (*MKL1*) was associated with CAD in Japanese and Korean populations. Functional analysis of the promoter SNP supported the role of *MKL1* in the atherosclerosis.

Materials and methods

Subjects

A total of 2,566 CAD cases and 4,100 control subjects from two Japanese populations and one Korean population were the subjects in this study. One Japanese population (group A) consisted of 629 (579 MI and 50 AP) cases and 1,232 controls. The cases were recruited from Kitasato University Hospital and related hospitals, while the controls were composed of healthy individuals selected at random ($n = 632$) and healthy-donor derived Epstein–Barr (EB) virus-transformed human B cell lines ($n = 600$) obtained from Japan Health Sciences Foundation. The other Japanese population (group B) composed of 1,109 MI cases and 2,122 controls who visited outpatient clinics of or were admitted to one of the five participating hospitals (Gifu Prefectural General Medical Center, Gifu Prefectural Tajimi Hospital, Hirosaki University Hospital, Reimeikyo Rehabilitation Hospital, and Hirosaki Stroke Center) because of various symptoms or for an annual health checkup. The control subjects had no history of coronary artery disease. The Korean population consisted of 828 (461 MI and 367 AP) cases and 746 controls including healthy individuals without ischemic heart diseases ($n = 222$) and cancer cases who carried no history of ischemic heart diseases ($n = 524$) recruited from Samsung Medical Center. The diagnosis of CAD was based on the standard criteria as described previously (Hohda et al. 2003). Severity of coronary atherosclerosis was classified according to the number of coronary vessels with significant stenosis (angiographic luminal stenosis >50%) as 0, 1, 2 or 3 vessel disease (VD). Informed consent was given from each participant and the study was approved by the Ethics Review Boards of Medical Research Institute of Tokyo Medical and Dental University, Kitasato University School of Medicine, Mie University, and Samsung Medical Center.

Strategy for screening of CAD loci

Multistep screening of CAD loci was done by using a total of 18,880 MS markers in the Japanese population (Japanese group A). Constitutions of three screening samples were as follows: the first set of 100 MI cases with 3 VD versus 100 controls, the second set of 100 MI cases with 2 VD versus 100 controls, and the third set of 192 MI cases consisting of two 3 VD, 46 2 VD and 144 1 VD with onset age younger than 50 years-old cases versus 192 controls) from the Japanese group A. Equal amounts of genomic DNAs from cases and controls were mixed to prepare a set of pooled DNA samples for cases and patients, respectively, according to the method described by Tamiya et al. (2005). In brief, genomic DNA concentration was measured in triplicate, in accordance with the methods of Collins et al. (2000) with the use of a double-stranded DNA quantification kit (Pico Green [Molecular Probes]) and each genomic DNA sample was adjusted to 8 ng/μl. As a pilot study, we tested two independently prepared pairs of pooled DNAs from the same set of panels for polymorphic pattern of 400 MS and found that the variation in distribution of polymorphic peak heights for each MS was less than 2%, suggesting the reproducibility of pooled DNA preparations. DNA samples

from cases in the first set were mixed to prepare a pooled case DNA for the first set. As well, DNA samples from cases in the second set and those from controls in the first and second sets were used to prepare the pooled case DNA for the second set and pooled control DNAs for the first and second sets, respectively.

In the first screening, we performed a case–control study using a set of pooled DNAs. The MS markers showing significant association with CAD ($P < 0.05$ in the 2-by-2 analysis and/or 2-by-multiple analysis) in the first screening were investigated for the association using another set of pooled DNAs prepared from different cases and controls (second screening). The MS markers showing statistical significance in both first and second screenings were examined for the association by individual typing of DNA samples from the first and second sets. The MS markers confirmed for the association by individual typing were subjected to the third screening using different cases and controls. Detailed information about the MS markers is available at http://www.tmd.ac.jp/mri/mri-mpath/MS-List.xls.

Analysis of SNPs in MKL1

We performed SNP analyses of *MKL1* in the CAD-associated locus. To further validate the association with specific SNP, replication studies using two different populations (Japanese group B and Korean group) were conducted to avoid the chance of false-positive findings (Hirschhorn and Daly 2005). All SNPs were genotyped by using TaqMan SNP genotyping assay (Applied Biosystems) in this study. Tag SNPs of *MKL1* were selected under the following conditions; r^2 cut off value of 0.8 and minor allele frequency cut off value of 0.05 in Japanese.

