occasionally presents [23]. Three genes, *KIRIT1* (*CCM1*) [13] at 7q21.2, *MGC4607* (*CCM2*) [5] at 7p13 and *PDCD10* (*CCM3*) [2] at 3q26.1, are responsible for cerebral cavernous malformation (hamartomatous vascular malformations). On the other hand, regarding familial AVM, only two linkage analyses using 6 small families have been published by a research group [11,25], showing seven possible disease-responsible regions, i.e., 6q25 with the highest LOD score, 3p27, 4q34, 7p21, 13q32–q33, 16p13–q12 and 20q11–q13, but failed to identify the causative mutation. In sporadic brain AVM, microarray study showed that the *VEGFA*, *ITGA5*, *ENG* and *MMP9* genes that may involve vascular development or maintenance, are highly expressed in AVM compared with normal brain parenchyma [8,22,24].

Here we report results of a genome-wide linkage analysis on an AVM family with four affected members in two successive generations.

2. Materials and methods

2.1. Subjects

A Japanese family consisting of 19 members across four generations included two patients with brain AVM, one patient with pulmonary AVM and one patient with both brain and pulmonary AVM (Fig. 1). The proband (III-3) first exhibited intractable epilepsy at 13 years old and was diagnosed by magnetic resonance imaging (MRI) as having a brain AVM of 2 cm in diameter located in the right frontal lobe (Fig. 2). Chest X-ray at the first visit detected a nodular shadow in the right lower lung field, and a diagnoses of pulmonary AVM with a 24% of shunt-rate was made following angiogram made (Fig. 2). This was resected when the proband was 14 years old. The proband's brain AVM was treated by gamma knife surgery when she was 19 years old, followed by treatment with antiepileptic medication. Her mother (II-3) died of intracranial hemorrhage due to brain AVM, and the maternal grandfather (I-1) died of a cancer. Another patient (III-5) had asymptomatic brain AVM, which was accidentally diagnosed by MRI. His father (II-5) had pulmonary AVM instead of brain AVM. These four members were assigned to "affected", six members (II-6, III-1, III-6, III-7, IV-1, and IV-2) without AVM confirmed by MRI were "unaffected", and the remaining three (I-2, II-1, and IV-3) who were not assessed by MRI but had neither past history of recurrent epistaxis or gastrointestinal tract bleedings were "unknown". None of the members had any AVM-related diseases, such as HHT. Evaluation of cutaneous

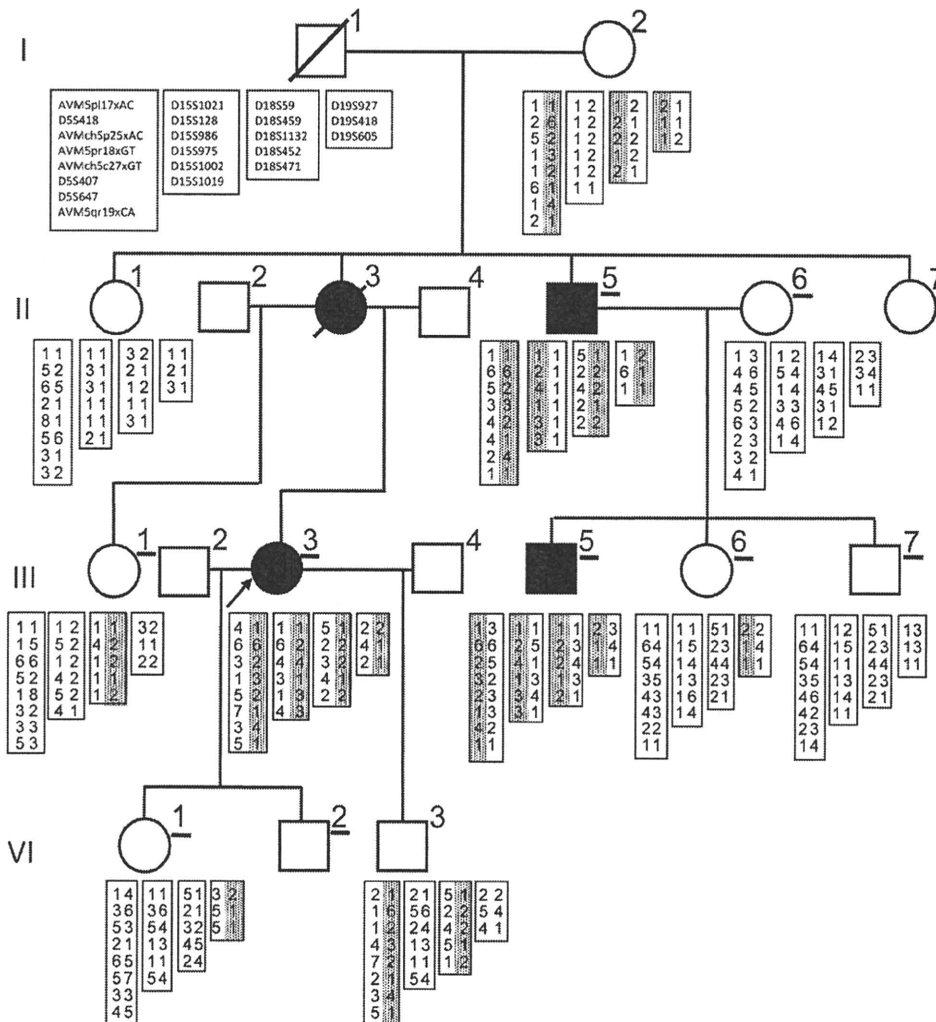

Fig. 1. Results of haplotype analysis at polymorphic loci in four regions, 5q13.2–q14.1, 15q11.2–q13.1, 18p11.32–p11.22 and 19q13.3–q13.42. Underlined individuals indicate those examined by MRI, and DNA was unavailable from individuals without haplotypes. Polymorphic alleles are numbered and candidate disease-associated haplotypes are shown by dotted boxes. Primer sequences designed for CA repeat amplification are available in Supplementary Table.

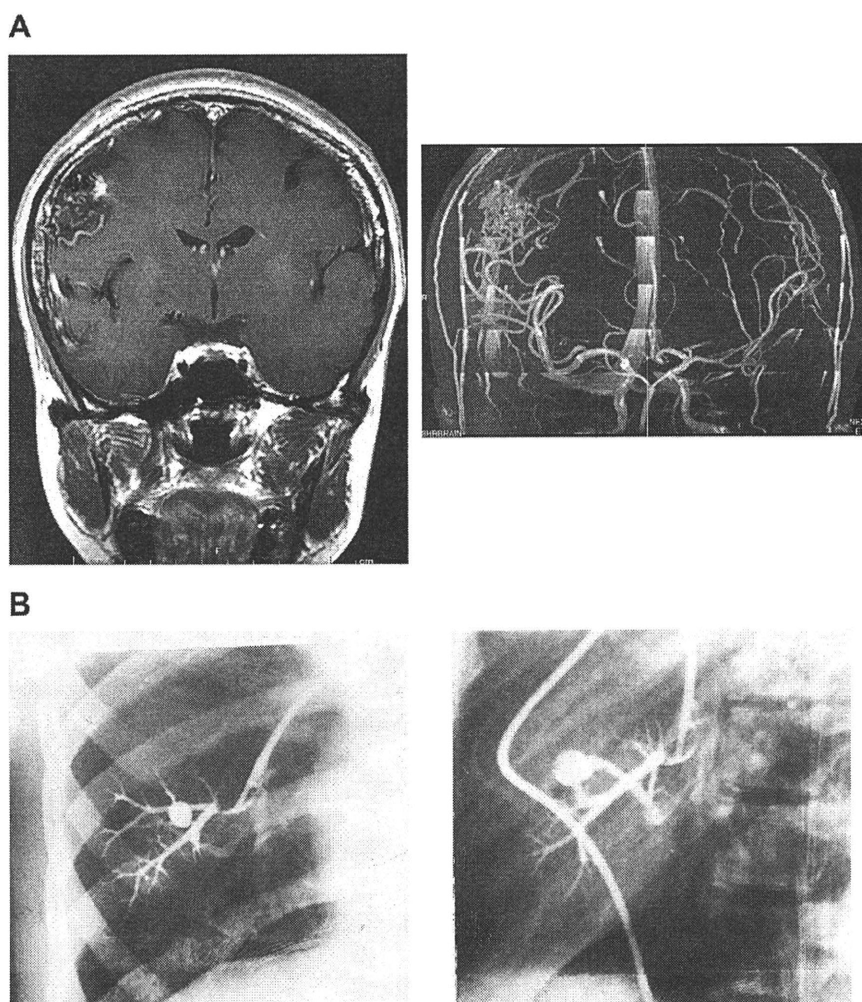

Fig. 2. Imaging of the brain and pulmonary AVM in the proband. (A) MRI scan and MR angiogram of the proband. The AVM is located right frontal lobe measured 2.0 × 1.3 cm. (B) Pulmonary angiograms of the proband. The pulmonary AVM is located in the right lower lobe (rtS8b) with 24% of shunt-rate.

lesions was conducted by examination of the proband and by detailed interview of the other family members by the proband and her sister (III-1), who is nurse. A total of 13 members participated in this study under informed consent. All experimental procedures for this study were approved by Committee for the Ethical Issues on Human Genome and Gene Analysis in Nagasaki University.

2.2. DNA extraction

As a blood sample was available only from the proband, clipped fingernail samples were obtained from 10 of the other 12 members instead. Genomic DNA was extracted from the fingernails using a buffer solution containing urea, DDT and proteinase K, as reported previously [16,18]. Briefly, clipped fingernails were once frozen in liquid nitrogen and crushed into fine powder using Multi-beads Shocker™ (Yasui Kikai, Osaka, Japan). The nail powder was lysed in a urea-lysis solution (2 M urea; 0.5% SDS; 10 mM Tris–HCl, pH 7.5; 0.1 M EDTA) containing 1 mg/ml proteinase K and 40 mM DDT at 55 °C overnight. Nail DNA was extracted with phenol/chloroform, and precipitated with ethanol and sodium acetate. Precipitated nail DNA was dissolved again in extraction buffer (0.5% SDS; 10 mM Tris–HCl, pH 7.5; 0.1 M EDTA) containing 1 mg/ml proteinase K, and incubated at 55 °C overnight. DNA was purified again as above, and was suspended in 30 μl of 1× TE buffer.

2.3. SNP genotyping with Affymetrix 10K 2.0 array

Blood DNA (250 ng) was processed according to the standard protocol provided by the GeneChip Mapping 10K Xba Assay Kit (Affymetrix, Santa Clara, CA). Fingernail DNA was processed in a similar manner but with the two following modifications to adapt to the oligonucleotide microarray system [15]. Prolongation of digestion time from 120 min as the standard protocol to overnight; and increase of the PCR cycle number from 35 to 45 cycles. Data acquired from the Affymetrix GeneChip Operating System were analyzed using the Affymetrix GeneChip Genotyping Analysis Software (GTYPE) 4.0 to call genotypes.

2.4. Linkage analysis with SNP-genotype data and haplotype analysis with microsatellite markers

Multipoint LOD scores were calculated using MERLIN software [1], under an assumption that AVM in the family is transmitted in an autosomal dominant mode with reduced penetrance ($p = 0.9$)

and with the disease allele frequency of 0.001. At loci with a positive LOD score by the GeneChip genotyping, possibly disease-associated haplotypes were constructed using SNP calls.

When SNP information was not informative, microsatellite markers were used for genotyping. Microsatellite markers used were referred to the National Center for Biotechnology Information (NCBI) database. One each of primer pairs for the markers was labeled with FAM, HEX, or NED (Supplementary Table 1), and PCR was performed in a 10 µl mixture containing 5 ng genomic DNA; 0.25 U ExTaq DNA polymerase HS-version (TAKARA Bio Inc., Kyoto, Japan); 200 µM dNTP; 0.5 µM primer; 1× ExTaq buffer on the T1 Thermocycler (Biometra, Goettingen, Germany). PCR products were separated on Genetic Analyzer 3130xl (AppliedBiosystems), and genotyping was carried out using GeneMapper software (AppliedBiosystems). At the regions where the affected individuals have a disease-associated haplotype, two-point LOD score was calculated by MLINK program (included in FASTLINK software version 4.0P) [14].

2.5. Mutation analysis

Some genes located within candidate regions identified by the linkage analysis were selected for further mutation analysis. A few other genes, albeit outside the regions, were also subjected to mutation analysis. Primer pairs for such genes were designed using Primer3-web 0.3.0 (http://frodo.wi.mit.edu/primer3/input.htm), according to their genomic sequences retrieved from the University of California, Santa Cruz (UCSC) Genome Browser Home (http://genome.cse.ucsc.edu/). PCR was carried out in a 15 µl reaction mixture containing 5 ng DNA; 0.25 U ExTaq DNA polymerase HS version; 200 mM dNTP; 0.5 µM each primer; 1× ExTaq buffer on the T1 Thermocycler. PCR products were subjected to direct sequencing, using BigDye Terminator v3.1 Cycle sequencing Kit (AppliedBiosystems) and Genetic Analyzer 3130xl. Electropherograms of sequences were aligned with ATGC software (GENETYX Corp., Tokyo, Japan) to inspect base alterations.

