10 Miyake, N., Harada, N., Shimokawa, O., Ohashi, H., Kurosawa, K., Matsumoto, T. *et al.* On the reported 8p22-p23.1 duplication in Kabuki make-up syndrome (KMS) and its absence in patients with typical KMS. *Am. J. Med. Genet.* **128A,** 170–172 (2004).
11 Hoffman, J. D., Zhang, Y., Greshock, J., Ciprero, K. L., Emanuel, B. S., Zackai, E. H. *et al.* Array based CGH and FISH fail to confirm duplication of 8p22-p23.1 in association with Kabuki syndrome. *J. Med. Genet.* **42,** 49–53 (2005).
12 Sanlaville, D., Genevieve, D., Bernardin, C., Amiel, J., Baumann, C., de Blois, M. C. *et al.* Failure to detect an 8p22–8p23.1 duplication in patients with Kabuki (Niikawa-Kuroki) syndrome. *Eur. J. Hum. Genet.* **13,** 690–693 (2005).
13 Kimberley, K. W., Morris, C. A. & Hobart, H. H. BAC-FISH refutes report of an 8p22–8p23.1 inversion or duplication in 8 patients with Kabuki syndrome. *BMC Med. Genet.* **7,** 46 (2006).
14 Maas, N. M., Van de Putte, T., Melotte, C., Francis, A., Schrander-Stumpel, C. T., Sanlaville, D. *et al.* The C20orf133 gene is disrupted in a patient with Kabuki syndrome. *J. Med. Genet.* **44,** 562–569 (2007).
15 Kuniba, H., Tsuda, M., Nakashima, M., Miura, S., Miyake, N., Kondoh, T. *et al.* Lack of C20orf133 and FLRT3 mutations in 43 patients with Kabuki syndrome in Japan. *J. Med. Genet.* **45,** 479–480 (2008).
16 Cuscó, I., del Campo, M., Vilardell, M., González, E., Gener, B., Galán, E. *et al.* Array-CGH in patients with Kabuki-like phenotype: identification of two patients with complex rearrangements including 2q37 deletions and no other recurrent aberration. *BMC Med. Genet.* **9,** 27 (2008).
17 Schoumans, J., Nordgren, A., Ruivenkamp, C., Brøndum-Nielsen, K., The, B. T., Annéren, G. *et al.* Genome-wide screening using array-CGH does not reveal microdeletions/microduplications in children with Kabuki syndrome. *Eur. J. Hum. Genet.* **13,** 260–263 (2005).
18 Miyake, N., Shimokawa, O., Harada, N., Sosonkina, N., Okubo, A., Kawara, H. *et al.* No detectable genomic aberrations by BAC Array CGH in Kabuki make-up syndrome patients. *Am. J. Med. Genet.* **140A,** 291–293 (2006).
19 Zhang, Z. F., Ruivenkamp, C., Staaf, J., Zhu, H., Barbaro, M., Petillo, D. *et al.* Detection of submicroscopic constitutional chromosome aberrations in clinical diagnostics: a validation of the practical performance of different array platforms. *Eur. J. Hum. Genet.* **16,** 786–792 (2008).
20 Chen, Y., Takita, J., Choi, Y. L., Kato, M., Ohira, M., Sanada, M. *et al.* Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* **455,** 971–975 (2008).
21 Nannya, Y., Sanada, M., Nakazaki, K., Hosoya, N., Wang, L., Hangaishi, A. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65,** 6071–6079 (2005).
22 Monis, P. T., Giglio, S. & Saint, C. P. Comparison of SYTO9 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of dye concentration on amplification and DNA melting curve analysis. *Anal. Biochem.* **340,** 24–34 (2005).
23 Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R. *et al.* PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* **13,** 2129–2141 (2003).
24 Marazita, M. L., Murray, J. C., Lidral, A. C., Arcos-Burgos, M., Cooper, M. E., Goldstein, T. *et al.* Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32–35. *Am. J. Hum. Genet.* **75,** 161–173 (2004).
25 Franke, L., de Kovel, C. G., Aulchenko, Y. S., Trynka, G., Zhernakova, A., Hunt, K. A. *et al.* Detection, imputation, and association analysis of small deletions and null alleles on oligonucleotide arrays. *Am. J. Hum. Genet.* **82,** 1316–1333 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

hybrid sterility involves both the unusual abundance and retention of OdsHmau protein in the *D. simulans* testis, as well as an unusual localization and possibly decondensation of the *D. simulans* Y chromosome. We conclude on the basis of these data that hybrid male sterility is caused by a gain-of-function interaction between OdsHmau and some component of the *D. simulans* Y chromosome heterochromatin, with this protein-DNA interaction representing the Dobzhansky-Muller incompatibility.

OdsH shares similarities with the hybrid sterility genes *Prdm9* (or *Meisetz*) in mouse (*23*) and *Overdrive* (*Ovd*) in *Drosophila* (*24*), all of which encode proteins with putative DNA-binding domains. Satellite DNAs have also been implicated in hybrid inviability, including a pericentric satellite locus (*Zhr*) (*25*, *26*) and a gene encoding a heterochromatin-binding protein (*Lhr*) (*27*). Thus, rapidly evolving repetitive DNA elements driven by genetic conflict may represent a major evolutionary force driving sequence divergence of speciation genes that would ultimately result in hybrid incompatibilities (*13*, *14*, *28*).

References and Notes

1. E. Mayr, *Systematics and the Origin of Species from the Viewpoint of a Zoologist* (Columbia Univ. Press, New York, 1942).
2. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer Associates, Sunderland, MA, 2004).
3. C. C. Laurie, *Genetics* **147**, 937 (1997).
4. R. M. Kliman *et al.*, *Genetics* **156**, 1913 (2000).
5. C. T. Ting, S. C. Tsaur, M. L. Wu, C. I. Wu, *Science* **282**, 1501 (1998).
6. S. Sun, C. T. Ting, C. I. Wu, *Science* **305**, 81 (2004).
7. D. E. Perez, C. I. Wu, *Genetics* **140**, 201 (1995).
8. D. E. Perez, C. I. Wu, N. A. Johnson, M. L. Wu, *Genetics* **134**, 261 (1993).
9. S. D. Hueber, I. Lohmann, *Bioessays* **30**, 965 (2008).
10. C. T. Ting *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12232 (2004).
11. K. Tabuchi, S. Yoshikawa, Y. Yuasa, K. Sawamoto, H. Okano, *Neurosci. Lett.* **257**, 49 (1998).
12. M. Nei, J. Zhang, *Science* **282**, 1428 (1998).
13. S. Henikoff, K. Ahmad, H. S. Malik, *Science* **293**, 1098 (2001).
14. S. Henikoff, H. S. Malik, *Nature* **417**, 227 (2002).
15. L. Fishman, A. Saunders, *Science* **322**, 1559 (2008).
16. A. Daniel, *Am. J. Med. Genet.* **111**, 450 (2002).
17. N. Aulner *et al.*, *Mol. Cell. Biol.* **22**, 1218 (2002).
18. M. Ashburner, K. G. Golic, R. S. Hawley, *Drosophila: A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 2, 2005).
19. G. Cenci, S. Bonaccorsi, C. Pisano, F. Verni, M. Gatti, *J. Cell Sci.* **107**, 3521 (1994).
20. B. D. McKee, *Curr. Top. Dev. Biol.* **37**, 77 (1998).
21. J. E. Tomkiel, *Genetica* **109**, 95 (2000).
22. J. Forejt, *Trends Genet.* **12**, 412 (1996).
23. O. Mihola, Z. Trachtulec, C. Vlcek, J. C. Schimenti, J. Forejt, *Science* **323**, 373 (2009).
24. N. Phadnis, H. A. Orr, *Science* **323**, 376 (2009).
25. K. Sawamura, M. T. Yamamoto, T. K. Watanabe, *Genetics* **133**, 307 (1993).
26. P. M. Ferree, D. A. Barbash, *PLoS Biol.* **7**, e1000234 (2009).
27. N. J. Brideau *et al.*, *Science* **314**, 1292 (2006).
28. H. S. Malik, S. Henikoff, *Cell* **138**, 1067 (2009).