Reporter-gene assay

Reporter-gene assays were performed using luciferase-reporter constructs containing −184C > T alleles. A 249-bp fragment of *MKL1* promoter region deduced from database of transcriptional start sites (http://dbtss.hgc.jp/) was amplified by PCR and cloned into pGL3-basic vector (Promega). Luciferase activity in each lysate was determined as described previously (Shibata et al. 2006). Data were obtained from four independent experiments (each in quadruplicate). Statistical analyses were done using Student's *t*-test.

Expression of MKL1 transcript in B-cell lines

Human B cell lines listed as International Histocompatibility Workshop cell lines (http://www.ihwg.org/cellbank/cell_lines/blcl.html) or those obtained from Japan Health Sciences Foundation were genotyped for the promoter polymorphism at −184 to select five B cell lines each with −184C/C or −184T/T genotype. Quantitative real-time PCR was performed as described previously (Shibata et al. 2006). Primer pairs used were as follows: 5′-TAGCCGA TGACCTCAATGAG-3′ with 5′-ATCGAAGGAAGAGC TGTCTG-3′ for *MKL1* and 5′-CTTCACCACCATGGAG AAGGC-3′ with 5′-GGCATGGACTGTGGTCATGAG-3′ for GAPDH. We repeated the experiments three times and studied each sample in triplicate in each experiment. Statistical analyses were done using Student's *t*-test.

Statistical analysis

P values less than 0.05 were considered to be significant throughout this study. To assess the extent of linkage disequilibrium (LD) across the 22q13 chromosomal region, D' and r^2 were calculated using Haploview software (Barrett et al. 2005). Recombination hot spots were estimated from HapMap data (Phase II, release 21). Genotype distribution and allele frequency of SNPs were compared between the cases and controls using a chi-square test. Strength of the association was expressed by odds ratio (OR). Meta-analyses were performed using a Mantel-Haenszel method (Mantel and Haenszel 1959; Helgadottir et al. 2007).

Results

Multistep screening of candidate CAD loci

We conducted a multistep screening of CAD loci using 18,880 MS markers by pooled DNA method in a Japanese population (Japanese group A, Table 1) to identify the candidate loci (Supplementary Text S1 and Supplementary Figure S1). Strategy for the screening is schematically shown in Fig. 1. As a result, we identified six candidate CAD loci (Supplementary Tables S1, S2). Linkage disequilibrium (LD) analysis of these six candidate loci based on HapMap JPT + CHB data identified that three MS markers in the candidate loci CAD3, CAD4, and CAD5 were located in the LD blocks containing no known genes, while MS markers representative of CAD1, CAD2, and CAD6 were within the LD-blocks carrying megakaryoblastic leukemia facter-1 gene (*MKL1*), proprotein convertase subtilisin/kexin type 9 (*PCSK9*), and AK125001, respectively. Because AK125001 was a gene of unknown function, we focused on the CAD1 (*MKL1*) and CAD2 (*PCSK9*) loci on the chromosome 22q13.3 and 1p32.3, respectively.

PCSK9 was reported to be associated with hyperlipidemia (Abifadel et al. 2003; Chen et al. 2005; Evans and Beil 2006; Willer et al. 2008), prompting us to investigate the association between CAD2-MS and CAD in relation

Table 1 Characteristics of cases and controls

	Japanese A		Japanese B		Korean	
	CAD	Control	CAD	Control	CAD	Control
	$n = 629$	$n = 1,232$	$n = 1,109$	$n = 2,122$	$n = 828$	$n = 746$
Age (years)	59.3 ± 10.1	39.0 ± 10.7	63.7 ± 10.6	67.6 ± 10.0	61.2 ± 11.1	58.4 ± 11.7
BMI	23.7 ± 2.9	n.a.	23.6 ± 3.2	23.5 ± 3.2	24.7 ± 2.6	n.a.
Gender (male %)	83.3	56.3	78.1	42.5	76.4	65.6
Smoking (%)	73.6	n.a.	21.6	24.5	69.3	n.a.
HT (%)	53.4	n.a.	72.5	30.3	45.2	n.a.
HC (%)	49.9	n.a.	57.0	26.0	43.5	n.a.
DM (%)	30.3	n.a.	48.6	7.9	28.3	n.a.