2.6. Search for genomic aberration

To search for copy number change within the candidate loci identified by linkage analysis, we used Affymetrix® Genome-Wide Human SNP Array 5.0 (920,568 probes; Affymetrix). Genomic DNA extracted from white blood cell of proband was processed according to manufacture's protocol. Intensity data from each probes were obtained from Affymetrix® Genotyping Console 3.0 as a CEL files. Unpaired copy number analysis of whole genome was carried out using Partek Genomics Suite (Partek, MO, USA) and regions with copy number change were determined by Hidden Markov Model at default settings.

3. Results

3.1. Linkage and haplotype analyses

The mean SNP call rate was 92.49% in 11 fingernail DNA samples, compared to 98.11% in a blood DNA sample from the proband. Incorrect SNP calls may result in seemingly inconsistent parent–child transmissions, but the call rates obtained are actually enough for further studies. We thus advanced to calculate LOD scores using these data.

The linkage analysis using MERLIN software revealed 18 regions with positive LOD scores (>0.00). Of the 18 regions, 14 with the following conditions were excluded: those without any functional full-length RefSeq genes; those in small size (<200 kb); and those in which some affected members did not have a common haplotype. Consequently, four loci, 5p13.2–q14.1, 15q11.2–q13.1, 18p11.32–p11.22 and 19q13.33–q13.42, remained as possibly linked regions (Figs. 1 and 3).

We then genotyped with microsatellite markers and calculated two-point LOD scores, considering the affected, unaffected, and the unknown family members. We confirmed three of the four candidate loci. They were a 48-Mb region between markers rs1366265 and rs1373965 at 5p13.2–q14.1, a 6-Mb region between rs850819

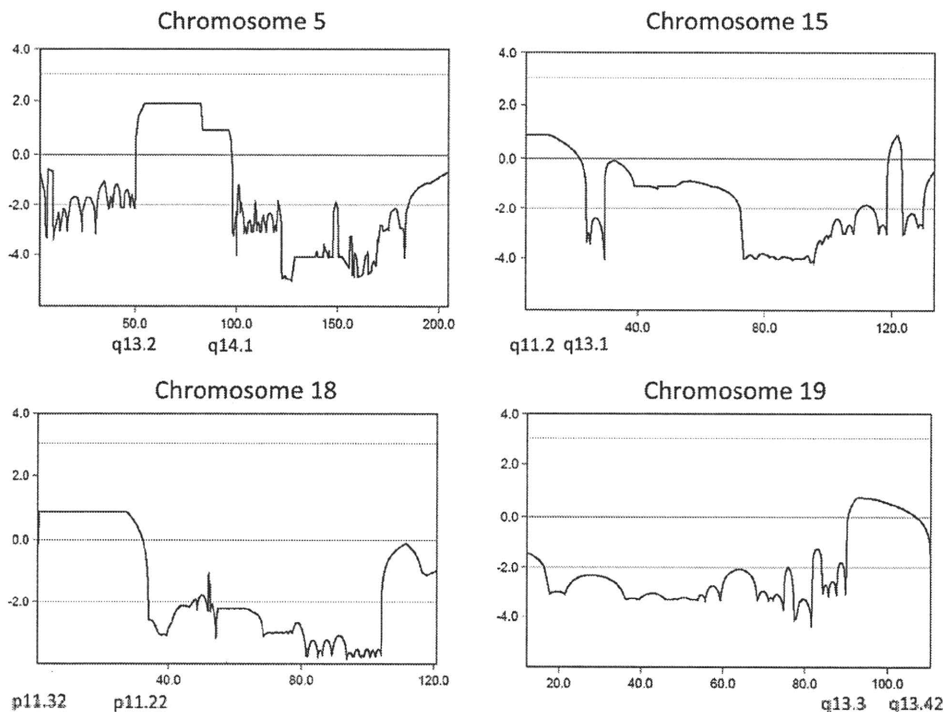

Fig. 3. Multipoint LOD scores calculated by MERLIN in four chromosomal regions, 5q13.2–q14.1, 15q11.2–q13.1, 18p11.32–p11.22 and 19q13.3–q13.42.

and rs818089 at 15q11.2–q13.1, both giving the maximum two-point LOD score of 1.632 ($\theta = 0$), and a 9-Mb region between rs486633 and rs1942150 at 18p11.32–p11.22 with the maximum LOD score of 0.851 ($\theta = 0$) (Table 1). As a possibly disease-associated haplotype on 19q13.33–q13.42 was transmitted to two definitively unaffected individuals (III-6 and IV-1), chromosome 19 was ruled out from the candidacy (Table 1, Fig. 1).

3.2. Mutation analysis of candidate genes

Within the 48-Mb region at 5p13.2–q14.1, there are about 200 RefSeq genes. Ten (*MAP3K1*, *DAB2*, *OCLN*, *FGF10*, *ESM1*, *ITGA1*, *ITGA2*, *EDFLAM*, *ERBB2IP*, and *PIK3R1*) from these genes were focused and selected as candidates for brain AVM, since they concern development or maintenance of vessels, are associated with other heritable vascular disorders such as HHT, or are expressed in the brain with AVM [8,22,24]. Mutation analyses in these 10 genes revealed no pathologic mutation in the proband, although other affected members were not examined because of insufficient amount of their DNA. Although the genes endoglin isoform 1 precursor (*ENG*), activin A receptor type II like 1 (*ALK1*) and RAS p21 protein activator 1 (*RASA1*) are not located in the candidate region, we investigated whether any of them are involved in the etiology of AVM in the family as a partial symptom of HHT or AVM-CM. Direct sequencing of these three genes failed to show any causative variants.

Copy number analysis of proband revealed one increased copy number loci at 12q and decreased at 2p, 3q, 4q, 6p, 7q and 22q (data not shown). But all these alterations were reported previously as copy number polymorphisms (http://projects.tcag.ca/variation/) and out of our candidate loci. In addition, neither deletions nor microdeletions were detected at 9q34.11 of *ENG*, 12q13.13 of *ALK1* and 5q14.3 of *RASA1*.

4. Discussion

We have reported a family consisting of two affected members with brain AVM, one with pulmonary AVM and one with both brain and pulmonary AVM. The condition in this family met the criteria of familial brain AVM and seems to be inherited in an autosomal dominant mode. We tried to assign the location of a putative disease-gene by linkage analysis and search for mutations by subsequent candidate gene approach.

Table 1
Two-point LOD scores for brain AVM at various loci.

Locus	Recombination fraction (θ)					
	0.00	0.01	0.02	0.03	0.04	0.05
AVM5pl17xAC	0.032	0.030	0.029	0.027	0.026	0.025
D5S418	0.551	0.535	0.518	0.501	0.484	0.467
AVMch5p25xAC	1.334	1.301	1.268	1.234	1.201	1.167
AVM5pr18xGT	0.511	0.491	0.472	0.452	0.433	0.414
AVMch5c27xGT	1.630	1.597	1.564	1.531	1.497	1.463
AVM5c18xAC	1.373	1.344	1.314	1.285	1.255	1.225
D5S407	1.632	1.599	1.566	1.532	1.499	1.465
D5S647	1.154	1.121	1.089	1.056	1.023	0.991
AVM5qr19xCA	0.810	0.790	0.769	0.748	0.727	0.706
D15S1021	0.171	0.164	0.157	0.150	0.143	0.137
D15S128	0.876	0.858	0.841	0.823	0.805	0.787
D15S986	0.812	0.791	0.770	0.749	0.728	0.707
D15S975	0.400	0.387	0.374	0.361	0.348	0.335
D15S1002	1.330	1.298	1.266	1.234	1.202	1.170
D15S1019	1.632	1.599	1.566	1.532	1.499	1.465
D18S59	0.199	0.214	0.225	0.234	0.241	0.246
D18S459	0.142	0.136	0.131	0.125	0.120	0.114
D18S1132	0.677	0.663	0.650	0.636	0.623	0.609
D18S452	0.851	0.832	0.813	0.794	0.774	0.755
D18S471	0.240	0.231	0.222	0.214	0.205	0.197
D19S927	−0.302	−0.277	−0.254	−0.234	−0.216	−0.200
D19S418	−2.655	−2.453	−2.257	−2.078	−1.919	−1.778
D19S605	−0.648	−0.574	−0.512	−0.460	−0.414	−0.374

The linkage analysis of the family revealed three candidate regions (5p13.2–q14.1, 15q11.2–q13.1, and 18p11.32–p11.22) with relatively high LOD scores of 1.632, 1.632 and 0.851, respectively (Table 1). However, neither region was conclusive. This insufficient mapping may have arisen from the small pedigree size, and/or from incomplete ascertainment of affected members, e.g., probable existence of asymptomatic affected persons among the "unknown" members. Indeed, as for a candidate locus at 5p13.2–q14.1, the proband's maternal grandmother (I-2) and son (IV-3) had a haplotype common to the three affected members (Fig. 1), but they were fallen into the "unknown" individuals. If DNA from IV-2 was available and if MRI examinations of VI-3 and I-2 were carried out, we would have obtained more definitive results. As we performed linkage analysis using high-density SNP genotyping, 14 small regions not containing RefSeq genes or miRNAs showed a positive LOD score. It is possible that an unidentified transcribed RNA in one of these regions could cause familial AVM, but these regions are candidate loci with a lower priority than those containing known genes. Thus, the three regions have remained at present as the equally possible loci for AVM. The three regions do not overlap with a previously reported candidate locus of familial brain AVM, i.e., 6p25 [11], and do not contain genes responsible for syndromic AVM (heritable disorders involving AVM) or cerebral cavernous malformations, such as *ENG* [17], *ALK1* [12], *RASA1* [3,6,20,21,26], and *PTEN* [23], *KRIT1* [13], *MGC407* [5], *PDCD10* [2].

We then searched for mutations in 10 genes within 5p13.2–q14.1, among which *MAP3K1*, *DAB2* and *OCLN* encode proteins playing roles in the TGF-β signaling pathway, and *FGF10*, *ESM1*, *ITGA1*, *ITGA2*, *EGFLAM*, *ERBB2IP* and *PIK3R1* were those expressed in brain AVM tissues by previous microarray analysis [8,22,24]. Nevertheless, no pathologic mutation was found in any of them. Because the presence of both brain AVM and pulmonary AVM in this pedigree is reminiscent of Hereditary Hemorrhagic Telangiectasia, we analyzed *ENG* and *ALK1* for mutations and genomic aberrations, which may cause HHT1 and HHT2 respectively [12,17]. The proband did not have any mutations in the coding exons or intron/exon boundaries of either gene, nor any genomic aberrations at those loci. We also analyzed *RASA1* because this may cause CM-AVM, which is characterized by multiple CM and AVM [3,6,20,21,26]. No causative mutation or genomic aberration was detected in the proband. Although other genes, such as *KRIT1*, *MGC407* and *PDCD10*, have been shown to cause slow-flow lesions i.e., cerebral cavernous malformation [2,5,13], they were not investigated in the present study, because the clinical manifestations in our family did not meet the criteria for these diseases.

Participation of family members and compliance with guidelines for human genome researches are critical to conduct a linkage analysis. Whole-blood samples cannot occasionally be available in some family members because of their far domicile. In such the case, fingernail DNA is useful, since clipped fingernails can be mailed in a usual way, and stored long at a room temperature, as indicated previously [16,19]. The present study is the first experience to adopt fingernail DNA to genome-wide high-density SNP microarray analysis. The performance obtained from fingernail DNA was sufficient, showing all SNP call rates of >86%. According to the manufacture's protocol, samples with an SNP call rate of <85% should further be evaluated before including the data in downstream analysis. Incorrect SNP calls may make serious problems in linkage analysis. For instance, SNPs with parent–child transmission inconsistency may be omitted, leading to a reduced LOD score.

In conclusion, we have assigned the familial AVM locus to three alternative regions, 5p13.2–q13.2, 15q11.2–q13.1 and 18p11.32–p11.22, by a genome-wide, high-density, SNP-based

linkage analysis with fingernail DNA in an AVM family. However, mutation analyses of some genes in the regions failed to identify any pathological changes.

Funding

This research was supported in part by Grants-in-Aid for Scientific Research (Nos. 17019055 and 19390095) from the Ministry of Education, Sports, Culture, Science and Technology of Japan, and by SORST from Japan Science and Technology Agency (JST). K.Y. was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare of Japan.

Competing interests

There are no competing interests.

Acknowledgments

We are grateful to the family members for their participation in this research. We also thank Ms Yasuko Noguchi, Ms Miho Ooga, and Ms Chisa Hayashida for their technical assistance.

Appendix. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.ejmg.2010.06.007.

References

[1] G.R. Abecasis, S.S. Cherny, W.O. Cookson, L.R. Cardon, Merlin–rapid analysis of dense genetic maps using sparse gene flow trees, Nat. Genet. 30 (2002) 97–101.

[2] F. Bergametti, C. Denier, P. Labauge, M. Arnoult, S. Boetto, M. Clanet, P. Coubes, B. Echenne, R. Ibrahim, B. Irthum, G. Jacquet, M. Lonjon, J.J. Moreau, J.P. Neau, F. Parker, M. Tremoulet, E. Tournier-Lasserve, Société Française de Neurochirurgie, Mutations within the programmed cell death 10 gene cause cerebral cavernous malformations, Am. J. Hum. Genet. 76 (2005) 42–51.

[3] L.M. Boon, J.B. Mulliken, M. Vikkula, RASA1: variable phenotype with capillary and arteriovenous malformation, Curr. Opin. Genet. Dev. 15 (2005) 265–269.

[4] P.M. Crawford, C.R. West, D.W. Chadwick, M.D.M. Shaw, Arteriovenous malformations of the brain: natural history in unoperated patients, J. Neurol. Neurosurg. Psychiatr. 49 (1986) 1–10.