29. We thank C-I. Wu for the *D. simulans* fertile and sterile introgression lines; C. Ting for scientific discussions and sharing data; G. Findlay for initial observations on OdsH cytology; and K. Ahmad, S. Biggins, N. Elde, S. Henikoff, N. Phadnis, T. Tsukiyama, and D. Vermaak for comments on the manuscript. Supported by NIH training grant PHS NRSA T32 GM07270 (J.J.B.), and grants from the Mathers foundation and NIH R01-GM74108 (H.S.M.). H.S.M. is an Early-Career Scientist of the Howard Hughes Medical Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1181756/DC1
Materials and Methods
Figs. S1 to S8
References

10 September 2009; accepted 13 October 2009
Published online 22 October 2009;
10.1126/science.1181756
Include this information when citing this paper.

Mapping Human Genetic Diversity in Asia

The HUGO Pan-Asian SNP Consortium*†

Asia harbors substantial cultural and linguistic diversity, but the geographic structure of genetic variation across the continent remains enigmatic. Here we report a large-scale survey of autosomal variation from a broad geographic sample of Asian human populations. Our results show that genetic ancestry is strongly correlated with linguistic affiliations as well as geography. Most populations show relatedness within ethnic/linguistic groups, despite prevalent gene flow among populations. More than 90% of East Asian (EA) haplotypes could be found in either Southeast Asian (SEA) or Central-South Asian (CSA) populations and show clinal structure with haplotype diversity decreasing from south to north. Furthermore, 50% of EA haplotypes were found in SEA only and 5% were found in CSA only, indicating that SEA was a major geographic source of EA populations.

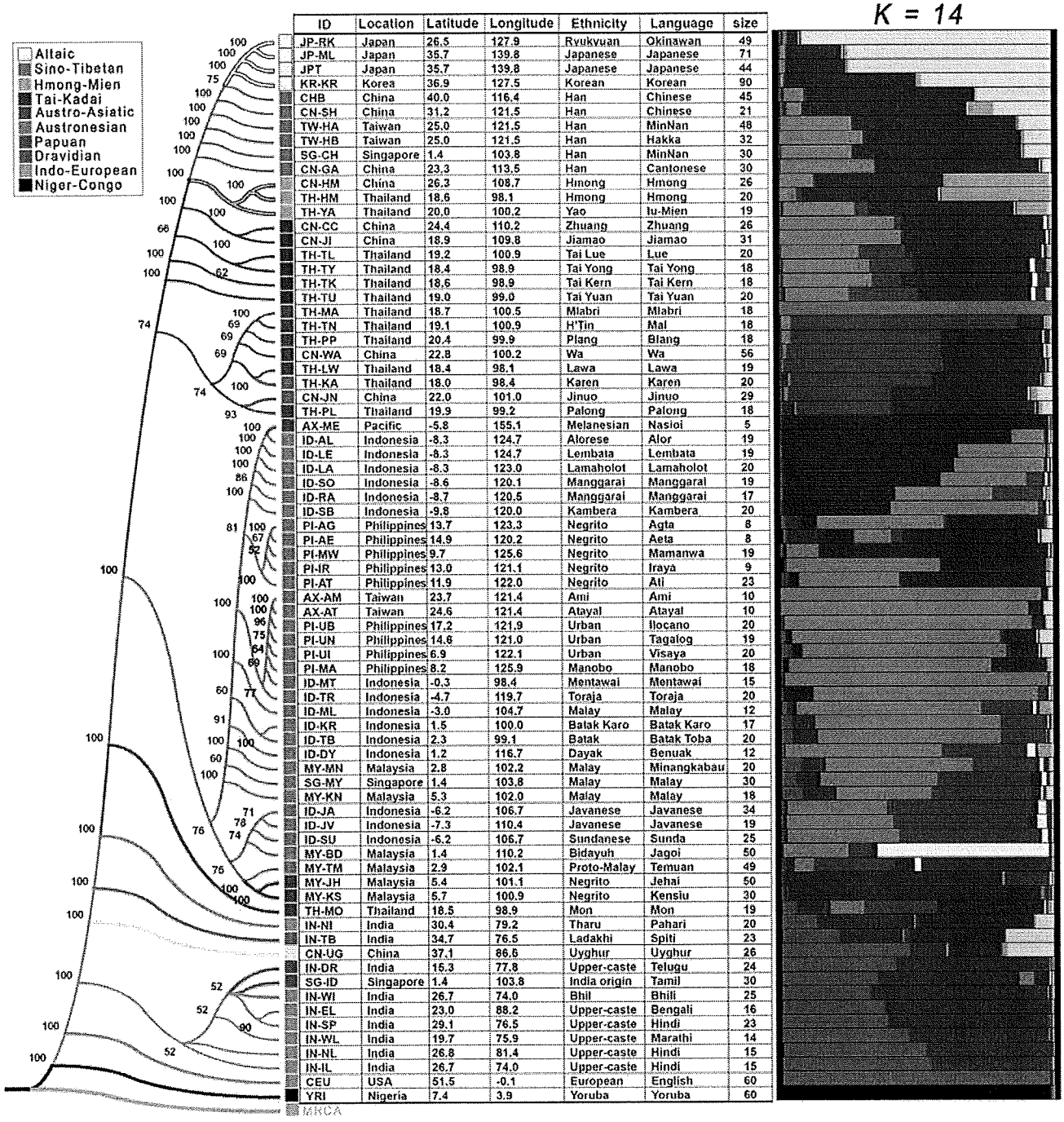
Several genome-wide studies of human genetic diversity focusing primarily on broad continental relationships, or fine-scale structure in Europe, have been published recently (*1*–*8*). We have extended this approach to Southeast Asian (SEA) and East Asian (EA) populations by using the Affymetrix GeneChip Human Mapping 50K Xba Array. Stringently quality-controlled genotypes were obtained at 54,794 autosomal single-nucleotide polymorphisms (SNPs) in 1928 individuals representing 73 Asian and two non-Asian HapMap populations (*9*). Apart from developing a general description of Asian population structure and its relation to geography, language, and demographic history, we concentrated on uncovering the geographic source(s) of EA and SEA populations.

We first performed a Bayesian clustering procedure using the STRUCTURE algorithm (*10*) to examine the ancestry of each individual. Each person is posited to derive from an arbitrary number of ancestral populations, denoted by K. We ran STRUCTURE from $K = 2$ to $K = 14$ using both the complete data set and SNP subsets to exclude those in strong linkage disequilibrium (Fig. 1 and figs. S1 to S13). At $K = 2$ and $K = 3$, all SEA and EA samples are united by predominant membership in a common cluster, with the other cluster(s) corresponding largely to Indo-European (IE) and African (AF) ancestries. At $K = 4$, a component most frequently found in Negrito populations that is also shared by all SEA populations emerges, suggesting a common SEA ancestry. Each value of K beyond 4 introduces a new component that tends to be associated with a group of populations united by membership in a linguistic family, by geographic proximity, by a known history of admixture, or, especially at higher Ks, by membership in a small population isolate. The results obtained using *frappe* (*11*), a maximum-likelihood–based clustering analysis, showed a general concordance with those of STRUCTURE (figs. S14 to S26). These analyses show that most individuals within a population share very similar ancestry estimates at all Ks, an observation that is consistent also with a phylogeny relating individuals (fig. S27) based on an allele-sharing distance (*12*). Therefore, we proceeded to evaluate the relationships among populations. A maximum-likelihood tree of populations, based on 42,793 SNPs whose ancestral states were known (Fig. 1), showed that all the SEA and EA populations make up a monophyletic clade that is supported by 100% of bootstrap replicates. This pattern remained even after data from 51 additional populations and 19,934 commonly typed SNPs from a recent study were integrated into the tree (fig. S28). These observations suggest that SEA and EA populations share a common origin.