Characteristics of panels used in this study are shown. The values are means ± SD and percentages where indicated. Smoking: current or former smoking of ≥10 cigarettes daily. *HT* hypertension, systolic blood pressure of ≥140 mmHg, diastolic blood pressure of ≥90 mmHg, or taking antihypertensive medication. *HC* hypercholesterolemia, serum total cholesterol of ≥5.72 mmol/l (220 mg/dl) or taking lipid-lowering medication. *DM* diabetes mellitus, fasting blood glucose of ≥6.93 mmol/l (126 mg/dl), glycosylated hemoglobin of ≥6.5%, or taking antidiabetes medication

with classical risk factors including hypercholesterolemia. We genotyped additional cases and controls for CAD2-MS and the data were combined with those obtained from the screening to investigate the association (Supplementary Table S3). It was demonstrated that the significant association with CAD2-MS was found exclusively in cases with hypercholesterolemia and not in normocholesterolemia cases, suggesting that the association between CAD2-MS and CAD was a reflection of the association between CAD2-MS and hypercholesterolemia. In addition, the analysis of tag SNPs in *PCSK9* revealed that one SNP (23968A > G) showed significant association with CAD (G allele frequencies in 430 cases and 332 controls were 0.054 and 0.032, respectively, giving OR = 1.73, $P = 0.039$).

On the other hand, the analysis of the data for CAD1-linked MS (CAD1-MS) demonstrated that the frequencies of $(ATAC)_9$ allele (A05 allele in Supplementary Table S2) were 0.132 and 0.090 in the cases and controls, respectively, showing a significant association with the susceptibility [odds ratio (OR) = 1.54, $P = 0.012$). In addition, the frequencies of $(ATAC)_{13}$ allele (A09 allele in Supplementary Table S2) were 0.118 and 0.173, respectively (OR = 0.64, $P = 0.004$). Recombination hot spots inferred from HapMap data were located on 40 kb upstream and in the 3′ side of *MKL1* (Fig. 2). CAD1-MS was located within *MKL1* gene and strong linkage disequilibrium (LD) was observed across the whole *MKL1* gene including promoter region, implying the existence of the responsible SNP in *MKL1*.

SNP study of MKL1 in association with CAD

To search for SNPs of *MKL1*, 32 Japanese samples were sequenced for exons and adjacent introns as well as upstream promoter region up to 1 kb and 12 SNPs, including seven novel SNPs, were identified (Supplementary Table S4). Two promoter SNPs, −184C > T and -173T/C (rs4140512), and two coding SNPs, 1848G > A and 1942A > G (rs878756), as well as seven other SNPs selected from the HapMap database were tested for the association with CAD using 432 cases and 432 controls from Japanese group A, and a significant association of CAD with −184C > T (allele frequencies were 0.104 and 0.080 in the cases and controls, respectively. OR = 1.44, $P = 0.033$) was observed, whereas the other SNPs did not show significant association (Supplementary Table S5).

Combined analysis of CAD1-MS and SNPs revealed that −184C > T was in tight LD with CAD1-MS A05 allele ($D' = 0.88$, $r^2 = 0.75$), suggesting that −184C > T might be responsible for the association of CAD1 locus with CAD. To further strengthen the finding, we examined −184C > T in all available samples from Japanese group A (629 cases and 1,232 controls, Table 1) and confirmed the significant association (OR = 1.30, 95% confidence interval (CI) = 1.04–1.64, $P = 0.024$; Table 2).

Validation of the observed association in different populations

To validate the observed association, we investigated −184C > T polymorphism in another Japanese population (group B; 1,109 cases and 2,122 controls) and Korean population (828 cases and 746 controls) (Table 1). As shown in Table 2, the association with −184C > T was replicated in both Japanese (OR = 1.25, 95% CI = 1.04–1.49, $P = 0.015$) and Korean (OR = 1.26, 95% CI = 1.01–1.58, $P = 0.042$) populations. Meta-analysis of data from Japanese A, Japanese B and Korean panels indicated that the association was significant ($P = 1.1 \times 10^{-4}$) with an OR of 1.27