[5] C. Denier, S. Goutagny, P. Labauge, V. Krivosic, M. Arnoult, A. Cousin, A.L. Benabid, J. Comoy, P. Frerebeau, B. Gilbert, J.P. Houtteville, M. Jan, F. Lapierre, H. Loiseau, P. Menei, P. Mercier, J.J. Moreau, A. Nivelon-Chevallier, F. Parker, A.M. Redondo, J.M. Scarabin, M. Tremoulet, M. Zerah, J. Maciazek, E. Tournier-Lasserve, Société Française de Neurochirurgie, Mutations within the MGC4607 gene cause cerebral cavernous malformations, Am. J. Hum. Genet. 74 (2004) 326–337.

[6] I. Eerola, L.M. Boon, J.B. Mulliken, P.E. Burrows, A. Dompmartin, S. Watanabe, R. Vanwijck, M. Vikkula, Capillary malformation-arteriovenous malformation, a new clinical and genetic disorder caused by RASA1 mutations, Am. J. Hum. Genet. 73 (2003) 1240–1249.

[7] I.G. Fleetwood, G.K. Steinberg, Arteriovenous malformations, Lancet 359 (2002) 863–873.

[8] T. Hashimoto, M.T. Lawton, G. Wen, G.Y. Yang, T. Chaly, C.L. Stewart, H.K. Dressman, N.M. Barbaro, D.A. Marchuk, W.L. Young, Gene microarray analysis of human brain arteriovenous malformations, Neurosurgery 54 (2004) 410–423 discussion 423–425.

[9] R. Herzig, S. Burval, V. Vladyke, L. Janouskova, P. Krivanek, B. Krupka, I. Vlachova, K. Urbanek, Familial occurrence of cerebral arteriovenous malformation in sisters: case report and review of the literature, Eur. J. Neurol. 7 (2000) 95–100.

[10] J.C. Horton, Arteriovenous malformations of the brain, N. Engl. J. Med. 357 (2007) 1774–1775.

[11] S. Inoue, W. Liu, K. Inoue, M. Youhei, K. Takenaka, H. Yamakawa, M. Abe, J.J. Jafar, R. Herzig, A. Koizumi, Combination of linkage and association studies for brain arteriovenous malformation, Stroke 38 (2007) 1368–1370.

[12] D.W. Johnson, J.N. Berg, M.A. Baldwin, C.J. Gallione, I. Marondel, S.-J. Yoon, T.T. Stenzel, M. Speer, M.A. Pericak-Vance, A. Diamond, A.E. Guttmacher, C.E. Jackson, L. Attisano, R. Kucherlapati, M.E.M. Porteous, D.A. Marchuk, Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2, Nat. Genet. 13 (1996) 189–195.

[13] S. Laberge-le Couteulx, H.H. Jung, P. Labauge, J.-P. Houtteville, C. Lescoat, C. Michaelle, E. Marechal, A. Joutel, J.-F. Bach, E. Tournier-Lasserve, Truncating mutations in CCM1, encoding KRIT1, cause hereditary cavernous angiomas, Nat. Genet. 23 (1999) 189–193.

[14] G.M. Lathrop, J.M. Lalouel, C. Julier, J. Ott, Strategies for multilocus linkage analysis in humans, Proc. Natl. Acad. Sci. U S A 81 (1984) 3443–3446.

[15] M. Lyons-Weiler, J. Hagenkord, C. Sciulli, R. Dhir, F.A. Monzon, Optimization of the Affymetrix GeneChip mapping 10K 2.0 Assay for routine clinical use on formalin-fixed paraffin-embedded tissues, Diagn. Mol. Pathol. 17 (2008) 3–13.

[16] N. Matsuzawa, K. Shimozato, N. Natsume, N. Niikawa, K. Yoshiura, A novel missense mutation in Van der Woude syndrome: usefulness of fingernail DNA for genetic analysis, J. Dent. Res. 85 (2006) 1143–1146.

[17] K.A. McAllister, K.M. Grogg, D.W. Johnson, C.J. Gallione, M.A. Baldwin, C.E. Jackson, E.A. Helmbold, D.S. Markel, W.C. McKinnon, J. Murrell, M.K. McCormick, M.A. Pericak-Vance, P. Heutink, B.A. Oostra, T. Haitjema, C.J.J. Westerman, M.E. Porteous, A.E. Guttmacher, M. Letarte, D.A. Marchuk, Endoglin, a TGF-beta binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1, Nat. Genet. 8 (1994) 345–351.

[18] M. Nakashima, M. Nakano, A. Hirano, T. Kishino, S. Kondoh, N. Miwa, N. Niikawa, K. Yoshiura, Genome-wide linkage analysis and mutation analysis of hereditary congenital blepharoptosis in a Japanese family, J. Hum. Genet. 53 (2008) 34–41.

[19] A.M. Oberbauer, D.I. Grossman, M.L. Eggleston, D.N. Irion, A.L. Schaffer, N.C. Pedersen, J.M. Belanger, Alternatives to blood as a source of DNA for large-scale scanning studies of canine genome linkages, Vet. Res. Commun. 27 (2003) 27–38.

[20] N. Revencu, L.M. Boon, J.B. Mulliken, O. Enjolras, M.R. Cordisco, P.E. Burrows, P. Clapuyt, F. Hammer, J. Dubois, E. Baselga, F. Brancati, R. Carder, J.M.C. Quintal, B. Dallapiccola, G. Fischer, I.J. Frieden, M. Garzon, J. Harper, J. Johnson-Patel, C. Labreze, L. Martorell, H.J. Paltiel, A. Pohl, J. Prendiville, I. Quere, D.H. Siegel, E.M. Valente, A. van Hagen, L. van Hest, K.K. Vaux, A. Vicente, L. Weibel, D. Chitayat, M. Vikkula, Parkes Weber syndrome, vein of Galen aneurysmal malformation, and other fast-flow vascular anomalies are caused by RASA1 mutations, Hum. Mutat. 29 (2008) 959–965.

[21] N. Revencu, L.M. Boon, M. Vikkula, From germline towards somatic mutations in the pathophysiology of vascular anomalies, Hum. Mol. Genet. 18 (R1) (2009) R65–R74.

[22] A. Sasahara, H. Kasuya, H. Akagawa, H. Ujiie, O. Kubo, T. Sasaki, H. Onda, Y. Sakamoto, B. Krischek, T. Hori, I. Inoue, Increased expression of ephrin A1 in brain arteriovenous malformation: DNA microarray analysis, Neurosurg. Rev. 30 (2007) 299–305 discussion 305.

[23] K. Suphapeetiporn, P. Kongkam, J. Tantivatana, T. Sinthuwiwat, S. Tongkobpetch, V. Shotelersuk, PTEN c.511C>T nonsense mutation in a BRRS family disrupts a potential exonic splicing enhancer and causes exon skipping, Jpn. J. Clin. Oncol. 36 (2006) 814–821.

[24] U. Sure, E. Battenberg, A. Dempfle, W. Tirakotai, S. Bien, H. Bertalanffy, Hypoxia-inducible factor and vascular endothelial growth factor are expressed more frequently in embolized than in nonembolized cerebral arteriovenous malformations, Neurosurgery 55 (2004) 663–669.

[25] K. Takenaka, J. Muroi, S. Yamada, H. Yamakawa, M. Abe, K. Tabuchi, A. Koizumi, Genetic dissection of the familial cerebral arteriovenous malformation, Jpn. J. Neurosurg. 13 (2004) 837–845.

[26] R. Thiex, J.B. Mulliken, N. Revencu, L.M. Boon, P.E. Burrows, M. Cordisco, Y. Dwight, E.R. Smith, M. Vikkula, D.B. Orbach, A novel association between RASA1 mutations and spinal rteriovenous anomalies, AJNR. Am. J. Neuroradiol. 31 (2010) 775–779.

[27] J. van Beijnum, H.B. van der Worp, H.M. Schippers, O. van Nieuwenhuizen, L.J. Kappelle, G.J.E. Rinkel, J.W.B. van der Sprenkel, C.J.M. Klijn, Familial occurrence of brain arteriovenous malformations: a systematic review, J. Neurol. Neurosurg. Psychiatr. 78 (2007) 1213–1217.

Failure to Confirm CNVs as of Aetiological Significance in Twin Pairs Discordant for Schizophrenia

Shinji Ono,[1,2] Akira Imamura,[2] Shinya Tasaki,[2] Naohiro Kurotaki,[2] Hiroki Ozawa,[2] Koh-ichiro Yoshiura,[1] and Yuji Okazaki[3]

[1] *Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Japan*

[2] *Department of Neuropsychiatry, Nagasaki University Graduate School of Biomedical Sciences, Japan*

[3] *Tokyo Metropolitan Matsuzawa Hospital, Tokyo, Japan*

Copy number variations (CNVs) are common structural variations in the human genome that strongly affect genomic diversity and can play a role in the development of several diseases, including neurodevelopmental disorders. Recent reports indicate that monozygotic twins can show differential CNV profiles. We searched CNVs in monozygotic twins discordant for schizophrenia to identify susceptible loci for schizophrenia. Three pairs of monozygotic twins discordant for schizophrenia were subjected to analysis. Genomic DNA samples were extracted from peripheral blood lymphocytes. We adopted the Affymetrix Genome-Wide Human SNP (Single Nucleotide Polymorphism) Array 6.0 to detect copy number discordance using Partek Genomics Suite 6.5 beta. In three twin pairs, however, validations by quantitative PCR and DNA sequencing revealed that none of the regions had any discordance between the three twin pairs. Our results support the hypothesis that epigenetic changes or fluctuation in developmental process triggered by environmental factors mainly contribute to the pathogenesis of schizophrenia. Schizophrenia caused by strong genetics factors including copy number alteration or gene mutation may be a small subset of the clinical population.

Keywords: CNVs, schizophrenia, genotype, monozygotic twin, epigenetic change

Schizophrenia is a chronic, debilitating psychiatric illness with a 1% worldwide prevalence. Genetic studies have shown that schizophrenia has a high heritability, with strong genetic factors involved in its etiology. Twin studies have played an important role in the elucidation of the genetic factors underlying neurodevelopmental disorders. Several twin studies have revealed that the concordance rate between monozygotic twins is 41–79% for schizophrenia, whereas the concordance rate between dizygotic twins is 0–14% (Shih et al., 2004; Kakiuchi et al., 2008). The higher concordance rate in monozygotic rather than dizygotic twins for schizophrenia suggests the contribution of genetic factors. Phenotypically discordant monozygotic twins are especially interesting resources for genetic studies, and twin studies could facilitate the identification of the causative genes of phenotypes. Kondo et al. (2002) reported that a nonsense mutation in *IRF6*, which is a causative gene for Van der Woude syndrome, was found in one affected individual in monozygotic phenotypically discordant twins. In relation to neurodevelopmental disorders, Bruder et al. (2008) reported that discordant monozygotic twins with parkinsonism showed different copy number variation (CNV) profiles.

CNVs are the most prevalent type of structural variations in the human genome that largely contribute to genomic diversity. Redon et al. (2006) and Carter et al. (2007) showed that as much as 12% of the human genome and thousands of genes are variable in copy number. A great number of CNVs may not be pathogenic but simply contribute to human genome diversity not related to phenotype. Meanwhile, some CNVs have been proven a significant factor related to disease susceptibility. Recent studies reported that CNVs contribute to genetic vulnerability factors and can play an important role in the etiology of several neurodevelopmental disorders (Sebat et al., 2007; 2009). Xu et al. (2008) found that de novo copy number mutations were about eight times more frequent in patients with sporadic schizophrenia. Numerous copy number analyses in schizophrenia revealed that genes that were disrupted by CNVs, which include *TBX1*, *ERBB4*, *SLC1A3*, *RAPGEF4*, *CIT*, *NRXN1*, and 16p11.2 region, were candidate genes and regions for schizophrenia (Cook et al., 2008; McCarthy et al., 2009; Merikangas et al., 2009; Walsh et al., 2008); however, most of these are rare copy number variants and the contribution of those genes to schizophrenia is restricted to a tiny part of etiologies.

Received 8 June, 2010; accepted 30 July, 2010.

Address for correspondence: Akira Imamura, MD, PhD, Sakamoto 1-7-1, Nagasaki 852-8501, Japan. E-mail: f1042@cc.nagasaki-u.ac.jp

To date, numerous causative genes for schizophrenia have been identified; however, because of genetic heterogeneity, there is still a long path to the elucidation of the pathogenesis of schizophrenia. To identify causative genes for schizophrenia, we have utilized the Affymetrix Genome-Wide Human SNP Array 6.0 in three pairs of monozygotic twins discordant for schizophrenia. Here, we describe the results of CNV and genotype profiles in three pairs of monozygotic twins.

Methods

Subjects

Three pairs of monozygotic twins discordant for schizophrenia participated in this study. Ten years had passed after the onset of schizophrenia in the affected individuals in all twin pairs. All of the twins were male, and their mean age was 53 years old. Two well-trained psychiatrists diagnosed the twins by structured clinical interview, and all affected individuals corresponded to the DSM-IV-TR criteria for the undifferentiated type of schizophrenia.

DNA Microarrays

Ten ml of peripheral blood samples was collected after obtaining written informed consent, and genomic DNA was extracted from blood lymphocytes using QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany). Experimental procedures were approved by the Committee for the Ethical Issues on Human Genome and Gene Analysis at Nagasaki University.