STRUCTURE/*frappe* and principal components analyses (PCA) (*13*) (Figs. 1 and 2 and figs. S1 to S26) identify as many as 10 main population components. Each component corresponds largely to one of the five major linguistic groups (Altaic, Sino-Tibetan/Tai-Kadai, Hmong-Mien, Austro-Asiatic, and Austronesian), three ethnic categories (Philippine Negritos, Malaysian Negritos, and East Indonesians/Melanesians) and two small population isolates (the Bidayuh of Borneo and the hunter-gatherer Mlabri population of central and northern Thailand). The STRUCTURE results

*All authors with their affiliations appear at the end of this paper.
†To whom correspondence should be addressed. E-mail: ljin007@gmail.com (L.J.); liue@gis.a-star.edu.sg (E.T.L.); seielstadm@gis.a-star.edu.sg (M.S.); xushua@picb.ac.cn (S.X.)

(Fig. 1 and figs. S1 to S13), population phylogenies (Fig. 1 and figs. S27 and S28), and PCA results (Fig. 2) all show that populations from the same linguistic group tend to cluster together. A Mantel test confirms the correlation between linguistic and genetic affinities ($R^2 = 0.253$; $P < 0.0001$ with 10,000 permutations), even after controlling for geography (partial correlation = 0.136; $P < 0.005$ with 10,000 permutations). Nevertheless, we identified eight population outliers whose linguistic and genetic affinities are inconsistent [Affymetrix-Melanesian (AX-ME), Malaysia-Jehai (MY-JH)



ID	Location	Latitude	Longitude	Ethnicity	Language	size
JP-RK	Japan	26.5	127.9	Ryukyuan	Okinawan	49
JP-ML	Japan	35.7	139.8	Japanese	Japanese	71
JPT	Japan	35.7	139.8	Japanese	Japanese	44
KR-KR	Korea	36.9	127.5	Korean	Korean	90
CHB	China	40.0	116.4	Han	Chinese	45
CN-SH	China	31.2	121.5	Han	Chinese	21
TW-HA	Taiwan	25.0	121.5	Han	MinNan	48
TW-HB	Taiwan	25.0	121.5	Han	Hakka	32
SG-CH	Singapore	1.4	103.8	Han	MinNan	30
CN-GA	China	23.3	113.5	Han	Cantonese	30
CN-HM	China	26.3	108.7	Hmong	Hmong	26
TH-HM	Thailand	18.6	98.1	Hmong	Hmong	20
TH-YA	Thailand	20.0	100.2	Yao	Iu-Mien	19
CN-CC	China	24.4	110.2	Zhuang	Zhuang	26
CN-JI	China	18.9	109.8	Jiamao	Jiamao	31
TH-TL	Thailand	19.2	100.9	Tai Lue	Lue	20
TH-TY	Thailand	18.4	98.9	Tai Yong	Tai Yong	18
TH-TK	Thailand	18.6	98.9	Tai Kern	Tai Kern	18
TH-TU	Thailand	19.0	99.0	Tai Yuan	Tai Yuan	20
TH-MA	Thailand	18.7	100.5	Mlabri	Mlabri	18
TH-TN	Thailand	19.1	100.9	H'Tin	Mal	18
TH-PP	Thailand	20.4	99.9	Plang	Blang	18
CN-WA	China	22.8	100.2	Wa	Wa	56
TH-LW	Thailand	18.4	98.1	Lawa	Lawa	19
TH-KA	Thailand	18.0	98.4	Karen	Karen	20
CN-JN	China	22.0	101.0	Jinuo	Jinuo	29
TH-PL	Thailand	19.9	99.2	Palong	Palong	18
AX-ME	Pacific	-5.8	155.1	Melanesian	Nasioi	5
ID-AL	Indonesia	-8.3	124.7	Alorese	Alor	19
ID-LE	Indonesia	-8.3	124.7	Lembata	Lembata	19
ID-LA	Indonesia	-8.3	123.0	Lamaholot	Lamaholot	20
ID-SO	Indonesia	-8.6	120.1	Manggarai	Manggarai	19
ID-RA	Indonesia	-8.7	120.5	Manggarai	Manggarai	17
ID-SB	Indonesia	-9.8	120.0	Kambera	Kambera	20
PI-AG	Philippines	13.7	123.3	Negrito	Agta	8
PI-AE	Philippines	14.9	120.2	Negrito	Aeta	8
PI-MW	Philippines	9.7	125.6	Negrito	Mamanwa	19
PI-IR	Philippines	13.0	121.1	Negrito	Iraya	9
PI-AT	Philippines	11.9	122.0	Negrito	Ati	23
AX-AM	Taiwan	23.7	121.4	Ami	Ami	10
AX-AT	Taiwan	24.6	121.4	Atayal	Atayal	10
PI-UB	Philippines	17.2	121.9	Urban	Ilocano	20
PI-UN	Philippines	14.6	121.0	Urban	Tagalog	19
PI-UI	Philippines	6.9	122.1	Urban	Visaya	20
PI-MA	Philippines	8.2	125.9	Manobo	Manobo	18
ID-MT	Indonesia	-0.3	98.4	Mentawai	Mentawai	15
ID-TR	Indonesia	-4.7	119.7	Toraja	Toraja	20
ID-ML	Indonesia	-3.0	104.7	Malay	Malay	12
ID-KR	Indonesia	1.5	100.0	Batak Karo	Batak Karo	17
ID-TB	Indonesia	2.3	99.1	Batak	Batak Toba	20
ID-DY	Indonesia	1.2	116.7	Dayak	Benuak	12
MY-MN	Malaysia	2.8	102.2	Malay	Minangkabau	20
SG-MY	Singapore	1.4	103.8	Malay	Malay	30
MY-KN	Malaysia	5.3	102.0	Malay	Malay	18
ID-JA	Indonesia	-6.2	106.7	Javanese	Javanese	34
ID-JV	Indonesia	-7.3	110.4	Javanese	Javanese	19
ID-SU	Indonesia	-5.2	105.7	Sundanese	Sunda	25
MY-BD	Malaysia	1.4	110.2	Bidayuh	Jagoi	50
MY-TM	Malaysia	2.9	102.1	Proto-Malay	Temuan	49
MY-JH	Malaysia	5.4	101.1	Negrito	Jehai	50
MY-KS	Malaysia	5.7	100.9	Negrito	Kensiu	30
TH-MO	Thailand	18.5	98.9	Mon	Mon	19
IN-NI	India	30.4	79.2	Tharu	Pahari	20
IN-TB	India	34.7	76.5	Ladakhi	Spiti	23
CN-UG	China	37.1	86.6	Uyghur	Uyghur	26
IN-DR	India	15.3	77.8	Upper-caste	Telugu	24
SG-ID	Singapore	1.4	103.8	India origin	Tamil	30
IN-WI	India	26.7	74.0	Bhil	Bhili	25
IN-EL	India	23.0	88.2	Upper-caste	Bengali	16
IN-SP	India	29.1	76.5	Upper-caste	Hindi	23
IN-WL	India	19.7	75.9	Upper-caste	Marathi	14
IN-NL	India	26.8	81.4	Upper-caste	Hindi	15
IN-IL	India	26.7	74.0	Upper-caste	Hindi	15
CEU	USA	51.5	-0.1	European	English	60
YRI	Nigeria	7.4	3.9	Yoruba	Yoruba	60

Fig. 1. Maximum-likelihood tree of 75 populations. A hypothetical most-recent common ancestor (MRCA) composed of ancestral alleles as inferred from the genotypes of one gorilla and 21 chimpanzees was used to root the tree. Branches with bootstrap values less than 50% were condensed. Population identification numbers (IDs), sample collection locations with latitudes and longitudes, ethnicities, language spoken, and size of population samples are shown in the table adjacent to each branch in the tree. Linguistic groups are indicated with colors as shown in the legend. All population IDs except the four HapMap samples are denoted by four characters. The first two letters indicate the country where the samples were collected or (in the case of Affymetrix) genotyped, according to the following convention: AX, Affymetrix; CN, China; ID, Indonesia; IN, India; JP, Japan; KR, Korea; MY, Malaysia; PI, the Philippines; SG, Singapore; TH, Thailand; and TW, Taiwan. The last two letters are unique IDs for the population. To the right of the table, an averaged graph of results from STRUCTURE is shown for $K = 14$.