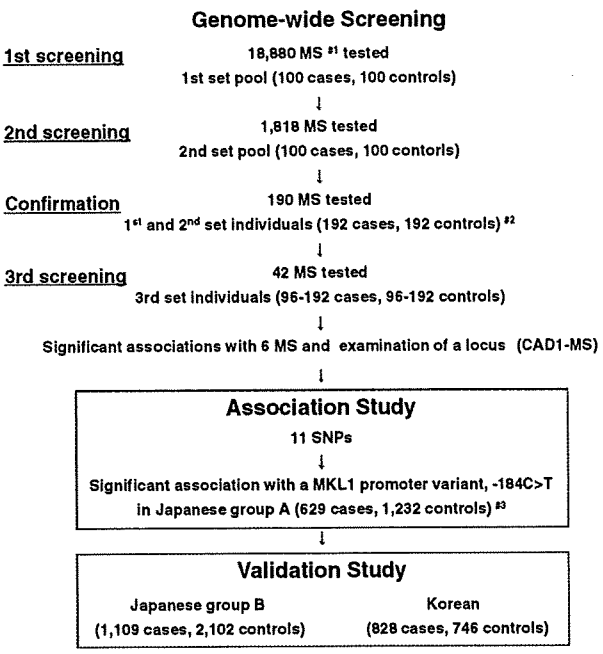

Fig. 1 Study design for identification and validation of sequence variants associated with CAD. #1; MS is an abbreviation of microsatellite marker. #2; Samples selected at random from the panel used in the first and second screenings. #3, Study population (Japanese A, Table 1) includes the samples used in the screening. First screening was done for 18,880 MS between the pooled DNA from 100 MI cases and that from 100 controls, resulting in the identification of 1,818 MS showing significant association. These 1,818 MS were tested for the association with CAD (MI) by using pooled DNAs from 100 MI cases and 100 controls, obtaining 190 MS demonstrating the significant association. Both the cases and controls in the second screening were different from those in the first screening. These 190 MS were investigated in 192 MI cases and 192 controls randomly selected from the panels used in the first and second screenings, and 42 MS were confirmed for the significant association with MI. These 42 MS were tested in the third set of screening in which 192 (or 96) MI cases and 192 (or 96) controls were analyzed, resulting in the identification of 6 MS as the candidate loci for CAD. One locus (CAD1) was selected for further association study of 11 SNPs and a promoter SNP, −184C > T, was found to be significantly associated with CAD in a Japanese population (Japanese group A). The association was validated in another Japanese population (Japanese group B) and a Korean population

(95%CI = 1.12–1.43) for the risk allele −184T. To investigate the mode of inheritance, we calculated genotype-specific ORs for the −184C > T and the results showed that ORs for CC, CT and TT genotypes were 0.79, 1.24 and 1.77, respectively, suggesting the additive model (Table 3).

To address the correlation between the risk allele and clinical background of CAD, we investigated the association between −184C > T and severity of coronary atherosclerosis. According to the number of significantly affected vessels, cases were classified into four groups, 0VD, 1VD, 2VD, and 3VD. There was a strong association between −184C > T and 3VD in each population (Table 2). In the combined data analysis, −184 T allele showed high risk in 3VD group (OR = 1.47, $P = 2.7 \times 10^{-5}$, Table 2). Stratified analyses of −184C > T with classical CAD risk factors including hypertension, hypercholesterolemia, and diabetes mellitus showed no trend of stratified association with any specific risk factors (data not shown).

Expression analysis

Because −184C > T locates in the promoter region of *MKL1*, we investigated whether it would influence the transcription level by using reporter-gene assays. As shown in Fig. 3, it was observed that the luciferase activity driven by the susceptible allele T showed reproducibly higher transcriptional activity than the allele C in both K562 (Fig. 3a) and HeLa (Fig. 3b) cells. We next examined the steady-state mRNA levels of *MKL1* in human B lymphoblastoid cell lines homozygous for −184C or −184T. As shown in Fig. 3c, the T allele homozygotes showed about 1.8 (1.5–2.1) fold higher expression than the C allele homozygotes, indicating that the −184T allele was associated with higher expression of *MKL1*.

Discussion

In this study, a multistep screening for CAD loci by using 18,880 MS markers identified 6 candidate loci for CAD and subsequent SNP analyses deciphered *MKL1* as a novel CAD-related gene. In the screening procedures using MS markers, we analyzed only MI patients as the cases, but the controls were not age-matched to and younger than the patients. Therefore, the power to detect the CAD-associated loci was considered to be low since the controls should include individuals who would develop CAD in the future. In addition, power calculation for our multistep screening strategy estimated the power to detect a significant ($P < 0.05$) risk allele conferring OR over 2.0 in additive model by an MS marker in LD ($D' = 0.9$) was 0.53. It was then considered that a number of CAD loci with small contribution would not be captured in our study design. However, multistep screening strategy using MS markers and subsequent validation studies in two independent populations should give highly-reliable data. Recent studies with similar approach have successfully identified, albeit not all, the genes responsible for multifactorial diseases (Tamiya et al. 2005; Shiffman et al. 2005; Kawashima et al. 2006).