DNA microarray experiments were performed using Affymetrix Genome-Wide Human SNP Array 6.0 (SNP Array 6.0) (Affymetrix, Santa Clara, CA, USA). We performed a paired analysis for loss of heterozygosity (LOH) and an unpaired analysis for copy number analysis using control individuals' data. All of the computer analyses were performed using Genotyping Console (Affymetrix) and Genomics Suite version 6.5 beta software (Partek, St. Louis, MO, USA). Genomic copy number data were analyzed with Partek Genomics Suite software using a segmentation algorithm with stringent p value cutoff.

Quantification of Genome Copy Number

We performed real-time quantitative PCR using an intercalating dye, SYTO13 (Molecular Probes, Eugene, OR, USA), which is an alternative to SYBR green I, or using Universal Probe Library (Roche Diagnostics, Mannheim, Germany) to verify copy number changes suggested by the microarray analyses. Primers and probes were designed using the website software Universal ProbeLibrary Assay Design Center (https://www.roche-applied-science.com). Real-time PCR amplification was run on a LightCycler480 Real-Time PCR System (Roche Diagnostics, Mannheim, Germany). All samples were measured in tetraplicates.

DNA Sequencing

To verify the SNP genotypes obtained by SNP Array 6.0, we performed direct sequencing of PCR-amplified genomic DNA fragments including SNPs that showed discordant allele calls in each twin pair. The amplified fragments were directly sequenced after purification with exonuclease I and NTPhosTM Thermolabile Phosphatase (Epicentre, Madison, WI, USA) using the BigDye Terminator v3.1 Cycle Sequencing Kit and run on an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems). DNA sequences were analyzed using Variant Reporter (Applied Biosystems) and ATGC version 6.0 (Software Development, Tokyo, Japan).

Results

Microarray Analysis Results

Quality control (QC) data obtained from the SNP Array 6.0 are summarized in Table 1. The call rate and contrast QC in SNP Array 6.0 data were > 95% and > 1.50, respectively, for all samples, and both values indicated experiments using the SNP Array 6.0 were done well.

Copy number analysis of microarray data using Partek Genomics Suite showed some deleted or amplified regions in each twin pair (data not shown). Regions with discordant genotyping between twins from microarray data are summarized in Table 2.

Unpaired analysis of 6 individuals in comparison with ethnically-matched normal controls (HapMap-JPT) revealed that an approximately 3 kb region within the *SLC25A37* gene was deleted in two of the three schizophrenia twin pairs, 11A/B and 31A/B. The deleted region (chromosome 8:23460969–23463786) was not registered in the Database of Genomic Variants (http://projects.tcag.ca/variation/).

Quantitative PCR Results

We verified the copy number state by real-time PCR of the regions with discordant copy number, including genes, by paired analysis using SNP Array 6.0. Primers were designed for the middle position of the regions.

Table 1

Summary of Twin Samples and Affymetrix GeneChip Genotyping Results

Samples	Sex	Phenotype	SNP call rate	Contrast QC*
11A / B	Male	Schizophrenia/unaffected	99.444 / 99.516	2.38 / 2.48
21A / B	Male	Schizophrenia/unaffected	98.974 / 99.175	1.88 / 2.22
31A / B	Male	Schizophrenia/unaffected	99.199 / 99.179	2.26 / 1.60

Note: *Contrast QC (Quality Control) is per sample Quality Control test metric for SNP Array 6.0 intensity data. In high-quality data sets, the Contrast QC metric is higher than the 0.4 threshold according to user manual provided by the manufacturer.

Table 2

List of Loss of Heterozygosity Regions Derived from Microarray Data

chr. #	Physical position		Twin #	Validated SNPs	Overlapping genes
	Start	End			
1	4309356	4465925	11A / B	rs7521665, rs4654438	LOC284661
	45006976	45050681	31A / B	rs6676749	BEST4, PLK3, RPS8, SNORD38A, SNORD38B, SNORD46, SNORD55
	170792582	170870563	31A / B	rs2472550	Region overlaps with 70.55% of C1orf9
2	50182138	50311147	31A / B	rs1452762, rs6712119	Contained within NRXN1
	142093343	142097262	31A / B	rs355581	Contained within LRP1B
3	3693732	3821526	31A / B	rs7613060, rs769806	Region overlaps with 4.23% of LRRN1
	123371895	123393318	31A / B	rs1501900	Region overlaps with 37.81% of CASR
4	24093201	24162064	31A / B	rs4697063	Region overlaps with 34.68% of DHX15
	81368193	81418990	11A / B	rs10518238, rs1458046	Region overlaps with 24.07% of FGF5
	101451872	101646851	31A / B	rs3756037	Region overlaps with 57.10% of EMCN
	109080142	109167540	31A / B	rs4395588	Region overlaps with 15.93% of CYP2U1 and 42.51% of HADH
	126258785	126764905	31A / B	rs7660602	FAT4
5	38382422	38389445	11A / B	rs9292705	Contained within EGFLAM
	166816487	166823787	31A / B	rs17068499	Contained within ODZ2
6	35297977	35376388	31A / B	rs3800385	ZNF76, region overlaps with 3.59% of DEF6 and 36.49% of SCUBE3
	119363250	119468737	31A / B	rs6913082	Contained within FAM184A and 74.19% of FAM184A
8	86383578	87077669	31A / B	rs1845891, rs1553015, rs6605618	CA1, CA2, CA3, REXO1L1, REXO1L2P
9	207826	208183	31A / B	rs10964703	Contained within DOCK8
	3900136	3920251	31A / B	rs630219	Contained within GLIS3
	7154039	7156090	31A / B	rs1556100	Contained within KDM4C
	112777053	112781741	31A / B	rs3758281, rs16915618	Contained within LPAR1
10	68497020	68657339	31A / B	rs10822972	Contained within CTNNA3, region overlaps with 21.12% of LRRTM3
	100181485	100219522	31A / B	rs11599112	Region overlaps with 28.02% of HPSE2 and 39.99% of HPS1
11	8896463	9040536	11A / B	rs4929922	C11orf17, region overlaps with 29.17% of SCUBE2
	19449860	19466526	31A / B	rs11820210	Contained within NAV2
12	21894811	21895465	31A / B	rs4148673	Contained within ABCC9
	33716220	36801139	31A / B	rs11052835, rs2387324	ALG10
	38818800	39404433	21A / B	rs7132869	LRRK2, region overlaps with 5.43% of CNTN1
	63692809	63739310	11A / B	rs4964104	Region overlaps with 18.58% of WIF1
	69385261	69392041	31A / B	rs10879183	Contained within PTPRR
	77123022	77139445	31A / B	rs9971904	Region overlaps with 48.10% of NAV3
	120088239	120155175	31A / B	rs25643	Region overlaps with 29.88% of P2RX7 and 34.55% of P2RX4
13	102227169	102252370	31A / B	rs9514058	KDELC1, region overlaps with 11.79% of BIVM
16	13150832	13161027	31A / B	rs4781419	Contained within SHISA9
22	24570234	24607029	31A / B	rs6004793	Contained within MYO18B
	36847351	36893417	31A / B	rs2076116	Contained within PLA2G6

Note: Chr. # means the number of chromosome.

Quantitative PCR was performed for a total of 120 regions. However, we could not reconfirm the differences between twins in all 120 tested regions. In addition, quantitative PCR within the *SLC25A37* gene revealed no loss or gain of the genome in comparison with ethnically matched control individuals.

Sequencing Results

DNA sequencing was performed for a total of 37 regions surrounding SNPs that had shown discordant genotype calls from microarray analysis within twin pairs. We selected one or more SNP(s) called discordant genotype in each LOH region. Sequencing revealed all of the SNPs were concordant between twin pairs (data not shown).

Discussion

In this study, we analyzed genomic alterations, CNVs and genotypes, in three pairs of monozygotic twins discordant for schizophrenia. None of the regions of copy number difference between twins shown by SNP Array 6.0 were reverified by quantitative PCR, and none of the genotype discordance was reverified by sequencing. Additionally, no novel CNVs was detected in the identified CNVs between twins. To our knowledge, this is the first report verifying the data from high-density and high-resolution DNA microarrays by quantitative PCR and DNA sequencing. Our results indicate that genomic alterations including CNVs and gene mutations contribute minimally to etiologies of

schizophrenia. The large genome-wide study by The Wellcome Trust Case Control Consortium (WTCCC) revealed CNVs is not main cause of bipolar disorder, which is one of the neurodevelopmental disorders (WTCCC, 2010). This report have a different concept from our study because our study aimed to find copy number alteration as a single gene disorder, however, WTCCC report could not discover the CNV contributing to the bipolar disorder. Our results may support the hypothesis that epigenetic changes (Roth et al., 2009), which can influence the expression of genes without affecting the DNA sequence, mainly contribute to the pathogenesis of schizophrenia.

SNP Array 6.0 allows us to detect different genotype or copy number neutral LOH regions. In our twin comparison, copy number neutral LOH would indicate segmental uniparental disomy (UPD) in twin pairs. Actually, UPD of the paternal allele at 11p15 in the affected twin caused discordance for hemihypertrophy in monozygotic twins (West et al., 2003). Furthermore, recent studies revealed that UPD was associated with schizophrenia. UPD on chromosome 1 and 5q32-qter in a patient with schizophrenia has been described in 2004 and 2006, respectively (Abecasis et al., 2004; Seal et al., 2006). But no genotype difference between twins was confirmed in this study.

Here, we presented no genomic discordance between twins; hereinafter, we will discuss some speculation about the relation between genetic factors and phenotypic discordance. First, it is possible that mosaicism at specific tissues (i.e., brain) because of postzygotic genomic rearrangements causes discordant phenotypes between monozygotic twins. Although we used DNA samples extracted from peripheral blood cells in this study, mosaic genomic rearrangement could be detected in brain. It is clear that the ideal source for studies of neurodevelopmental disorders is brain tissue. Nonetheless, it is practically impossible to harvest the brain tissue of twins (Kato et al., 2005). Olfactory sensory neurons have recently been shown to be easily accessible neuronal cells that have been useful for studies on schizophrenia (Arnold et al., 2001), enabling the study of neuronal cell character including genotype and copy number state. Second, it is possible that smaller-scale genomic aberrations below detection sensitivity influence the discordant phenotype of monozygotic twins. SNP Array 6.0 is one of the highest resolution platforms commercially available and allows us to identify CNVs much smaller than 10 kb. However, McCarroll et al. (2008) showed that the detection rate using the SNP Array 6.0 sharply diminished for CNVs <4 kb. To increase sensitivity, the use of many more detection probes is needed, and more than one experimental platform should be performed in future studies.

Bruder et al. (2009) successfully detected many copy number changes in peripheral blood using a Bacterial Artificial Chromosome (BAC)-array at mosaic state (~20%) in nine monozygotic twins discordant for parkinsonism. All of the nine pairs had copy number discordancy in their reports. Because their results suggested copy number change could be found in the mosaic state, tissue-specific mosaicism is a possible explanation for psychiatric disorders. We may have overlooked copy number change in a mosaic state in peripheral blood with the use of the SNP Array 6.0 instead of the BAC-array because the SNP Array 6.0 is a powerful tool to identify small regions with copy number change but is not suitable to detect copy number in a mosaic state.

It seems most likely that epigenetic changes between monozygotic twins influence the phenotypic discordance of monozygotic twins. Several recent studies indicate that epigenetic changes contribute to the etiology of schizophrenia. Rett syndrome and Fragile X syndrome are neurodevelopmental disorders caused by a single gene defect and dysregulation of DNA methylation very early in life (Amir et al., 1999; Das et al., 1997). Kaminsky et al. (2009) have shown that differences in DNA methylation profiles increase in monozygotic twins along with aging. Because the onset of schizophrenia is later than Rett syndrome and Fragile X syndrome, it is possible that cumulative epigenetic modifications could be one cause of schizophrenia development. Furthermore, a recent study by Roth et al. (2009) suggested that DNA methylation and histone modification triggered by influence of environmental factors is responsible for the difference in onset age between these disorders. Akbarian et al. (2005) indicated that histone modification contributes to the pathogenesis of prefrontal dysfunction in patients with schizophrenia based on the finding that the level of H3-(methyl)arginine 17 in patients with schizophrenia exceeded control values by 30%. Thus, genome-wide DNA methylation and genome-wide histone modification studies for monozygotic twins discordant for phenotypes may be promising techniques in future twin studies. In fact, Baranzini et al. (2010) reported genomic sequence and epigenetics (methyl-cytosine) analysis of monozygotic twin discordant for multiple sclerosis using next generation sequencer. They could not find reproducible differences between twins, but these comprehensive analyses including genome and epigenome sequence are just started. As Crow (2002) discussed, it is important to analyze the genetic and epigenetic influence to the species-specific characteristics. Comprehensive genetic and epigenetic analysis of discordant monozygotic twins will be advanced using next generation technologies.

In summary, we did not detect genomic alterations including CNVs and gene mutations between twins discordant for phenotype. Our results indicate that schizophrenia caused by genomic alterations may be a small subset of the clinical population and may support the hypothesis that epigenetic mechanisms triggered by the influence of environmental factors are associated with the etiology of schizophrenia. Experimental investigations of epigenetic mechanisms

such as expression analysis, methylation site sequence and histone modification studies using DNA samples extracted from olfactory sensory neurons are needed to identify the differences responsible for discordant phenotypes in future studies.

Acknowledgments

We are grateful to the subjects and their families for their participation in this research. We especially thank Ms Miho Ooga and Ms Chisa Hayashida for their technical assistance. K.Y. was supported partly by a Grant-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare and partly by grants from the Takeda Scientific Foundation and the Naito Foundation.