(Negrito), Malaysia-Kensiu (MY-KS) (Negrito), Thailand-Mon (TH-MO), Thailand-Karen (TH-KA), China-Jinuo (CN-JN), India-Spiti (IN-TB), and China-Uyghur (CN-UG); see table S3]. These linguistic outliers tend to cluster with their geographic neighbors or [especially evident in the principal component (PC) plots of Fig. 2] occupy an intermediate position between their geographic neighbors and the more-distant members of their linguistic group. These patterns are consistent either with substantial recent admixture among the populations (*14*–*16*), a history of language replacement (*17*), or uncertainties in the linguistic classifications themselves (for example, the controversial Altaic family, which groups Korean and Japanese with Uyghur).

Considerable gene flow among Asian populations was observed among subpopulations in these clusters, including those groups believed to practice endogamy based on linguistic, cultural, and ethnic information. In fact, most populations studied, even at lower *K*s, show evidence of admixture in the STRUCTURE analyses. For example, the Han Chinese have grown to become the largest ethnic group today in a demographic expansion that has occurred mostly within historical times. STRUCTURE reveals that the six Han Chinese population samples in our study show varying degrees of admixture (Fig. 1 and figs. S1 to S26) between a northern Altaic cluster and a Sino-Tibetan/Tai-Kadai cluster, which most frequently appears in the ethnic groups sampled from southern China and northern Thailand. Finally, most of the Indian populations showed evidence of shared ancestry with European populations, which is consistent with the recent observations (*18*) and our understanding of the expansion of Indo-European–speaking populations (Fig. 1 and figs. S1 to S26).

The geographic source(s) contributing to EA populations have long been debated. One hypothesis suggests that all SEA and EA populations derive primarily from a single initial migration, which entered the continent along a southern, largely coastal route (*19*, *20*). Another hypothesis argues for at least two independent migrations into East Asia, first along a southern route, followed later by a series of migrations along a more northern route that served to bridge European and EA populations, but with little contribution to populations in Southeast Asia (*20*). The topology of a maximum-likelihood tree (Fig. 1 and fig. S28) displays a largely south-to-north ordering of the populations, and a plot of the first two PCs (Fig. 2) similarly orients most populations according to their geographic coordinates. The average

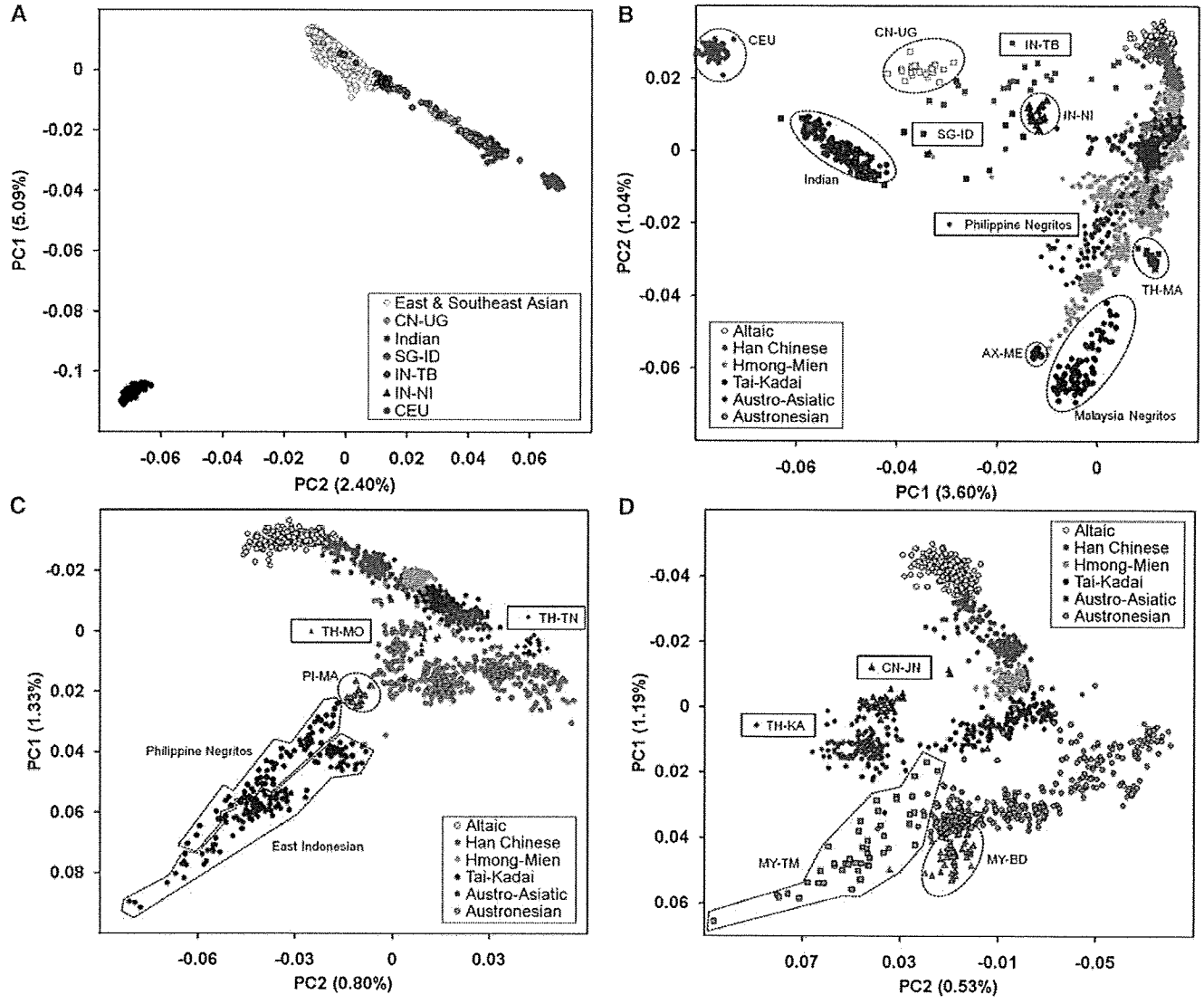
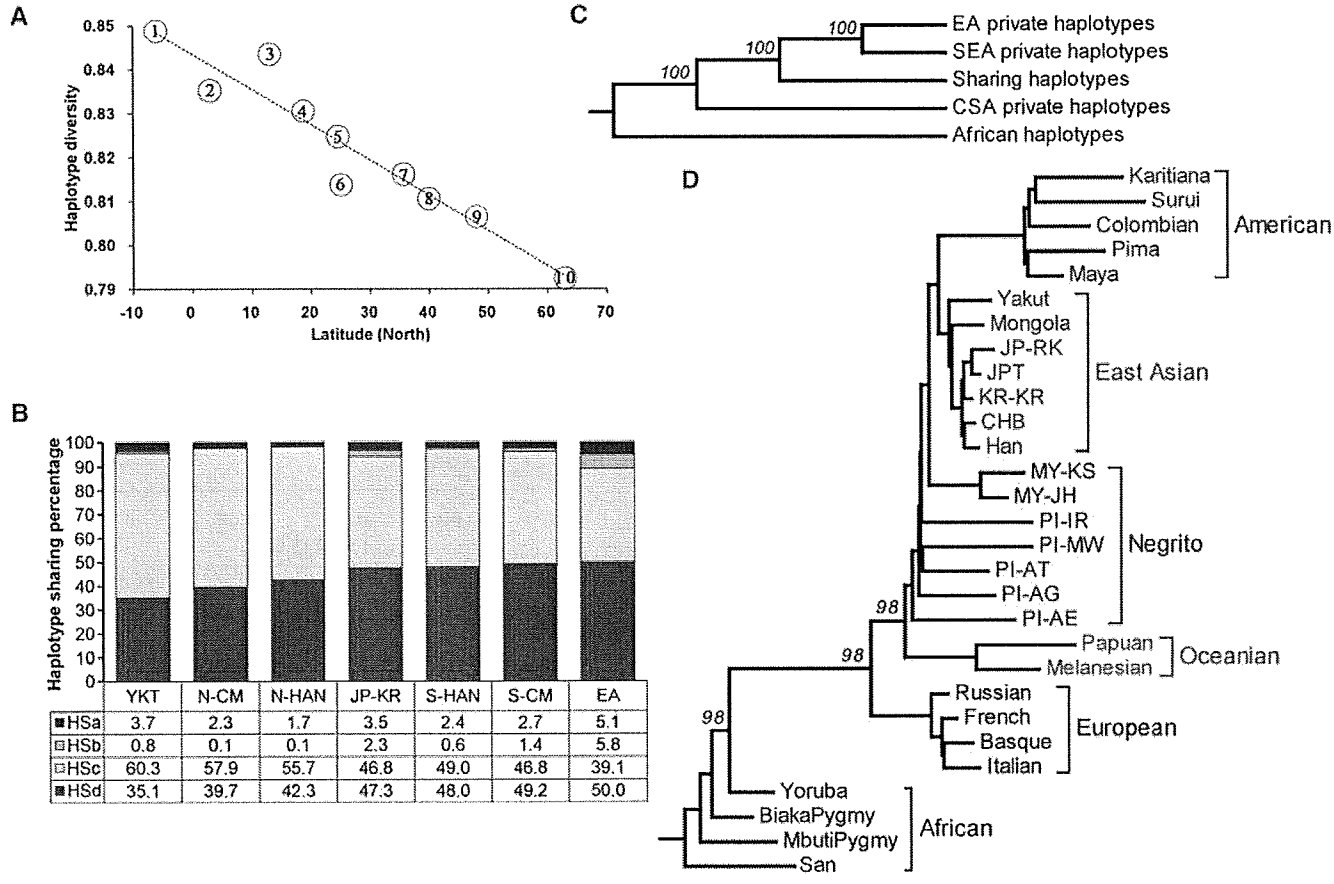