There are several genome-wide association studies for CAD in Caucasoid populations (Wellcome Trust Case Control Consortium 2007; Helgadottir et al. 2007; Samani et al. 2007; Erdmann et al. 2009; Trégouët et al. 2009; Myocardial Infarction Genetics Consortium 2009; Coronary Artery Disease Consortium 2009) and Asian populations (Ozaki et al. 2002; Ebana et al. 2007; Ozaki et al. 2009), reporting

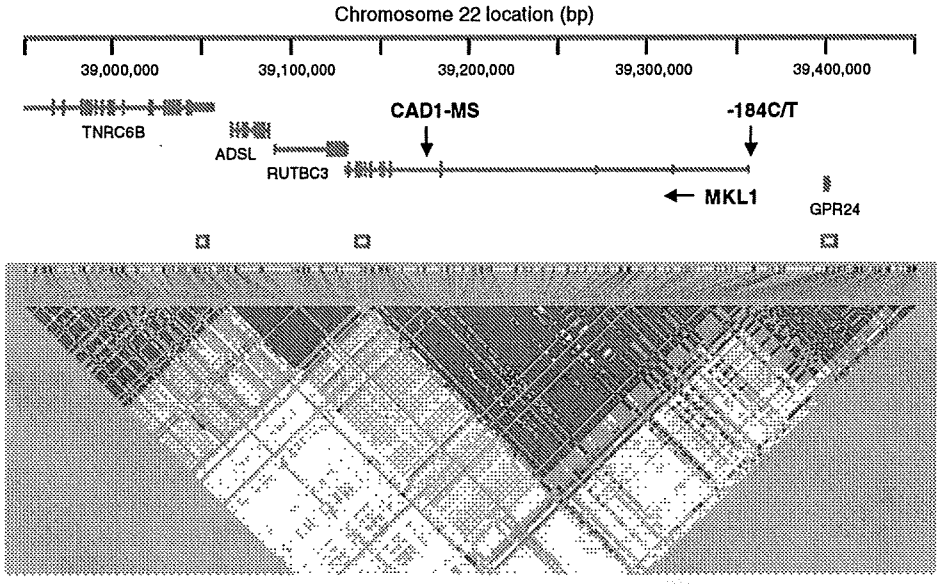

Fig. 2 LD structure around *MKL1* in Mongoloid populations. Upper panel shows the location of genes according to the Reference Sequence collection of the National Center for Biotechnology Information (NCBI). Data were obtained from the Genome Browser of the University of California-Santa Cruz (NCBI build 35). Red boxes represent recombination hot spots deduced from the HapMap data (Phase II, release 21). Lower panel demonstrates the LD structure around *MKL1* indicated by D' values for all SNPs from HapMap JPT (Japanese) + CHB (Chinese) data. Strength of the LD between SNPs increased from *white* to *blue* to *red* (white, $D' < 1$ and LOD score <2; *blue*, $D' = 1$ and LOD score <2; *pink* or *light red*, $D' < 1$ and LOD score ≥ 2; and *bright red*: $D' = 1$ and LOD score ≥ 2)

Table 2 Association of −184 T allele of *MKL1* with CAD

Study population (*n*1/*n*2/*n*3)	Control	All CAD cases			3 VD group		
	Frequency	Frequency	OR (95% CI)	*P*	Frequency	OR (95% CI)	*P*
Japanese group A (1,232/629/174)	0.083	0.106	1.30 (1.04–1.64)	0.024	0.129	1.64 (1.16–2.31)	0.0047
Japanese group B (2,122/1,109/394)	0.079	0.097	1.25 (1.04–1.49)	0.015	0.104	1.35 (1.05–1.74)	0.020
Korean (746/828/121)	0.099	0.122	1.26 (1.01–1.58)	0.042	0.149	1.59 (1.07–2.35)	0.020
Combined							
Japanese (3,354/1,738/568)	0.081	0.100	1.27 (1.10–1.46)	9.5×10^{-4}	0.112	1.44 (1.18–1.78)	4.2×10^{-4}
All groups (4,100/2,566/689)	0.084	0.107	1.27 (1.12–1.43)	1.1×10^{-4}	0.118	1.47 (1.23–1.76)	2.7×10^{-5}

Frequency of −184T allele in the control was shown as compared with those in all CAD cases and cases with 3 vessel disease (VD) along with odd risk (OR), 95% confidence interval (95% CI), and *P* values. Number of controls (*n*1), all CAD cases (*n*2) and CAD cases with 3VD (*n*3) are shown in the parentheses