References

Abecasis, G. R., Burt, R. A., Hall, D., Bochum, S., Doheny, K. F., Lundy, S. L., Torrington, M., Roos, J. L., Gogos, J. A., & Karayiorgou, M. (2004). Genomewide scan in families with schizophrenia from the founder population of Afrikaners reveals evidence for linkage and uniparental disomy on chromosome 1. *American Journal of Human Genetics, 74,* 403–417.

Akbarian, S., Ruehl, M. G., Bliven, E., Luiz, L.A., Peranelli, A. C., Baker, S. P., Roberts, R. C., Bunney, W. E. Jr., Conley, R. C., Jones, E. G., Tamminga, C. A., & Guo, Y. (2005). Chromatin alterations associated with down-regulated metabolic gene expression in the prefrontal cortex of subjects with schizophrenia. *Archives of General Psychiatry, 62,* 829–840.

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics, 23,* 185–188.

Arnold, S. E., Han, L. Y., Moberg, P. J., Turetsky, B. I., Gur, R. E., Trojanowski, J. Q., & Hahn, C. G. (2001). Dysregulation of olfactory receptor neuron lineage in schizophrenia. *Archives of General Psychiatry, 58,* 829–835.

Baranzini, S. E., Mudge, J., van Velkinburgh, J. C., Khankhanian, P., Khrebtukova, I., Miller, N. A., Zhang, L., Farmer, A. D., Bell, C. J., Kim, R. W., May, G. D., Woodward, J. E., Caillier, S. J., McElroy, J. P., Gomez, R., Pando, M. J., Clendenen, L. E., Ganusova, E. E., Schilkey, F. D., Ramaraj, T., Khan, O. A., Huntley, J. J., Luo, S., Kwok, P. Y., Wu, T. D., Schroth, G. P., Oksenberg, J. R., Hauser, S. L., & Kingsmore, S. F. (2010). Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature, 464,* 1351–1356.

Bruder, C. E., Piotrowski, A., Gijsbers, A. A., Andersson, R., Erickson, S., de Ståhl, T. D., Menzel, U., Sandgren, J., von Tell, D., Poplawski, A., Crowley, M., Crasto, C., Partridge, E. C., Tiwari, H., Allison, D. B., Komorowski, J., van Ommen, G. J., Boomsma, D. I., Pedersen, N. L., den Dunnen, J. T., Wirdefeldt, K., & Dumanski, J. P. (2008). Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *American Journal of Human Genetics, 82,* 763–771.

Carter, N. P. (2007). Methods and strategies for analyzing copy number variation using DNA microarrays. *Nature Genetics, 39,* S16–21.

Cook, E. H., Jr & Scherer, S. W. (2008). Copy-number variations associated with neuropsychiatric conditions. *Nature, 455,* 919–923.

Crow, T. J. (2002). Handedness, language lateralization and anatomical asymmetry: Relevance of protocadherinXY to hominid speciation and the aetiology of psychosis. *British Journal of Psychiatry, 181,* 295–297.

Das, S., Kubota, T., Song, M., Daniel, R., Berry-Kravis, E. M., Prior, T. W., Popovich, B., Rosser, L., Arinami, T., & Ledbetter, D. H. (1997). Methylation analysis of the fragile X syndrome by PCR. *Genetic Testing, 1,* 151–155.

Kakiuchi, C., Ishiwata, M., Nanko, S., Ozaki, N., Iwata, N., Umekage, T., Tochigi, M., Kohda, K., Sasaki, T., Imamura, A., Okazaki, Y., & Kato, T. (2008). *Up*-regulation of *ADM* and *SEPX1* in the lymphoblastoid cells of patients in monozygotic twins discordant for schizophrenia. *American Journal of Medical Genetics Part B, 147B,* 557–564.

Kaminsky, Z. A., Tang, T., Wang, S. C., Ptak, C., Oh, G. H., Wong, A. H., Feldcamp, L. A., Virtanen, C., Halfvarson, J., Tysk, C., McRae, A. F., Visscher, P. M., Montgomery, G. W., Gottesman, I. I., Martin, N. G., & Petronis, A. (2009) DNA methylation profiles in monozygotic and dizygotic twins. *Nature Genetics, 41,* 240–245.

Kato, T., Iwamoto, K., Kakiuchi, C., Kuratomi, G., & Okazaki, Y. (2005). Genetic or epigenetic difference causing discordance between monozygotic twins as a clue to molecular basis of mental disorders. *Molecular Psychiatry, 10,* 622–630.

Kondo, S., Schutte, B. C., Richardson, R. J., Bjork, B. C., Knight, A. S., Watanabe, Y., Howard, E., de Lima, R. L., Daack-Hirsch, S., Sander, A., McDonald-McGinn, D. M., Zackai, E. H., Lammer, E. J., Aylsworth, A. S., Ardinger, H. H., Lidral, A. C., Pober, B. R., Moreno, L., Arcos-Burgos, M., Valencia, C., Houdayer, C., Bahuau, M., Moretti-Ferreira, D., Richieri-Costa, A., Dixon, M. J., & Murray, J. C. (2002). Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nature Genetics, 32,* 622–628.

McCarroll, S. A., Kuruvilla, F. G., Korn, J. M., Cawley, S., Nemesh, J., Wysoker, A., Shapero, M. H., de Bakker, P. I., Maller, J. B., Kirby, A., Elliott, A. L., Parkin, M., Hubbell, E., Webster, T., Mei, R., Veitch, J., Collins, P. J., Handsaker, R., Lincoln, S., Nizzari, M., Blume, J., Jones, K. W., Rava, R., Daly, M. J., Gabriel, S. B., & Altshuler, D. (2008). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics, 40,* 1166–1174.

McCarthy, S. E., Makarov, V., Kirov, G., Addington, A. M., McClellan, J., Yoon, S., Perkins, D. O., Dickel, D. E., Kusenda, M., Krastoshevsky, O., Krause, V., Kumar, R. A., Grozeva, D., Malhotra, D., Walsh, T., Zackai, E. H., Kaplan, P., Ganesh, J., Krantz, I. D., Spinner, N. B., Roccanova, P., Bhandari, A., Pavon, K., Lakshmi, B., Leotta, A., Kendall, J., Lee, Y. H., Vacic, V., Gary, S., Iakoucheva, L.M., Crow, T. J., Christian, S. L., Lieberman, J. A., Stroup, T. S., Lehtimäki, T., Puura, K., Haldeman-Englert, C., Pearl, J., Goodell, M., Willour, V. L., Derosse, P., Steele, J., Kassem, L., Wolff, J., Chitkara, N., McMahon, F. J., Malhotra, A. K., Potash, J. B., Schulze, T. G., Nöthen, M. M., Cichon, S., Rietschel, M., Leibenluft, E., Kustanovich, V., Lajonchere, C. M., Sutcliffe, J. S., Skuse, D., Gill, M., Gallagher, L., Mendell, N. R., Wellcome Trust Case Control Consortium, Craddock, N., Owen, M. J., O'Donovan, M. C., Shaikh, T. H., Susser, E., Delisi, L. E., Sullivan, P. F., Deutsch, C. K., Rapoport, J., Levy, D. L., King, M. C., & Sebat, J. (2009). Microduplications of 16p11.2 are associated with schizophrenia. *Nature Genetics*, *41*, 1223–1227.

Merikangas, A. K., Corvin, A. P., & Gallagher, L. (2009). Copy-number variants in neurodevelopmental disorders: Promises and challenges. *Trends in Genetics, 25*, 536–544.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W., Cho, E. K., Dallaire, S., Freeman, J. L., González, J. R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J. R., Marshall, C. R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M. J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D. F., Estivill, X., Tyler-Smith, C., Carter, N. P., Aburatani, H., Lee, C., Jones, K. W., Scherer, S. W., & Hurles, M. E. (2006). Global variation in copy number in the human genome. *Nature*, *444*, 444–454.

Roth, T. L., Lubin, F. D., Sodhi, M., & Kleinman, J. E. (2009). Epigenetic mechanisms in schizophrenia. *Biomedica et Biophysica Acta*, *1790*, 869–877.

Seal, J. L., Gornick, M. C., Gogtay, N., Shaw, P., Greenstein, D. K., Coffey, M., Gochman, P. A., Stromberg, T., Chen, Z., Merriman, B., Nelson, S. F., Brooks, J., Arepalli, S., Wavrant-De, Vrièze, F., Hardy, J., Rapoport, J. L., & Addington, A. M. (2006). Segmental uniparental isodisomy on 5q32-qter in a patient with childhood-onset schizophrenia. *Journal of Medical Genetics*, *43*, 887–892.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y. H., Hicks, J., Spence, S. J., Lee, A.T., Puura, K., Lehtimäki, T., Ledbetter, D., Gregersen, P. K., Bregman, J., Sutcliffe, J. S., Jobanputra, V., Chung, W., Warburton, D., King, M.C., Skuse, D., Geschwind, D. H., Gilliam, T. C., Ye, K., & Wigler, M. (2007). Strong association of de novo copy number mutations with autism. *Science*, *316*, 445–449.

Sebat, J., Levy, D. L., & McCarthy, S. E. (2009). Rare structural variants in schizophrenia: One disorder, multiple mutations; one mutation, multiple disorders. *Trends in Genetics*, *25*, 528–535.

Shih, R. A., Belmonte, P.L., & Zandi, P. P. (2004). A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *International Review of Psychiatry*, *16*, 260–283.

The Wellcome Trust Case Control Consortium. (2010). Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature*, *464*, 713–720.

Xu, B., Roos, J. L., Levy, S., van Rensburg, E. J., Gogos, J. A., & Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics*, *40*, 880–885.

Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., Nord, A. S., Kusenda, M., Malhotra, D., Bhandari, A., Stray, S. M., Rippey, C. F., Roccanova, P., Makarov, V., Lakshmi, B., Findling, R. L., Sikich, L., Stromberg, T., Merriman, B., Gogtay, N., Butler, P., Eckstrand, K., Noory, L., Gochman, P., Long, R., Chen, Z., Davis, S., Baker, C., Eichler, E. E., Meltzer, P. S., Nelson, S. F., Singleton, A. B., Lee, M. K., Rapoport, J. L., King, M. C., & Sebat, J. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*, *320*, 539–543.

West, P. M., Love, D. R., Stapleton, P. M., & Winship, I. M. (2003). Paternal uniparental disomy in monozygotic twins discordant for hemihypertrophy. *Journal of Medical Genetics*, *40*, 223–226.

Clinical Chemistry 56:11
1767–1771 (2010)

Brief Communication

Identification of Pregnancy-Associated MicroRNAs in Maternal Plasma

Kiyonori Miura,[1*] Shoko Miura,[1] Kentaro Yamasaki,[1] Ai Higashijima,[1] Akira Kinoshita,[2] Koh-ichiro Yoshiura,[2] and Hideaki Masuzaki[1]
Departments of [1] Obstetrics and Gynecology and [2] Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; * address correspondence to this author at: Department of Obstetrics and Gynecology, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8501, Japan. Fax +81-95-819-7365; e-mail kiyonori@nagasaki-u.ac.jp.

BACKGROUND: Several placental microRNAs (miRNAs) have been identified as pregnancy-associated molecules with the potential for use in estimating the condition of the placenta. Our understanding of these novel molecules is still limited, however. The aim of this study was to isolate and characterize pregnancy-associated miRNAs in maternal plasma.

METHODS: By microarray-based screening of 723 human miRNAs, we selected miRNAs that exhibited signal intensities >100 times higher in placental tissues than in the corresponding whole blood samples. Subsequent quantitative real-time reverse-transcription PCR revealed miRNAs produced predominantly in the placenta that showed significantly decreased concentrations in maternal plasma after delivery. These miRNAs were identified as pregnancy-associated miRNAs.

RESULTS: We selected 82 miRNAs produced predominantly in the placenta and identified 24 as pregnancy-associated miRNAs. The genes encoding these miRNAs included 16 that are clustered on 19q13.42 and 5 clustered on 14q32. As the pregnancy progressed into the third trimester, the plasma concentrations of cell-free chromosome 19–derived miRNAs (has-miR-515-3p, has-miR-517a, has-miR-517c, has-miR-518b, and has-miR-526b) increased significantly ($P = 0.0284$, 0.0069, 0.0125, 0.0284, and 0.0093, respectively, Wilcoxon signed rank test), whereas that of cell-free has-miR-323-3p on chromosome 14q32.31 showed no change ($P = 0.2026$).

CONCLUSIONS: In addition to the known pregnancy-associated miRNAs, we identified new pregnancy-associated miRNAs with our microarray-based approach. Most of the genes encoding these miRNAs were clustered on 19q13.42 or 14q32, which are critical regions for placental and embryonic development. These new pregnancy-associated miRNAs may be useful molecular markers for monitoring pregnancy-associated diseases.

MicroRNAs (miRNAs) are nonprotein-coding small RNAs (20–24 nucleotides) that function as regulators of gene expression via antisense complementarity to specific mRNAs *(1–3)*. Therefore, miRNAs produced predominantly in the placenta are probably involved in placental differentiation and the maintenance of pregnancy.