Fig. 2. Analysis of the first two PCs. **(A)** 1928 individuals representing all 75 populations. **(B)** 1868 individuals representing 74 populations (excluding YRI). **(C)** 1471 individuals representing 58 populations (excluding all Indians, CN-UG, TH-MA, AX-ME, and Negritos from Malaysia). **(D)** 1235 individuals representing 44 populations (excluding Philippine Negritos, PI-MA, and East Indonesians).

value of the first PC is highly correlated with the latitude at which the populations were sampled (R^2 = 0.79, P < 0.0001). Such a pattern could result simply from isolation-by-distance (IBD), as suggested by Ding *et al.* (*21*), although a recent study failed to detect IBD in East Asia with data from the Human Genome Diversity Project (*22*).

In an effort to distinguish between long-term historical divergence and the effects of IBD, we applied partial and multiple Mantel tests to the data (*23*) [see supporting online material (SOM) text for details]. The primary approach was to ascertain the differential correlation between genetic distance, geographical distance, and a group indicator matrix as an indication of prehistoric population divergence. The partial correlation coefficient of genetic and geographic distances was 0.228 (P < 0.0006), after controlling for the group indicator matrix (inferred from STRUCTURE/*frappe* analyses), whereas the partial correlation of the genetic and group indicator matrices was 0.403 (P < 0.0001) after controlling for geography. The superior association between genetic distance and the group indicator matrix as measured by the correlation coefficients suggests that prehistorical population divergence is the favored model over IBD in explaining the data (*24*). This conclusion is supported by simulation studies that also suggest that the observed patterns cannot be explained by simple IBD effects alone (see SOM text for details).

To further refine the analysis, we looked to haplotype organization to limit the effect of fluctuations in single-nucleotide determinations and to increase the resolution around genetic diversity. The IBD model predicts a correlation of genetic distance with geographical distance but not genetic diversity and geographic distance (*24*). By contrast, we found (Fig. 3A) that haplotype diversity is strongly correlated with latitude (R^2 = 0.91, P < 0.0001), with diversity decreasing from south to north, which is consistent with a loss of diversity as populations moved to higher latitudes. In estimating the contribution of SEA and Central-South Asian (CSA) haplotypes to the EA gene pool by haplotype sharing analyses (*16*), we found that more than 90% of haplotypes in EA populations could be found in SEA and CSA populations, of which about 50% were found in SEA and EA only and 5% found in CSA only (Fig. 3B, see also SOM text). Phylogenetic analysis of private haplotypes indicates greater similarity between EA and SEA populations relative to EA and CSA populations (Fig. 3C). These observations suggest that the geographic source(s) contributing to EA populations were mainly from SEA populations, with rather minor contributions from CSA,



	YKT	N-CM	N-HAN	JP-KR	S-HAN	S-CM	EA
HSa	3.7	2.3	1.7	3.5	2.4	2.7	5.1
HSb	0.8	0.1	0.1	2.3	0.6	1.4	5.8
HSc	60.3	57.9	55.7	46.8	49.0	46.8	39.1
HSd	35.1	39.7	42.3	47.3	48.0	49.2	50.0

Fig. 3. Analysis of haplotype diversity, haplotype sharing, and population phylogeny. **(A)** Haplotype diversity versus latitudes. Haplotypes were estimated from combined data, and diversity was measured by heterozygosity of haplotypes. HSa, b, c, and d and the corresponding colors show the percentages of EA group haplotypes in each class: HSa, found in CSA only; HSb, found in neither CSA nor SEA; HSc, found in both CSA and SEA; HSd, found in SEA only. Latitudes (*y* axis) for groups were obtained from the center of sample collection locations. Circled numbers are as follows: 1, Indonesian; 2, Malay; 3, Philippine; 4, Thai; 5, Southern Chinese minorities; 6, Southern Han Chinese; 7, Japanese and Korean; 8, Northern Han Chinese; 9, Northern Chinese minorities; and 10, Yakut. Haplotype heterozygosity of each group was estimated from 100-kb bins and taking together all haplotypes within each group. R^2 for the regression line is 0.91 (P < 0.0001). **(B)** Haplotype sharing analysis for EA populations and groups. YKT, Yakut; N-CM, Northern Chinese minorities; N-HAN, Northern Han Chinese; JP-KR, Japanese and Korean; S-HAN, Southern Han Chinese; S-CM, Southern Chinese minorities; EA, East Asian. **(C)** Phylogeny of group private haplotypes. EA private haplotypes: haplotypes found only in EA samples; SEA private haplotypes: haplotypes found only in SEA samples; CSA private haplotypes: haplotypes found only in CSA samples; Shared haplotypes: haplotypes found in all EA, SEA, and CSA samples; African haplotypes were used as outgroup. **(D)** Maximum-likelihood tree of 29 populations. The tree is based on data from 19,934 SNPs. Bootstrap values were based on 100 replicates. Only values on splitting of African and non-African, European and Oceanian and Asian, and Oceanian and Asian are shown.

and that this clinal structure of EA populations arose from prehistoric population divergence rather than IBD or gene flow from CSA populations.