Upper three columns indicate the data from different study panels; Japanese group A, Japanese group B and Korean. Data for the combined analysis of two Japanese panels and that of all three panels are indicated in the lower two columns

the association of CAD with SNPs in the regions or genes on the chromosomes 1p13.3 (*PSRC1-CELRS2-MYBPHL-SORT1*), 1p32 (*PCSK9*), 1q41 (*MIA3*), 1q43, 2q33 (*WDR12*), 2q36.3, 3p21.1 (*ITIH3*), 3q22.3 (*MRAS*), 5q21, 6p21 (*LTA*), 6p24 (*PHACTR1*), 6q25.1 (*MTHFD1L*), 6q25.3 (*SLC22A3-LPAL2-LPA*), 9p21 (*CDK2A-CDK2B-ARF-MTAP-ANRIL*), 10q11.2 (*CXCL12*), 12q24.1 (*BRAP*), 12q24.3 (*HNF1A*), 15q22.3 (*SMAD3*), 16q23, 19q12, 19q13 (*LDLR*), and 21q22 (*SLC5A3-MPRS6-KCNE2*). Among them, only *PSCK9* locus was overlapped with the CAD loci identified in this study. The reason why the other loci were not detected in our study might be that the power of our screening design was not sufficient to detect the loci, especially those with relatively small contribution to the disease as discussed earlier or that not all of the disease-related genes were common among different ethnic groups. Nevertheless, the detection of *PCSK9* as the CAD2 locus in our study supported the usefulness of our strategy in detecting the disease genes for CAD.

This study unraveled that the *MKL1* promoter SNP correlated with high expression level was associated with the susceptibility to CAD. In addition, our study showed that the association was prominent in the cases with severe atherosclerosis (3 VD group), suggesting a potential role

Table 3 Genotype-specific OR for −184C > T

Study population (*n*1/*n*2)	Genotype-specific OR (95% CI)		
	CC	CT	TT
Japanese A (1,232/629)	0.75 (0.59–0.96)	1.32 (1.02–1.70)	1.38 (0.52–3.63)
Japanese B (2,122/1,109)	0.80 (0.66–0.97)	1.23 (1.01–1.50)	1.56 (0.82–2.66)
Korean (746/828)	0.81 (0.63–1.03)	1.17 (0.91–1.50)	2.96 (0.96–9.11)
Combined			
Japanese (3,354/1,738)	0.78 (0.67–0.91)	1.26 (1.08–1.48)	1.49 (0.83–2.67)
All groups (4,100/2,566)	0.79 (0.69–0.90)	1.24 (1.08–1.41)	1.77 (1.06–2.95)

Odds risks (OR) for −184 genotypes are indicated along with 95% confidence intervals (95%CI) in parentheses. Numbers of controls (*n*1) and CD cases (*n*2) in three different populations are shown in the parentheses after the name of studied panels. Upper three columns indicate the data from Japanese group A, Japanese group B and Korean. Data for the combined analysis of two Japanese panels and that of all three panels are indicated in the lower two columns