Cell-free placental DNA and/or mRNAs in maternal plasma are possible molecular markers for noninvasive prenatal monitoring or early detection of pregnancy-associated diseases *(4–8)*. A recent search of a panel of 157 miRNA assays and sequencing of a small-RNA library revealed placental miRNAs as pregnancy-associated markers that are present at consistent concentrations in the maternal circulation *(9, 10)*. Similarly to cell-free placental DNA/mRNAs, the measurement of placental miRNA concentrations in maternal plasma may offer a noninvasive test for monitoring pregnancy-associated diseases. Our understanding of pregnancy-associated miRNAs is still limited, however. In this study, we used a microarray-based approach to identify pregnancy-associated miRNAs that occur in the maternal plasma. We then investigated the chromosomal locations of the miRNA-encoding genes and the time dependency of the changes in circulating miRNA concentrations occurring in maternal plasma.

For the miRNA microarray analysis, we used 2 sets of placental tissue samples and corresponding maternal blood samples collected in the first trimester (12–13 weeks of pregnancy) and another 2 sets collected in the third trimester (38–39 weeks). Placental tissue samples were obtained immediately after the termination of pregnancy and placed in RNAlater™ (Ambion/Applied Biosystems). Blood samples (7 mL) were collected before the termination of pregnancy and placed in tubes containing EDTA. miRNA was extracted from placental tissues and maternal blood cells immediately after sampling.

For the quantitative analysis of miRNAs in maternal plasma, we collected nonserial blood samples from 10 women in the first trimester of pregnancy (12–13 weeks) and from 10 women in the third trimester (38–39 weeks) before delivery and 1 day after pregnancy termination. Serial blood samples were also collected from 10 pregnant women during the first and third trimesters and after the termination of pregnancy. All of these women were admitted to Nagasaki University Hospital. All samples were obtained after we received written informed consent, and the study pro-

tocol was approved by the Institutional Review Board for Ethical, Legal and Social Issues of Nagasaki University.

Cell-free plasma samples were prepared from maternal blood by the double-centrifugation method described previously *(11)*. After a first centrifugation at 3000*g* for 10 min, we centrifuged the supernatant at 16 000*g* for 10 min to remove blood cells. Total RNA containing small RNA molecules was extracted from 5.0 mL maternal blood and 3.0 mL maternal plasma with a *mir*Vana miRNA Isolation Kit (Ambion/Applied Biosystems) according to the manufacturer's instructions.

Quality assessment and concentration measurements of total RNA, including small RNAs, were performed with a Bioanalyzer (Agilent Technologies) and a NanoDrop spectrophotometer (Thermo Fisher Scientific), respectively. We subjected 100 ng total RNA from the placenta or from the corresponding maternal blood sample of each trimester to microarray analysis. For each sample, 1 μg total RNA (which included small RNAs) was labeled with cyanine 3–cytidine bisphosphate (pCp-Cy3). The miRNA Labeling Reagent and Hybridization Kit (Agilent Technologies) was then used according to the manufacturer's instructions to hybridize the labeled RNA for 20 h to Human miRNA Microarray Version 2 (Agilent Technologies), which included 723 miRNAs encoded by genes located across all chromosomes except chromosome Y. The number of miRNAs located on each chromosome in the Human miRNA Microarray Version 2 varied from 7 to 86, and the numbers of miRNAs located on chromosomes 19 (n = 86), 14 (n = 75), and X (n = 86) were relatively larger than the numbers for the other chromosomes. Slides were washed for 10 min at room temperature in 6× standard saline citrate (0.9 mol/L NaCl and 0.09 mol/L sodium citrate) containing 50 μL/L Triton X-102 and then for 5 min in 0.1× standard saline citrate (0.015 mol/L NaCl and 0.0015 mol/L sodium citrate) containing 50 μL/L Triton X-102. The slides were scanned on an Agilent microarray scanner (model G2565A) at 100% and 5% sensitivity settings. Agilent Feature Extraction software (version 9.5) was used for image analysis. miRNAs exhibiting signal intensities in placental samples >100 times those of the corresponding whole blood samples were selected as placentally produced miRNAs. These miRNAs were considered candidate pregnancy-associated miRNAs (see Fig. 1 in the Data Supplement that accompanies the online version of this Brief Communication at http://www.clinchem.org/content/vol56/issue11).

For quantitative real-time reverse-transcription PCR analysis, we used a TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions and 2.5 μL total RNA in a total reaction volume of 25 μL. We then performed quantitative PCR with the TaqMan Universal PCR Master Mix (Applied Biosystems). For each miRNA assay, we prepared a calibration curve by 10-fold serial dilution of single-stranded cDNA oligonucleotides corresponding to each miRNA sequence over the concentration interval of 1.0×10^3 to 1.0×10^7 copies/mL. Each sample and each calibration dilution were analyzed in triplicate. Each assay could detect the RNA concentration down to 1000 copies/mL of plasma. Every batch of amplifications included 3 water blanks as negative controls for each of the reverse-transcription and PCR steps. All data were collected and analyzed with an ABI Prism 7900 Sequence Detection System (Applied Biosystems). miRNAs were selected as pregnancy-associated miRNAs when circulating placenta-produced miRNAs were detected (>1000 copies/mL) in all 10 maternal plasma samples before pregnancy termination and showed a significantly decreased concentration after the termination of the pregnancy ($P < 0.05$, Wilcoxon signed rank test; see Fig. 1 in the online Data Supplement).

Chromosomal localization of the encoding gene and tissue production were evaluated for each miRNA by searching the microRNA.org (http://www.microrna.org/) and miRBase (http://www.mirbase.org/index.shtml) databases. Statistical analysis was performed with StatView 5.0 software (SAS Institute). Differences were evaluated with the Wilcoxon signed rank test, and statistical significance was set at P values <0.05.

Of 723 human miRNAs, 82 were produced predominantly in the placenta according to our comparisons of the miRNA profiles of placental tissues for the first and the third trimesters and those of the corresponding samples of maternal blood cells. These placental miRNAs occurred with signal intensities >100-fold higher than those in maternal blood cells. Of these 82 miRNAs (see Table 1 in the online Data Supplement), 12 were detected only in the first trimester, 25 were detected only in the third trimester, and the remaining 45 miRNAs were detected in both trimesters (see Table 1 in the online Data Supplement).

Our analysis of the chromosomal locations of the genes encoding the 82 miRNAs that were predominantly produced in the placenta showed that 44 (53.7%) were clustered on chromosome 19q13 and 13 (15.9%) were clustered on 14q32 (see Table 1 in the online Data Supplement). Thirty-six miRNAs (81.8%) encoded by genes clustered on 19q13 and 2 miRNAs (15.4%) encoded by genes clustered on 14q32 were detected in both trimesters.

Of the 82 miRNAs that were produced predominantly in the placenta, 24 showed significantly decreased concentrations in the maternal plasma after de-

Table 1. List of pregnancy-associated miRNAs in maternal plasma.

miRNA probe set ID[a]	Pattern[b]	Chromosomal location	Trimester[c]	TaqMan ID	Circulating miRNA, copies/mL[d]		P[f]	CV[g]
					Before TOP[e]	After TOP		
hsa-miR-515-3p	S	19q13.42	F, T	002369	18 100 (9100–83 500)	1700 (1200–4200)	0.0051	6.3%
hsa-miR-515-5p	P	19q13.42	F, T	002369	13 100 (6200–38 900)	1700 (0–6800)	0.0069	5.2%
hsa-miR-517a	S	19q13.42	F, T	002402	27 100 (11 300–73 100)	1600 (1200–2300)	0.0051	5.2%
hsa-miR-517c	S	19q13.42	F, T	001153	29 300 (12 300–68 400)	1400 (1200–4500)	0.0051	9.0%
hsa-miR-517*	S	19q13.42	F, T	001113	42 100 (13 300–97 300)	1400 (1300–18 400)	0.0069	6.6%
hsa-miR-518b	S	19q13.42	F, T	001156	26 000 (5300–78 200)	1600 (1400–2300)	0.0051	2.6%
hsa-miR-518c	S	19q13.42	F, T	002401	23 000 (7100–93 900)	2000 (0–23 200)	0.0166	6.6%
hsa-miR-518e	S	19q13.42	F, T	002395	22 400 (1800–33 800)	1600 (1200–10 800)	0.0166	3.4%
hsa-miR-519a	S	19q13.42	F, T	002415	10 100 (2700–84 300)	1750 (1100–7100)	0.0125	5.0%
hsa-miR-519d	S	19q13.42	F, T	002403	4000 (1800–9300)	1800 (0–2800)	0.0069	1.3%
hsa-miR-520a-5p	P	19q13.42	F, T	001168	9100 (1300–34 300)	1400 (1300–3900)	0.0093	4.5%
hsa-miR-525-5p	S	19q13.42	F, T	001174	9100 (3100–94 800)	4500 (0–8500)	0.0367	8.7%
hsa-miR-526b	S	19q13.42	F, T	002382	25 300 (12 300–53 100)	1600 (1200–2300)	0.0051	10%
hsa-miR-498	S	19q13.42	F	001044	28 700 (3800–94 900)	5500 (0–9800)	0.0284	2.3%
hsa-miR-525-3p	P	19q13.42	F	002385	9300 (1800–43 000)	1900 (1300–3400)	0.0069	4.2%
hsa-miR-526b*	S	19q13.42	T	002383	4300 (1300–9500)	1600 (0–1900)	0.0166	10%
hsa-miR-323-3p	P	14q32.31	F, T	002227	18 600 (7700–35 900)	2100 (1500–5400)	0.0051	3.5%
hsa-miR-433	P	14q32.2	F, T	001028	9000 (4200–18 300)	2200 (1400–5400)	0.0125	2.9%
hsa-miR-411	P	14q32.31	F	001610	4900 (1300–9300)	2000 (1200–4300)	0.0218	0.8%
hsa-miR-487a	P	14q32.31	F	001279	6700 (2100–11 900)	5600 (0–8400)	0.0367	7.6%
hsa-miR-154*	P	14q32.31	T	000478	8800 (4900–15 900)	7000 (1500–9300)	0.0125	8.2%
hsa-miR-218	P	4p15.31	F, T	000521	5800 (2700–12 000)	4200 (0–9900)	0.0166	5.1%
hsa-miR-204	P	9q21.12	T	000508	3900 (1300–8700)	2500 (0–3900)	0.0367	4.6%
hsa-miR-141	P	12p13.31	T	002145	17 000 (5000–281 200)	7400 (1800–14 900)	0.0069	5.3%

[a] Each asterisk (*) below is part of the microRNA name.
[b] S, specific expression pattern from placenta tissue; P, predominant expression pattern from placenta tissue. See http://www.microrna.org/microrna/getExprForm.do.
[c] F, first trimester; T, third trimester.
[d] Circulating miRNA concentrations in maternal plasma are presented as the median (range).
[e] TOP, termination of pregnancy.
[f] Wilcoxon signed rank test.
[g] Intraassay CV in each quantitative reverse-transcription PCR assay. The SD was obtained from a single triplicate measurement.

livery of the placenta (Table 1); we therefore identified them as pregnancy-associated miRNAs. Of the genes encoding these 24 pregnancy-associated miRNAs, 21 (87.5%) were clustered on 19q13.42 or 14q32, 16 (66.7%) were located on 19q13.42, and 5 (20.8%) were on 14q32. The mean intraassay CV, which is the ratio of the standard deviation to the mean, for the probes in the quantitative real-time reverse-transcription PCRs was 4.7% (range, 0.6%–11%) (Table 1). A search of the microRNA.org database (http://www.microrna.org/) for miRNA production in tissues indicated that 13 of the 24 genes encoding pregnancy-associated miRNAs had a "specific expression pattern from placenta tissue," which indicated that these miRNA genes were not expressed in tissues other than placenta (Table 1). All of the genes encoding these miRNAs are located on 19q13.42. The genes encoding the remaining 11 pregnancy-associated miRNAs (3 on 19q13.42 and 8 on other chromosomes) indicated a "predominant expression pattern from placenta tissue," which indicated miRNAs that were produced predominantly in placenta tissue rather than in blood cells (Table 1).