On the basis of increased cultural, linguistic, and genetic diversity, the origins of SEA populations are thought to be more complex than the origins of those to their north. Notably, the Negritos of the Philippines and Malaysia differ from neighboring populations in aspects of their physical appearance, prompting intense speculation about models of human settlement in Southeast Asia. The two-wave hypothesis, which suggests that ancestral Negrito populations settled in Southeast Asia, Australia, and Oceania before a more northerly migration originating in or near the Middle East, and spreading both toward Europe and Northeast Asia via Central Asia (*25*), has been supported by phylogenetic trees constructed from data on a limited number of protein markers (*24*, *25*). The topology of our population trees, both with and without the data from additional European and Asian populations discussed in (*1*), is inconsistent with regard to this genetic similarity of European and EA populations (Figs. 1 and 3D). Instead, on the basis of variation at a large number of independent SNPs, we observed that there is substantial genetic proximity of SEA and EA populations (fig. S28). An identical pattern is seen in the population tree of Li *et al.* (*1*) based on all of their 642,690 SNPs. Our forward-time simulation results under extreme ascertainment scenarios (SOM text) show that the observed phylogeny is not the result of ascertainment bias. Simulation studies also suggest that substantial levels of migration between populations after their initial separation are unlikely to distort the topology of the phylogeny (SOM text).

To unambiguously infer population histories represents a considerable challenge (*26*). Although this study does not disprove a two-wave model of migration, the evidence from our autosomal data and the accompanying simulation studies (figs. S29 and S30) point toward a history that unites the Negrito and non-Negrito populations of Southeast and East Asia via a single primary wave of entry of humans into the continent.

References and Notes

1. J. Z. Li *et al.*, *Science* **319**, 1100 (2008).
2. M. Kayser *et al.*, *Am. J. Hum. Genet.* **82**, 194 (2008).
3. N. A. Rosenberg *et al.*, *PLoS Genet.* **1**, e70 (2005).
4. N. A. Rosenberg *et al.*, *Science* **298**, 2381 (2002).
5. J. Novembre *et al.*, *Nature* **456**, 98 (2008).
6. M. Nelis *et al.*, *PLoS One* **4**, e5472 (2009).
7. C. Tian *et al.*, *PLoS Genet.* **4**, e4 (2008).
8. O. Lao *et al.*, *Curr. Biol.* **18**, 1241 (2008).
9. The International HapMap Consortium, *Nature* **426**, 789 (2003).
10. J. K. Pritchard, M. Stephens, P. Donnelly, *Genetics* **155**, 945 (2000).
11. H. Tang, J. Peng, P. Wang, N. J. Risch, *Genet. Epidemiol.* **28**, 289 (2005).
12. J. L. Mountain, L. L. Cavalli-Sforza, *Am. J. Hum. Genet.* **61**, 705 (1997).
13. N. Patterson, A. L. Price, D. Reich, *PLoS Genet.* **2**, e190 (2006).
14. S. Xu, L. Jin, *Am. J. Hum. Genet.* **83**, 322 (2008).
15. S. Xu, W. Huang, J. Qian, L. Jin, *Am. J. Hum. Genet.* **82**, 883 (2008).
16. S. Xu, W. Jin, L. Jin, *Mol. Biol. Evol.* **26**, 2197 (2009).
17. L. Reid, in *Language Contact and Change in the Austronesian World.* T. Dutton, T. Tryon, Eds. (Mouton de Gruyter, Berlin, 1994) pp. 443–475.
18. Indian Genome Variation Consortium, *J. Genet.* **87**, 3 (2008).
19. J. Y. Chu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11763 (1998).
20. B. Su *et al.*, *Am. J. Hum. Genet.* **65**, 1718 (1999).
21. Y. C. Ding *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14003 (2000).
22. A. Manica, F. Prugnolle, F. Balloux, *Hum. Genet.* **118**, 366 (2005).
23. M. P. Telles, J. A. Diniz-Filho, *Genet. Mol. Res.* **4**, 742 (2005).
24. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1993).
25. L. L. Cavalli-Sforza, M. W. Feldman, *Nat. Genet.* **33**, 266 (2003).
26. G. Hellenthal, A. Auton, D. Falush, *PLoS Genet.* **4**, e1000078 (2008).

27. The entire consortium thanks all individuals who volunteered their DNA for this project. It is this collaboration between scientists and the public that is essential to progress in our field. All SNP data have been submitted to dbSNP with the submission handle PASNPI and will become accessible in dbSNP Build 131. See SOM text for a complete listing of all acknowledgments.

The HUGO Pan-Asian SNP Consortium

Mahmood Ameen Abdulla,[1] Ikhlak Ahmed,[2] Anunchai Assawamakin,[3,4] Jong Bhak,[5] Samir K. Brahmachari,[2] Gayvelline C. Calacal,[6] Amit Chaurasia,[2] Chien-Hsiun Chen,[7] Jieming Chen,[8] Yuan-Tsong Chen,[7] Jiayou Chu,[9] Eva Maria C. Cutiongco-de la Paz,[10] Maria Corazon A. De Ungria,[6] Frederick C. Delfin,[6] Juli Edo,[1] Suthat Fuchareon,[3] Ho Ghang,[5] Takashi Gojobori,[11,12] Junsong Han,[13] Sheng-Feng Ho,[7] Boon Peng Hoh,[14] Wei Huang,[15] Hidetoshi Inoko,[16] Pankaj Jha,[2] Timothy A. Jinam,[1] Li Jin,[17,38]† Jongsun Jung,[18] Daoroong Kangwanpong,[19] Jatupol Kampuansai,[19] Giulia C. Kennedy,[20,21] Preeti Khurana,[22] Hyung-Lae Kim,[18] Kwangjoong Kim,[18] Sangsoo Kim,[23] Woo-Yeon Kim,[5] Kuchan Kimm,[24] Ryosuke Kimura,[25] Tomohiro Koike,[11] Supasak Kulawonganunchai,[4] Vikrant Kumar,[8] Poh San Lai,[26,27] Jong-Young Lee,[18] Sunghoon Lee,[5] Edison T. Liu,[8]† Partha P. Majumder,[28] Kiran Kumar Mandapati,[22] Sangkot Marzuki,[29] Wayne Mitchell,[30,31] Mitali Mukerji,[2] Kenji Naritomi,[32] Chumpol Ngamphiw,[4] Norio Niikawa,[40] Nao Nishida,[25] Bermseok Oh,[18] Sangho Oh,[5] Jun Ohashi,[25] Akira Oka,[16] Rick Ong,[8] Carmencita D. Padilla,[10] Prasit Palittapongarnpim,[33] Henry B. Perdigon,[6] Maude Elvira Phipps,[1,34] Eileen Png,[8] Yoshiyuki Sakaki,[35] Jazelyn M. Salvador,[6] Yuliana Sandraling,[29] Vinod Scaria,[2] Mark Seielstad,[8]† Mohd Ros Sidek,[14] Amit Sinha,[2] Metawee Srikummool,[19] Herawati Sudoyo,[29] Sumio Sugano,[37] Helena Suryadi,[29] Yoshiyuki Suzuki,[11] Kristina A. Tabbada,[6] Adrian Tan,[8] Katsushi Tokunaga,[25] Sissades Tongsima,[4] Lilian P. Villamor,[6] Eric Wang,[20,21] Ying Wang,[15] Haifeng Wang,[15] Jer-Yuarn Wu,[7] Huasheng Xiao,[13] Shuhua Xu,[38]† Jin Ok Yang,[5] Yin Yao Shugart,[39] Hyang-Sook Yoo,[5] Wentao Yuan,[15] Guoping Zhao,[15] Bin Alwi Zilfalil,[14] Indian Genome Variation Consortium[2]