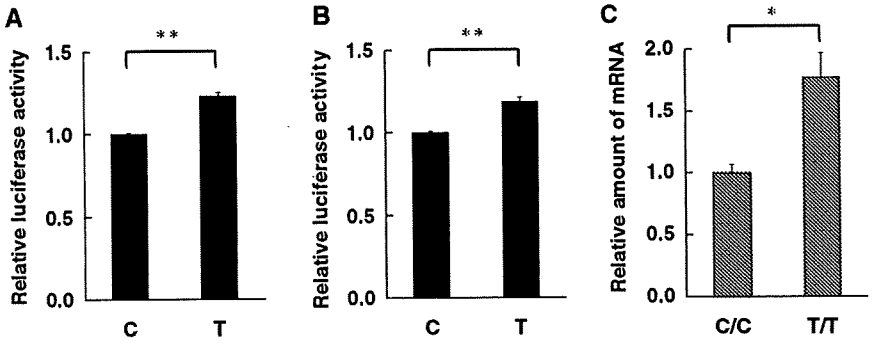

Fig. 3 Transcriptional activity of *MKL1* in association with −184C > T. Luciferase activities expressed from reporter gene containing each promoter allele transfected into K562 cell (**a**) and HeLa cell (**b**). A 249 bp fragment containing the *MKL1* promoter region with −184C or −184T was connected to firefly luciferase gene in pGL3-basic vector. The construct was transfected into the cells and the luciferase activities of transfected cells were measured along with renilla luciferase activities as the internal controls to correct the transfection efficacy. Relative luciferase activity (firefly luciferase activity/renilla luciferase activity) was calculated and that for −184C allele was arbitrarily defined as 1.0. The experiments were repeated four times with each sample in quadruplicate. *Mean error bars* represent SD. **c**; Quantification of *MKL1* mRNA in B cell lines. *Error bars* represent SEM. The values were obtained by dividing the amount of *MKL1* mRNA by amount of *GAPDH* mRNA in the same sample. Data from five B cell lines with −184CC genotype were combined and mean value was arbitrarily defined as 1.0, while the data from other five B cell lines with −184TT genotype were combined. *An asterisk* indicates $P < 0.05$ while *double asterisks* correspond to $P < 0.01$

of *MKL1* in the pathogenesis of atherosclerosis. Given that atherosclerosis is the most common underlying cause of CAD (Ross 1999), *MKL1* expression level may contribute to the progression of atherosclerosis. It has been reported that *MKL1* is a member of myocardin family serving as a SRF cofactor that strongly activates genes under the control of SRF (Cen et al. 2004). *MKL1* was reported to induce the differentiation of smooth muscle cells via transactivation of SRF-dependent transcriptional regulatory elements (Du et al. 2004). In addition, *MKL1* played crucial roles in regulating differentiation of human somatic stem cells into smooth muscle cells, because overexpression of *MKL1* in undifferentiated embryonic stem cells induced expression of genes encoding smooth muscle cell differentiation markers in the SRF-dependent manner (Du et al. 2004). It was recently demonstrated that among blood cells there might be progenitors of smooth-muscle cells, which attached to the endothelium and differentiated into smooth-muscle cells, and these progenitors might contribute to neointima formation in the coronary artery (Sata et al. 2002). In addition, *MKL1* is tightly linked to RhoA-dependent and TGF-beta-dependent signaling pathways in smooth muscle cells, both of which are important in the progression of atherosclerosis (Cen et al. 2004; Elberg et al. 2008). Therefore, these lines of evidence suggested that high expression of *MKL1* was involved in the pathogenesis of atherosclerosis and CAD.

On the other hand, it was reported that *MKL1* was required for the development of cardiovascular system, because *MKL1* null mice were born at less than the predicted Mendelian frequency (Parmacek 2007). Fetal loss was attributed to dilated cardiac chambers, aorta, and pulmonary arteries accompanied by heart failure, which was observed in 35% of *MKL1* null embryos. Approximately

40% of *MKL1* null mice embryos suffered from lethal cardiac cell necrosis with the mitochondrial dysfunction. These observations suggested important roles of *MKL1* in the development, formation and maintenance of arterial tissues, and hence *MKL1* might also play an additional role in the coronary atherosclerosis.

In conclusion, multistep screening approach identified a candidate CAD loci containing *MKL1* on chromosome 22q13. Subsequent SNP analyses revealed that a *MKL1* promoter SNP was associated with the susceptibility to CAD. The *MKL1* promoter SNP conferred high expression of *MKL1*. These observations strongly suggested the involvement of *MKL1* in the pathogenesis of CAD.

Acknowledgments We are grateful Dr. Gen Tamiya for his help in the MS analysis and Dr. Akihiko Hamamatsu for the preparation of coronary artery sections. This work was supported in part by Grant-in-Aids for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, a research grant from the Ministry of Health, Labour and Welfare, Japan, research grants for Japan-Korea collaboration research from Japan Society for the Promotion of Science and by a KOSEF grant, the Korea-Japan Basic Scientific Cooperation Program.

References

Abifadel M, Varret M, Rabès JP, Allard D, Ouguerram K et al (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 34:154–156

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

Cardon LR, Bell JI (2001) Association study designs for complex diseases. Nat Rev Genet 2:91–99

Cen B, Selvaraj A, Prywes R (2004) Myocardin/MKL family of SRF coactivators: key regulators of immediate early and muscle specific gene expression. J Cell Biochem 93:74–82

Chen SN, Ballantyne CM, Gotto AM Jr, Tan Y, Willerson JT et al (2005) A common PCSK9 haplotype, encompassing the E670G coding single nucleotide polymorphism, is a novel genetic marker for plasma low-density lipoprotein cholesterol levels and severity of coronary atherosclerosis. J Am Coll Cardiol 45:1611–1619

Ciruzzi M, Schargrodsky H, Rozlosnik J, Pramparo P, Delmonte H et al (1997) Frequency of family history of acute myocardial infarction in patients with acute myocardial infarction. Am J Cardiol 80:122–127