Of the 24 pregnancy-associated miRNAs, 6 miRNAs (has-miR-515-3p, has-miR-517a, has-miR-517c, has-miR-518b, has-miR-526b, and has-miR-323-3p) that showed the most significantly decreased concentrations in maternal plasma after pregnancy termination were selected for an analysis of time-

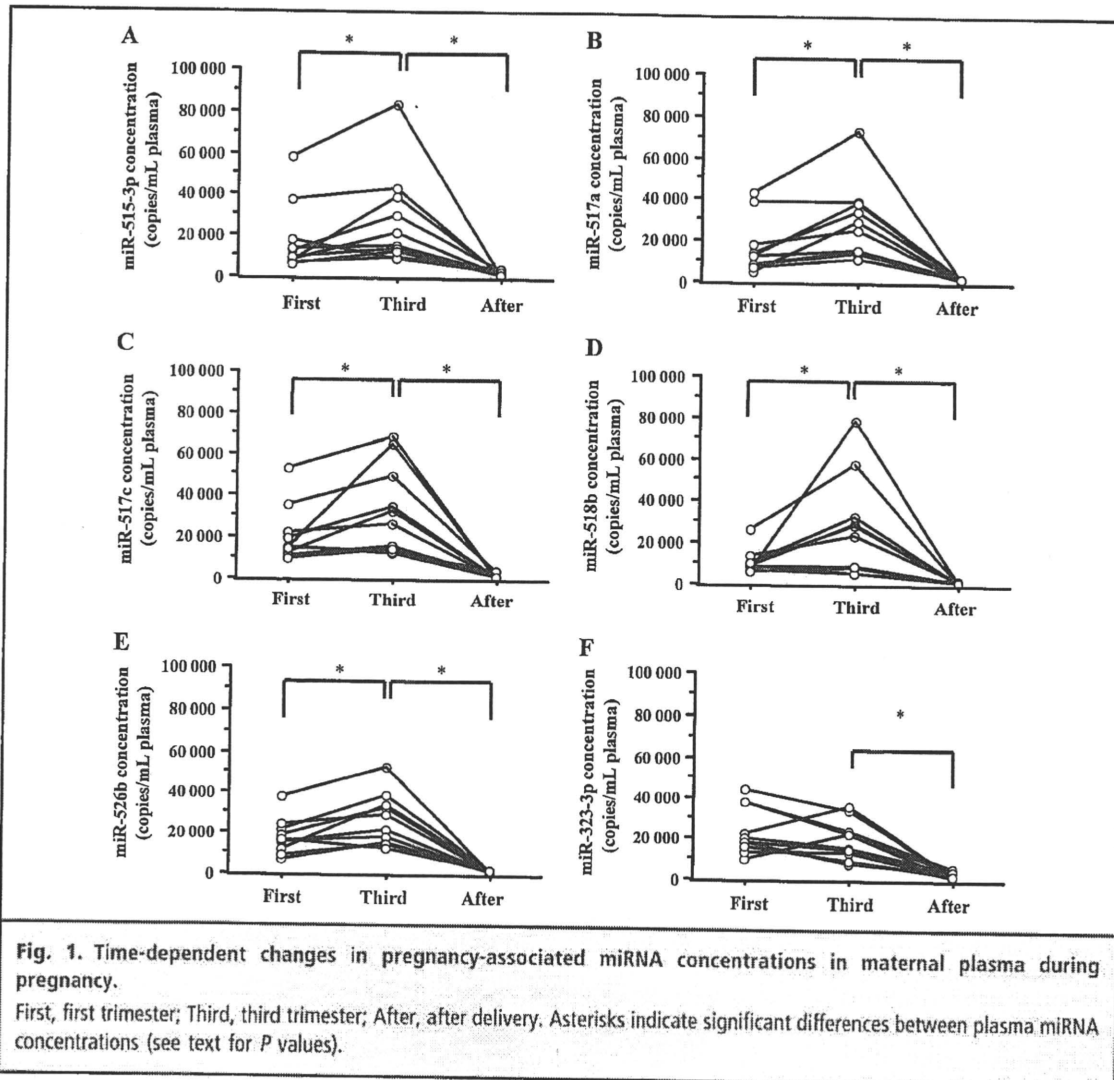

Fig. 1. Time-dependent changes in pregnancy-associated miRNA concentrations in maternal plasma during pregnancy.

First, first trimester; Third, third trimester; After, after delivery. Asterisks indicate significant differences between plasma miRNA concentrations (see text for *P* values).

dependent changes in cell-free miRNA concentrations in maternal plasma during pregnancy ($P = 0.0051$, Wilcoxon signed rank test; Table 1). As the pregnancy progressed into the third trimester, the plasma concentrations of cell-free chromosome 19q13.42–derived miRNAs (has-miR-515-3p, has-miR-517a, has-miR-517c, has-miR-518b, and has-miR-526b) increased significantly ($P = 0.0284$, 0.0069, 0.0125, 0.0284, and 0.0093, respectively, Wilcoxon signed rank test), whereas the concentration of cell-free has-miR-323-3p encoded by the corresponding miRNA gene on 14q32.31 showed no change ($P = 0.2026$) (Fig. 1).

Most of the initial 82 miRNAs with predominant production in placental tissue showed no post-delivery clearance, suggesting that maternal tissues other than blood cells contributed to their presence in maternal plasma. It is noteworthy that 21 of 24 genes encoding pregnancy-associated miRNAs present in maternal plasma were clustered on 19q13.42 or 14q32, which are critical regions for placental growth and embryonic development *(12–15)*. The concentrations of circulating pregnancy-associated miRNAs in maternal plasma showed a time dependence as pregnancy progressed into the third trimester. Therefore, the pregnancy-associated miRNAs we identified may be useful molecular markers for the early detection or prenatal monitoring of pregnancy-associated diseases, such as preeclampsia, placenta accreta, intrauterine growth restriction, and hydatidiform moles. Further understanding of the clinical importance of the pregnancy-associated miRNAs we detected will re-

quire clarification about how the alterations in the concentrations of these miRNAs in maternal plasma affect specific gene expression in some maternal tissues.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: K. Miura, Grant-in-Aid for Young Scientists (B) (no. 21791567) from the Ministry of Education, Sports, Culture, Science and Technology of Japan; Seed Grant (no. J089500122) from the Japan Science and Technology Agency (JST); grants from the Naito Foundation; and a grant for Child Health and Development (no. 20C-1) from the Ministry of Health, Labour and Welfare, Japan; K. Yoshiura, Grants-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare, Japan.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993;75:843–54.
2. Mendell JT. MicroRNAs: critical regulators of development, cellular physiology and malignancy. Cell Cycle 2005;4:1179–84.
3. Plasterk RH. Micro RNAs in animal development. Cell 2006;124:877–81.
4. Chiu RW, Lui WB, Cheung MC, Kumta N, Farina A, Banzola I, et al. Time profile of appearance and disappearance of circulating placenta-derived mRNA in maternal plasma. Clin Chem 2006;52:313–6.
5. Purwosunu Y, Sekizawa A, Farina A, Wibowo N, Okazaki S, Nakamura M, et al. Cell-free mRNA concentrations of CRH, PLAC1, and selectin-P are increased in the plasma of pregnant women with preeclampsia. Prenat Diagn 2007;27:772–7.
6. Lo YM, Chiu RW. Prenatal diagnosis: progress through plasma nucleic acids. Nat Rev Genet 2007;8:71–7.
7. Maron JL, Bianchi DW. Prenatal diagnosis using cell-free nucleic acids in maternal body fluids: a decade of progress. Am J Med Genet C Semin Med Genet 2007;145:5–17.
8. Miura K, Miura S, Yamasaki K, Yoshida A, Yoshiura K, Nakayama D, et al. Increased level of cell-free placental mRNA in a subgroup of placenta previa that needs hysterectomy. Prenat Diagn 2008;28:805–9.
9. Chim SS, Shing TK, Hung EC, Leung TY, Lau TK, Chiu RW, Lo YM. Detection and characterization of placental microRNAs in maternal plasma. Clin Chem 2008;54:482–90.
10. Luo SS, Ishibashi O, Ishikawa G, Ishikawa T, Katayama A, et al. Human villous trophoblasts express and secrete placenta-specific microRNAs into maternal circulation via exosomes. Biol Reprod 2009;81:717–29.
11. Chiu RW, Poon LL, Lau TK, Leung TN, Wong EM, Lo YM. Effects of blood-processing protocols on fetal and total DNA quantification in maternal plasma. Clin Chem 2001;47:1607–13.
12. Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, et al. Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet 2005;37:766–70.
13. Tsai KW, Kao HW, Chen HC, Chen SJ, Lin WC. Epigenetic control of the expression of a primate-specific microRNA cluster in human cancer cells. Epigenetics 2009;4:587–92.
14. Seitz H, Royo H, Bortolin ML, Lin SP, Ferguson-Smith AC, Cavaillé J. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res 2004;14:1741–8.
15. Royo H, Bortolin ML, Seitz H, Cavaillé J. Small non-coding RNAs and genomic imprinting. Cytogenet Genome Res 2006;113:99–108.

Previously published online at
DOI: 10.1373/clinchem.2010.147660

Association between Breast Cancer Risk and the Wild-type Allele of Human ABC Transporter *ABCC11*

IKUKO OTA[1,2], AKI SAKURAI[3], YU TOYODA[3], SATOSHI MORITA[4], TAKESHI SASAKI[5], TAKASHI CHISHIMA[1], MINORU YAMAKADO[6], YUKI KAWAI[7], TAKEFUMI ISHIDAO[7], ALEXANDER LEZHAVA[8], KOH-ICHIRO YOSHIURA[9], SHINJI TOGO[1], YOSHIHIDE HAYASHIZAKI[8], TAKASHI ISHIKAWA[2], TOSHIHISA ISHIKAWA[3,8], ITARU ENDO[1] and HIROSHI SHIMADA[1]

[1]*Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine and Departments of* [2]*Breast and Thyroid Surgery,* [4]*Biostatistics and Epidemiology and* [5]*Pathology, Yokohama City University Medical Center, Yokohama, Japan;*
[3]*Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan;*
[6]*Center for Multiphasic Health Testing and Services, Mitsui Memorial Hospital, Tokyo, Japan;*
[7]*K.K. DNAFORM, Yokohama, Japan;*
[8]*Omics Science Centre, RIKEN Yokohama Institute, Yokohama, Japan;*
[9]*Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan*

Abstract. *Background: International mortality and frequency rates for breast cancer have been associated with the wet type of human earwax. It was recently found that earwax type is determined by a single nucleotide polymorphism (SNP), 538G>A (Gly180Arg), in ABCC11. The G allele determines the wet type of earwax as a Mendelian trait with a dominant phenotype. The present study examined the association between the frequency rate of breast cancer and the frequency of the G allele of ABCC11. Patients and Methods: Using blood samples from patients with invasive breast cancer (n=270) and control volunteers (n=273), the 538G>A SNP in ABCC11 was genotyped using the SmartAmp method. Results: The frequency of the G allele in breast cancer patients was higher than that in healthy controls. The odds ratio for the genotypes (G/G+G/A) to develop breast cancer was estimated to be 1.63 (p-value=0.026), suggesting that the G allele in ABCC11 is associated with breast cancer risk. Conclusion: This study showed that Japanese women with wet earwax have a higher relative risk of developing breast cancer than those with dry earwax. The ABCC11 SNPs that determine these phenotypes should be further investigated in order to obtain insights into the mechanisms by which breast cancer develops and progresses.*

Correspondence to: Toshihisa Ishikawa, Ph.D., Omics Science Centre, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. Tel: +81 455039222; Fax: +81 455039216, e-mail: toshi-i@gsc.riken.jp

Key Words: Apocrine gland, axillary osmidrosis, earwax, mammary gland, SNP, ABC transporter *ABCC11*, breast cancer risk.

Breast cancer is the most common cancer among women in the industrialised world, where it accounts for 22% of all cancers in women. There has been increased interest in the genetic predisposition for many common cancer types, including breast cancer (1). In 1971, Nicholas L. Petrakis first reported that international mortality and frequency rates for breast cancer appeared to be associated with the frequency of the allele for wet-type earwax (2). Caucasian and African-American women in the USA as well as German women exhibited approximately four-fold higher rates of breast cancer mortality than Japanese and Taiwanese women (2). Nevertheless, the phenotypic association of the wet type of earwax with breast cancer has remained controversial (2, 3).

Recent studies (4, 5) have provided evidence that the type of earwax is determined by one single nucleotide polymorphism (SNP), 538G>A (Gly180Arg), in the ATP-binding cassette (ABC) transporter *ABCC11* located on human chromosome 16q12.1 (6). The G/G and G/A genotypes correspond to the wet type of earwax, whereas A/A corresponds to the dry type (5). Wide ethnic differences have been observed in the frequencies of those alleles (4, 7). Among worldwide populations, the 'G' (wild-type) allelic frequency shows striking downward geographical gradient distributions from Africa to Far East Asia, supporting previous phenotypic observations (8). Interestingly, there are strong associations among wild-type *ABCC11*, earwax type (4), axillary osmidrosis

0250-7005/2010 $2.00+.40

(5, 9, 10), and apocrine colostrum secretion from the mammary gland (11). Furthermore, *ABCC11* mRNA is highly expressed in breast tumours (6, 12, 13).

At the present time, it is not well understood whether wild-type *ABCC11* actually influences breast cancer risk. The present study therefore genotyped the 538G>A SNP in 543 Japanese women to examine the association between the frequency rate of breast cancer and the allele frequency of the wild-type. For this purpose, the SmartAmp method was used to measure 538G>A SNP frequency in *ABCC11* (14, 15). The results suggest that the wild-type allele in *ABCC11* is associated with breast cancer risk.

Patients and Methods

Collection of blood samples from breast cancer patients and control volunteers. Blood samples from 270 Japanese female patients with invasive breast cancer who had been diagnosed at Yokohama City University Medical Centre from 1991 to 2008. In addition, blood samples were also collected from 273 Japanese female volunteers as controls. All blood samples were collected in 2Na-EDTA-coated blood collection tubes. All study participants provided written informed consent and protocols for the present study were approved by the Institutional Review Boards at both Yokohama City University Medical Centre and Mitsui Memorial Hospital. This clinical research study was conducted according to the Declaration of Helsinki Principles. Genotyping of *ABCC11* in the blood samples by the SmartAmp method was approved by the Research Ethical Committee at RIKEN Yokohama Institute.

Clinicopathological data. For breast cancer patients, clinicopathological data was acquired such as age, body mass index (BMI), tumour size, lymph node metastasis (N), the status of the oestrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2), triple negative (ER^- PR^- $HER2^-$) tumour phenotype, nuclear grade, and the St. Gallen risk assignment criteria. ER, PR, and HER2 status in breast cancer was determined by immunohistochemistry.

Genotyping of ABCC11 by the SmartAmp method. Detection of the 538G>A SNP in *ABCC11* was performed as described previously (5). Prior to the SmartAmp reaction, the blood samples were incubated at 98°C for 3 min to destroy RNA and to denature proteins and genomic DNA. After chilling on ice, each sample (1 μl) was added directly into the SmartAmp reaction mixture (final total volume of 25 μl). The reaction mixtures were then incubated at 60°C for 40 min under an isothermal condition in a real-time PCR model Mx3000P system (Agilent technologies, La Jolla, CA, USA), where the fluorescence intensity of SYBR® Green I dye indicating DNA amplification was monitored during the reaction.