[1]Department of Molecular Medicine, Faculty of Medicine, and the Department of Anthropology, Faculty of Arts and Social Sciences, University of Malaya, Kuala Lumpur, 50603, Malaysia. [2]Institute of Genomics and Integrative Biology, Council for Scientific and Industrial Research, Mall Road, Delhi 110007, India. [3]Mahidol University, Salaya Campus, 25/25 M. 3, Puttamonthon 4 Road, Puttamonthon, Nakornpathom 73170, Thailand. [4]Biostatistics and Informatics Laboratory, Genome Institute, National Center for Genetic Engineering and Biotechnology, Thailand Science Park, Pathumtani 12120, Thailand. [5]Korean BioInformation Center (KOBIC), Korea Research Institute of Bioscience and Biotechnology (KRIBB), 111 Gwahangno, Yuseong-gu, Deajeon 305-806, Korea. [6]DNA Analysis Laboratory, Natural Sciences Research Institute, University of the Philippines, Diliman, Quezon City 1101, Philippines. [7]Institute of Biomedical Sciences, Academia Sinica, 128 Sec 2 Academia Road Nangang, Taipei City 115, Taiwan. [8]Genome Institute of Singapore, 60 Biopolis Street 02-01, 138672, Singapore. [9]Institute of Medical Biology, Chinese Academy of Medical Science, Kunming, China. [10]Institute of Human Genetics, National Institutes of Health, University of the Philippines Manila, 625 Pedro Gil Street, Ermita Manila 1000, Philippines. [11]Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Research Organization of Information and Systems, 1111 Yata, Mishima, Shizuoka 411-8540, Japan. [12]Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan. [13]National Engineering Center for Biochip at Shanghai, 151 Li Bing Road, Shanghai 201203, China. [14]Human Genome Center, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia. [15]MOST-Shanghai Laboratory of Disease and Health Genomics, Chinese National Human Genome Center Shanghai, 250 Bi Bo Road, Shanghai 201203, China. [16]Department of Molecular Life Science Division of Molecular Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara-A Kanagawa-Pref A259-1193, Japan. [17]State Key Laboratory of Genetic Engineering and MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, 220 Handan Road, Shanghai 200433, China. [18]Korea National Institute of Health, 194, Tongil-Lo, Eunpyung-Gu, Seoul, 122-701, Korea. [19]Department of Biology, Faculty of Science, Chiang Mai University, 239 Huay Kaew Road, Chiang Mai 50202, Thailand. [20]Genomics Collaborations, Affymetrix, 3420 Central Expressway, Santa Clara, CA 95051, USA. [21]Veracyte, 7000 Shoreline Court, Suite 250, South San Francisco, CA 94080, USA. [22]The Centre for Genomic Applications (an IGIB-IMM Collaboration), 254 Ground Floor, Phase III Okhla Industrial Estate, New Delhi 110020, India. [23]Soongsil University, Sangdo-5-dong 1-1, Dongjak-gu, Seoul 156-743, Korea. [24]Eulji University College of Medicine, 143-5 Yong-du-dong Jung-gu, Dae-jeon City 301-832, Korea. [25]Department of Human Genetics, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [26]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, National University Hospital, 5 Lower Kent Ridge Road, 119074, Singapore. [27]Population Genetics Lab, Defence Medical and Environmental Research Institute, DSO National Laboratories, 27 Medical Drive, 117510, Singapore. [28]Indian Statistical Institute (Kolkata) 203 Barrackpore Trunk Road, Kolkata 700108, India. [29]Eijkman Institute for Molecular Biology, Jl. Diponegoro 69, Jakarta 10430, Indonesia. [30]Informatics Experimental Therapeutic Centre, 31 Biopolis Way, 03-01 Nanos, 138669, Singapore. [31]Division of Information Sciences, School of Computer Engineering, Nanyang Technological University, 50 Nanyang Avenue, 639798, Singapore. [32]Department of Medical Genetics, University of the Ryukyus Faculty of Medicine, Nishihara, 207 Uehara, Okinawa 903-0215, Japan. [33]National Science and Technology Development Agency, 111 Thailand Science Park, Pathumtani 12120, Thailand. [34]Monash University (Sunway Campus), Jalan Lagoon Selatan, 46150 Bandar Sunway, Selangor, Malaysia. [35]RIKEN Genomic Sciences Center, W502, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. [36]Department of Biochemistry, University of Hong Kong, 3/F Laboratory Block, Faculty of Medicine Building, 21 Sasson Road, Pokfulam, Hong Kong. [37]Laboratory of Functional Genomics, Department of Medical Genome Sciences Graduate School of Frontier Sciences, University of Tokyo (Shirokanedai Laboratory), 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [38]Chinese Academy of Sciences-Max Planck Society Partner Institute for Computational Biology, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, 320 Yueyang Rd., Shanghai 200031, China. [39]Genomic Research Branch, National Institute of Mental Health, National Institutes of Health, 6001 Executive Boulevard, Bethesda, MD 20892 USA. [40]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Tobetsu 061-0293, Japan.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5959/1541/DC1
Materials and Methods
SOM Text
Figs. S1 to S38
Tables S1 to S4

1 June 2009; accepted 13 October 2009
10.1126/science.1177074

The FASEB Journal • Research Communication

Earwax, osmidrosis, and breast cancer: why does one SNP (538G>A) in the human ABC transporter *ABCC11* gene determine earwax type?

Yu Toyoda,* Aki Sakurai,*,† Yasumasa Mitani,†,‡ Masahiro Nakashima,§ Koh-ichiro Yoshiura,‖ Hiroshi Nakagawa,* Yasuo Sakai,¶ Ikuko Ota,†,# Alexander Lezhava,† Yoshihide Hayashizaki,† Norio Niikawa,‖, and Toshihisa Ishikawa*,†,1**

*Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan; †Omics Science Center (OSC), RIKEN Yokohama Institute, Yokohama, Japan; ‡K.K. DNAFORM, Yokohama, Japan; §Tissue and Histopathology Section, Atomic Bomb Disease Institute, and ‖Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; ¶Department of Plastic and Reconstructive Surgery, Fujita Health University, Toyoake, Japan; #Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, Yokohama, Japan; and **Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

ABSTRACT **One single-nucleotide polymorphism (SNP), 538G>A (Gly180Arg), in the *ABCC11* gene determines the type of earwax. The G/G and G/A genotypes correspond to the wet type of earwax, whereas A/A corresponds to the dry type. Wide ethnic differences exist in the frequencies of those alleles, reflecting global migratory waves of the ancestors of humankind. We herein provide the evidence that this genetic polymorphism has an effect on the *N*-linked glycosylation of ABCC11, intracellular sorting, and proteasomal degradation of the variant protein. Immunohistochemical studies with cerumen gland-containing tissue specimens revealed that the ABCC11 WT protein was localized in intracellular granules and large vacuoles, as well as at the luminal membrane of secretory cells in the cerumen gland, whereas granular or vacuolar localization was not detected for the SNP (Arg180) variant. This SNP variant lacking *N*-linked glycosylation is recognized as a misfolded protein in the endoplasmic reticulum and readily undergoes ubiquitination and proteasomal degradation, which determines the dry type of earwax as a mendelian trait with a recessive phenotype. For rapid genetic diagnosis of axillary osmidrosis and potential risk of breast cancer, we developed specific primers for the SmartAmp method that enabled us to clinically genotype the *ABCC11* gene within 30 min.—Toyoda, Y., Sakurai, A., Mitani, Y., Nakashima, M., Yoshiura, K., Nakagawa, H., Sakai, Y., Ota, I., Lezhava, A., Hayashizaki, Y., Niikawa, N., Ishikawa, T. Earwax, osmidrosis, and breast cancer: why does one SNP (538G>A) in the human ABC transporter *ABCC11* gene determine earwax type? *FASEB J.* 23, 2001–2013 (2009)**