Collins HE, Li H, Inda SE, Anderson J, Laiho K et al (2000) A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. Hum Genet 106:218–226

Coronary Artery Disease Consortium (2009) Large scale association analysis of novel genetic loci for coronary artery disease. Arterioscler Thromb Vasc Biol 29:774–780

Du KL, Chen M, Li J, Lepore JJ, Mericko P et al (2004) Megakaryoblastic leukemia factor-1 transduces cytoskeletal signals and induces smooth muscle cell differentiation from undifferentiated embryonic stem cells. J Biol Chem 279:17578–17586

Ebana Y, Ozaki K, Inoue K, Sato H, Iida A et al (2007) A functional SNP in ITIH3 is associated with susceptibility to myocardial infarction. J Hum Genet 52:220–229

Elberg G, Chen L, Elberg D, Chan MD, Logan CJ et al (2008) MKL1 mediates TGF-beta1-induced alpha-smooth muscle actin expression in human renal epithelial cells. Am J Physiol Renal Physiol 294:F1116–F1128

Erdmann J, Grosshennig A, Braund PS, König IR, Hengstenberg C et al (2009) New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet 41:280–282

Evans D, Beil FU (2006) The E670G SNP in the PCSK9 gene is associated with polygenic hypercholesterolemia in men but not in women. BMC Med Genet 7:66

Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T et al (2007) A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 316:1491–1493

Hinohara K, Nakajima T, Takahashi M, Hohda S, Sasaoka T et al (2008) Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. J Hum Genet 53:357–359

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108

Hohda S, Kimura A, Sasaoka T, Hayashi T, Ueda K et al (2003) Association study of CD14 polymorphism with myocardial infarction in a Japanese population. Jpn Heart J 44:613–622

Kawashima M, Tamiya G, Oka A, Hohjoh H, Juji T et al (2006) Genomewide association analysis of human narcolepsy and a new resistance gene. Am J Hum Genet 79:252–263

Lopez AD, Mathers CD, Ezzati M, Jamiso DT, Murray CJ (2006) Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 367:1747–1757

Mantel N, Haenszel W (1959) Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 22:719–748

McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R et al (2007) A common allele on chromosome 9 associated with coronary heart disease. Science 316:1488–1491

Morgan TM, Krumholz HM, Lifton RP, Spertus JA (2007) Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. JAMA 297:1551–1561

Myocardial Infarction Genetics Consortium (2009) Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 41:334–341

Ohashi J, Tokunaga K (2003) Power of genomewide linkage disequilibrium testing by using microsatellite markers. J Hum Genet 48:487–491

Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R et al (2002) Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet 32:650–654

Ozaki K, Sato H, Inoue K, Tsunoda T, Sakata Y et al (2009) SNPs in BRAP associated with risk of myocardial infarction in Asian populations. Nat Genet 41:329–333

Parmacek MS (2007) Myocardin-related transcription factors: critical coactivators regulating cardiovascular development and adaptation. Circ Res 100:633–644

Ross R (1999) Atherosclerosis—an inflammatory disease. N Engl J Med 340:115–126

Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M et al (2007) Genomewide association analysis of coronary artery disease. N Engl J Med 357:443–453

Sata M, Saiura A, Kunisato A, Tojo A, Okada S et al (2002) Hematopoietic stem cells differentiate into vascular cells that participate in the pathogenesis of atherosclerosis. Nat Med 8:403–409

Shibata H, Yasunami M, Obuchi N, Takahashi M, Kobayashi Y et al (2006) Direct determination of single nucleotide polymorphism haplotype of NFKBIL1 promoter polymorphism by DNA conformation analysis and its application to association study of chronic inflammatory diseases. Hum Immunol 67:363–373

Shiffman D, Ellis SG, Rowland CM, Malloy MJ, Luke MM et al (2005) Identification of four gene variants associated with myocardial infarction. Am J Hum Genet 77:596–605

Tamiya G, Shinya M, Imanishi T, Ikuta T, Makino S et al (2005) Whole genome association study of rheumatoid arthritis using 27 039 microsatellites. Hum Mol Genet 14:2305–2321

Trégouët DA, König IR, Erdmann J, Munteanu A, Braund PS, Hall AS et al (2009) Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. Nat Genet 41:283–285

Wang Q (2005) Molecular genetics of coronary artery disease. Curr Opin Cardiol 20:182–188

Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678

Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL et al (2008) Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet 40:161–169