Statistical rationale. The sample size required for this clinical study was calculated by assuming that 26% of Japanese breast cancer patients and 16% of control volunteers carry G/G homozygote or G/A heterozygote genotypes. The rationale for this assumption originated from Petrakis' pilot study (2), which showed that more Japanese women with breast cancer (9 out of 31; 29%) had the wet type of earwax than did women of the control group (9 out of 52; 17%). The minimum number of subjects needed for confirming significant differences between the cancer and control groups with 80% power was calculated to be 270 in total for a two-sided model at a 5.0% significance level. Thus, for this clinical study, blood samples were collected from a total of 270 Japanese breast cancer patients and 273 Japanese control volunteers.

Statistical analysis. To evaluate the statistical significance of observed data, χ^2 tests were performed using the Dr. SPSS software (SPSS11.5J for Windows; SPSS Inc., Chicago, IL, USA) for univariate analysis and logistic regression for multivariate analyses.

Results

Detection of the 538G>A SNP in ABCC11 by the SmartAmp method. The SmartAmp method can genotype *ABCC11* in blood samples within 40 min (5). Figure 1A depicts the time courses for the SmartAmp reaction that clearly discriminated the three different genotypes (538G/G homozygote, 538G/A heterozygote, and 538A/A homozygote) in human *ABCC11*, where blood samples were pre-treated and incubated as described in the legend of Figure 1A.

The mean age was 52.9±12.52 and 53.3±9.96 years old (mean±S.D.) for the breast cancer patient group and control group, respectively. The mean body mass index (BMI) value was 23.0±3.87 for the breast cancer patient group and 21.0±3.02 for the control group.

Table I compares the *ABCC11* genotypes and earwax type in the study population. The odds ratio was calculated to be 1.63 (p-value=0.026) (Table I).

Figure 1B shows that the relative ratio of breast cancer patients carrying the homozygous 538G/G allele was 1.77-fold greater than that of the corresponding healthy volunteers. This relative ratio was even greater than that (1.41-fold) for breast cancer patients carrying the heterozygous 538G/A allele. The G allele appears to be positively related to breast cancer frequency in the groups of Japanese women studied.

Clinicopathological observations. The prognosis for breast cancer patients with wet earwax has been reported to be worse than for that for patients with dry earwax (2). In this context, it was anticipated in this study that the wet earwax genotypes might have some relation to specific clinicopathological features, such as the status of growth factor receptors or nuclear grade. Since histopathology data were available for all breast cancer patients involved in this clinical study, a possible relationship between *ABCC11* genotypes and clinicopathological features including tumour size, lymph node metastasis, ER, PR, and HER2 status, triple negative tumour phenotype, nuclear grade, tumour stage, and St. Gallen risk was investigated (Table II). Nevertheless, neither chi-square tests nor logistic regression analysis revealed any statistically significant difference between the wet earwax genotypes (538G/G+538G/A) and the dry earwax genotype (538A/A) with respect to the clinico-pathological features investigated in this study.

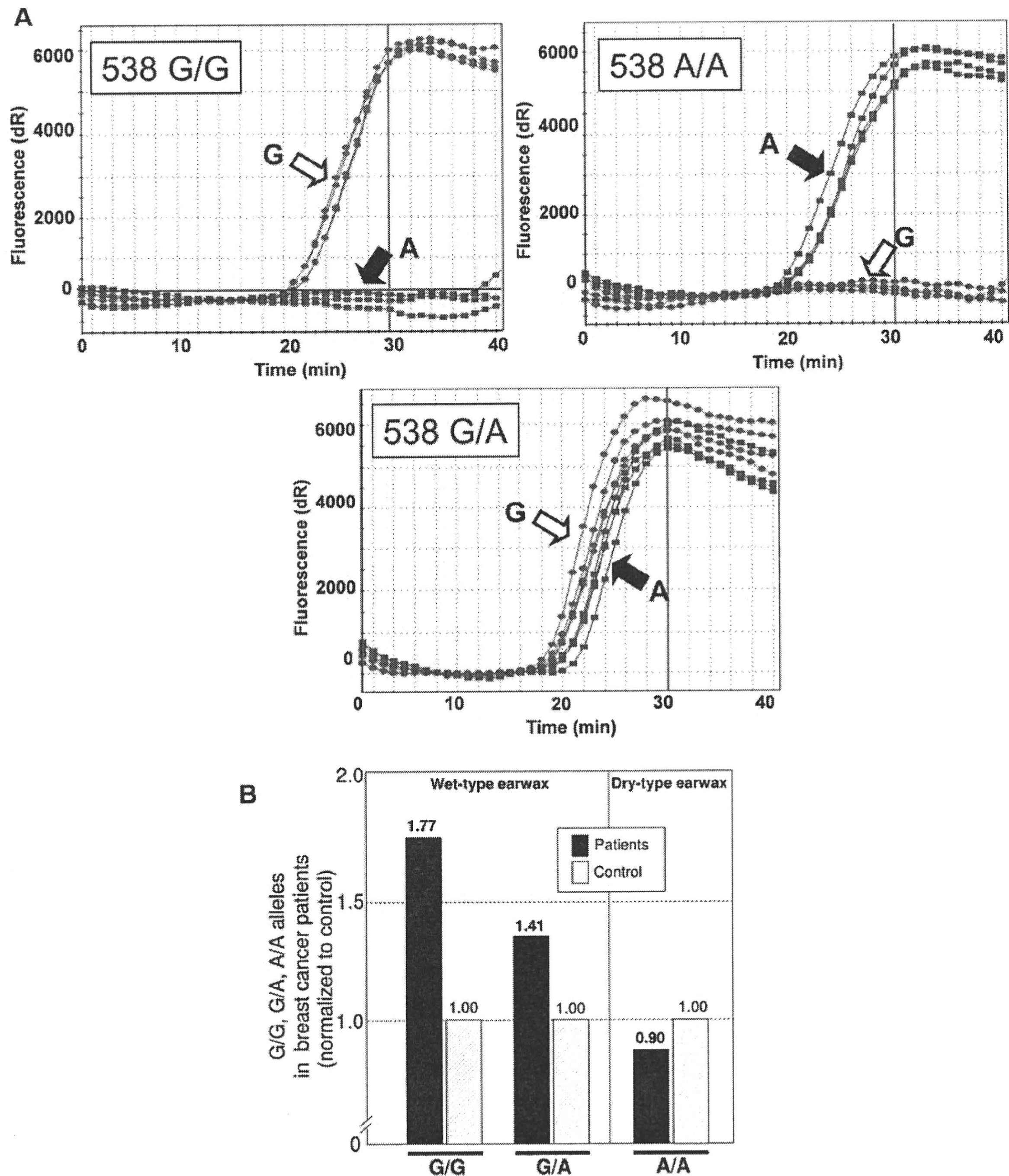

Figure 1. *SmartAmp method-based genotyping of human ABCC11 in blood samples from breast cancer patients and control subjects. A: Blood was mixed with 50 mM NaOH (1:2 v/v) and incubated at 98°C for 3 min. After chilling on ice, 1 µl of the sample was mixed with the SmartAmp reaction mixture (a final total volume of 25 µl) containing 2.0 µM folding primer, 2.0 µM turn-back primer, 1.0 µM boost primer, 0.25 µM of each outer primer, 20 µM competitive probe, 1.4 mM dNTPs, 5% DMSO, 20 mM Tris-HCl (pH 8.0), 10 mM KCl, 10 mM* $(NH_4)_2SO_4$, *8 mM* $MgSO_4$, *0.1% (v/v) Tween®20, SYBR® Green I (Takara Bio Inc., Shiga, Japan) diluted 1/100,000, and 0.24 unit/µl Aac DNA polymerase (K.K. DNAFORM, Yokohama, Japan). SmartAmp reaction mixtures were incubated at 60°C for 40 min in a real-time PCR model Mx3000P system (Agilent Technologies, La Jolla, CA, USA). The fluorescence intensity of the SYBR® Green I dye was monitored. The time courses represent SmartAmp reactions with ABCC11 allele-specific primers: 538G (red) and 538A (blue). B: Relative ratios of breast cancer patients carrying 538G/G, 538G/A, and 538 A/A alleles. Data are normalised to the controls (female volunteers).*

Table I. *Comparisons of ABCC11 genotypes and earwax type between Japanese breast cancer patients and control subjects.*

	Patients Number (%)	Controls Number (%)	Chi-square test *p*-value
SNP in *ABCC11*			
538G/G	14 (5.19)	8 (2.93)	
538G/A	53 (19.6)	38 (13.9)	
538A/A	203 (75.2)	227 (83.2)	
Phenotype of earwax			
Wet (538G/G + G/A)	67 (24.8)	46 (16.8)	0.026*
Dry (538A/A)	203 (75.2)	227 (83.2)	

*Statistically significant according to the Chi-square tests.

Table II. *Comparison of clinicopathological features between sub-populations with wet-type (538G/G + 538G/A in ABCCC11) and dry-type (538A/A in ABCC11) earwax in Japanese breast cancer patients.*

	538G/G + 538G/A Number (%)	538A/A Number (%)
Tumour size		
<2 cm	32 (51.6)	101 (51.5)
≥2 cm	30 (48.4)	95 (48.5)
Lymph node metastasis		
N (+)	30 (47.6)	81 (40.7)
N (–)	33 (52.4)	118 (59.3)
Receptors		
ER (+)	52 (78.8)	148 (74.0)
ER (–)	14 (21.2)	52 (26.0)
PR (+)	35 (53.8)	116 (58.0)
PR (–)	30 (46.2)	84 (42.0)
HER2 (+)	16 (24.2)	34 (17.3)
HER2 (–)	50 (75.8)	153 (82.7)
Triple negative	5 (7.58)	31 (15.7)
Non-triple negative	61 (92.4)	166 (84.3)
Nuclear grade		
Grade 1 & 2	38 (57.1)	128 (63.4)
Grade 3	21 (35.6)	69 (36.7)
Stage		
1 & 2A	36 (57.1)	128 (63.4)
>2B	27 (42.9)	74 (36.6)
St. Gallen risk		
Low	14 (22.6)	38 (19.3)
Intermediate	30 (49.4)	112 (56.9)
High	18 (29.0)	47 (23.6)

ER: Oestrogen receptor; PR: progesterone receptor. Triple negative means ER (–) PR (–) HER2 (–).

Discussion

Potential role of ABCC11 in breast cancer. The present study provides evidence that the wild type allele of the *ABCC11* gene is associated with breast cancer risk, at least in the Japanese population. About 40 years ago, Petrakis (2) assumed that genetically determined variation in the apocrine system might influence susceptibility to breast cancer, although the genetic determinant (538G>A SNP in *ABCC11*) was not known at that time. It is only recently that more than 10 nonsynonymous SNPs have been found in the human *ABCC11* gene (5, 7). Among those SNPs, one SNP (rs17822931; 538G>A, Gly180Arg) is thought to be a clinically important polymorphism that may related with breast cancer risk.

Genetic polymorphisms of ABCC11 gene. The wild-type (538G or Gly180) human ABC transporter *ABCC11* expressed in apocrine glands plays a pivotal role in earwax secretion (4, 5), axillary osmidrosis (5, 9, 10, 16), and apocrine colostrum secretion from the mammary gland (11). Human ABCC11 reportedly functions as an ATP-dependent efflux pump for amphipathic anions, including oestrone 3-sulfate, dehydroepiandrosterone 3-sulfate (DHEAS), and oestradiol 17-β-D-glucuronide (16-18), suggesting a potential role of ABCC11 in the secretion of steroid metabolites from secretory cells within apocrine glands (16). Indeed, the transport activity of ABCC11 appears to be related to the size of the apocrine glands (5). The genetic polymorphism, on the other hand, has an impact on the *N*-linked glycosylation of ABCC11, intracellular sorting, and proteasomal degradation of the variant protein (5). The SNP variant (538A; Arg180), which lacks *N*-linked glycosylation, is recognised in the endoplasmic reticulum as a misfolded protein that is readily ubiquitinated and proteasomally degraded. Thus, the dry type of earwax was determined to be a Mendelian trait with a recessive phenotype (5). As a consequence, the SNP variant (538A or Arg180) fails to perform its transport function (4, 5), and the apocrine glands that are formed are notably small in size (5). Therefore, it is hypothesised that the function of ABCC11 *per se* or a metabolite transported by ABCC11 may stimulate the proliferation of apocrine gland cells. As far as the cell cycle machinery is operating normally, proliferation of apocrine gland cells should stop at certain levels. However, once somatic mutation has occurred in *BRCA1, BRCA2, p21,* or *p53* gene, deleterious and unregulated proliferation of those cells may start.

It has also been reported that ABCC11 is potentially involved in drug resistance in breast cancer. *ABCC11* mRNA is highly expressed in breast tumours (6, 12, 13), in particular, in invasive ductal adenocarcinomas (https://www.oncomine.org/resource/logn.html). Its expression is reportedly regulated by ER-α (19) and induced by 5-fluorouracil (5-FU) (20). In addition, it has been reported that ABCC11 is directly involved in 5-FU resistance by the efflux transport of the active metabolite 5-fluoro-2'-deoxyuridine 5'-monophosphate (FdUMP) (20-22). It remains to be elucidated, however, whether the expression of