Key Words: endoplasmic reticulum-associated degradation · genetic polymorphism · N-linked glycosylation · protein instability

CELL SECRETION IS AN IMPORTANT physiological process that ensures smooth metabolic activities and tissue repair as well as growth and immunological functions in the body. Human earwax is a dimorphic trait consisting of wet or dry types. Wet earwax (cerumen) is a secretory product of ceruminous apocrine glands, whereas the dry type of earwax lacks oily components. In addition, apocrine glands can be found in the external auditory canal, breast, and axillary region; those physical characteristics also are concerned with apocrine glands. A positive association among the wet earwax type, axillary osmidrosis (1), colostrum secretion (2), and breast cancer risk (3, 4) has been suggested by phenotype-based analysis. Apocrine secretion occurs when the secretory process is accomplished with a partial loss of cell cytoplasm. The secretory materials may be contained within the secretory vesicles or dissolved in the cytoplasm, and during secretion they are released as cytoplasmic fragments into the glandular lumen or interstitial space (5).

Hitherto apocrine secretory mechanisms have not been well characterized (5). Although the biochemical and physiological pathways that regulate the apocrine secretory process are not clearly known, our recent finding that the nonsynonymous SNP 538G>A (rs17822931; Gly180Arg) in the *ABCC11* gene determines the type of earwax has shed light on the novel function of this ABC transporter in apocrine glands (6). In 2001, we and other groups independently cloned human ABCC11 and ABCC12 cDNAs (7-9). Both *ABCC11* and *ABCC12* genes are located on human chromosome 16q12.1 in a tail-to-head orientation with a separation distance of ~20 kb (7). Interestingly, there is an orthologous gene corresponding to

[1] Correspondence: Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259-B-60 Nagatsuta, Midori-ku, Yokohama 226-8501, Japan. E-mail: tishikaw@bio.titech.ac.jp

doi: 10.1096/fj.09-129098

the *ABCC12* gene in mammals (10, 11), but no gene orthologous to human *ABCC11* (11) has been found in mammals except for primates. Based on a database search, it is speculated that *ABCC11* is a paralogous gene generated by *ABCC12* gene duplication in the primate genome, including that of humans (12). Human ABCC11 reportedly functions as an ATP-dependent efflux pump for amphipathic anions, including cyclic nucleotides, leukotriene C_4, estrone 3-sulfate, dehydroepiandrosterone 3-sulfate (DHEAS), estradiol 17-β-D-glucuronide, and folic acid (13-16).

We have recently reported that one single-nucleotide polymorphism (SNP), 538G>A (Gly180Arg), in the *ABCC11* gene determines the type of earwax (6). This is the first example of a DNA polymorphism determining a visible genetic trait. The G/G and G/A genotypes correspond to the wet type of earwax, whereas A/A corresponds to the dry type. A functional assay has demonstrated that cells with allele A show a lower excretory activity for cGMP than those with allele G (6). Wide ethnic differences were also observed in the frequencies of those alleles (6). Furthermore, it has most recently been demonstrated that the genotype frequencies in the *ABCC11* and *EDAR* genes differ significantly between the Hondo and Ryukyu clusters in Japan (17).

At the present time, however, little is known about how genetic polymorphisms of the *ABCC11* gene affect the function of the ceruminous apocrine gland. The present study has two specific aims. First, to understand the molecular mechanism determining the earwax type, we investigated the effect of the SNP 538G>A (Gly180Arg) on the protein expression and intracellular localization of ABCC11. Second, to gain insight into the link between the genetic polymorphisms of *ABCC11* and the functions of apocrine glands, such as cerumen secretion and axillary osmidrosis, we designed a primer set for clinical genotyping.

In this article, we provide evidence that the SNP 538G>A (Gly180Arg) variant of human ABCC11 lacking *N*-linked glycosylation is recognized as a misfolded protein in the endoplasmic reticulum (ER) and readily undergoes proteasomal degradation. This ABCC11 protein degradation underlies the molecular mechanism to form the dry type of earwax as a mendelian trait with a recessive phenotype. Furthermore, we demonstrate that the wet type of earwax is genetically linked with axillary osmidrosis, which is recognized as a disease covered by the national health insurance system in Japan. Human ABCC11 is suggested to play a pivotal role in the secretion of steroid metabolites from secretory cells in the apocrine glands.

MATERIALS AND METHODS

Collection of genomic DNA from volunteers and preparation of DNA samples

Under written informed consent, we collected blood samples from 124 Japanese volunteers at Nagasaki University (Nagasaki, Japan) during a period from January 2004 to December 2005. Genomic DNA was extracted from whole blood by the standard method. Protocols for the present study were approved by the Committee for the Ethical Issues on Human Genome and Gene Analysis, Nagasaki University. The clinical investigation was conducted according to the Declaration of Helsinki Principles. Briefly, all blood samples were collected in standard 2Na-EDTA-coated blood collection tubes. The samples were subjected to proteinase K digestion, and then genomic DNA was isolated by phenol/chloroform extraction and subsequently by ethanol precipitation. Part of the *ABCC11* gene including the 538G>A alleles was amplified by PCR with the following primers: EWX1, 5′-TGCAAAGAGATTCCACCAGTT-3′; and EWX2, 5′-AAGGTCTTCATTTTCTAGACAGC-3′.

Diagnosis of axillary osmidrosis in Japanese patients and genotyping

All procedures for the diagnosis of axillary osmidrosis patients and their genotyping were performed according to the protocol approved by the Ethical Review Board of Fujita Health University School of Medicine (Nagoya, Japan). The clinical investigation was conducted according to the Declaration of Helsinki Principles. Fourteen axillary osmidrosis patients participated in this study under written informed consent. All of the patients expressed their subjective views on symptoms by answering survey inquiries. Axillary osmidrosis was diagnosed basically from both the self-declaration of patients and the odor-smelling test carried out by authorized medical doctors. Earwax type was determined from wax samples obtained by stirring a cotton swab in the external auditory canal of the patient. Individual genomic DNA was obtained from 2 ml of whole blood by using QuickGene-610L (Fujifilm Co., Tokyo, Japan) according to the manufacturer's protocol. To sequence the SNP 538G>A, the following PCR primers were designed: P1, 5′-TGTCACATGCAAAGAGATTCC-3′; and P2, 5′-CTCCTGGCATGGACTTGAACA-3′. To identify the Δ27 mutation, we designed one set of primers: P3, 5′-AGGTCTCTAGGGCCTGAAGTA-3′; and P4, 5′-AGCCTTCACCTTCCCATTGCC-3′. The experiments to determine the genotype of each patient with *ABCC11* gene amplification by PCR and DNA sequencing were carried out in a masked manner.

DNA sequence analysis

DNA sequence analysis was carried out by using BigDye Terminator 3.1 (Applied Biosystems Inc., Foster City, CA, USA) after ExoSAP-IT treatment, and with an Autosequencer Model 3100 (Applied Biosystems Inc.), according to the manufacturer's protocol. The sequences were aligned with an AutoAssembler (Applied Biosystems Inc.) and visualized with Sequencher 4.7 Demo (Hitachi Software Engineering Co., Ltd., Tokyo, Japan) to find SNPs or mutations.

Data analysis

SNP data on the polymorphisms of ABCC11 were obtained from the dbSNP database of the National Center for Biotechnology Information (NCBI; Bethesda, MD, USA). The hydropathy profile of the ABCC11 protein deduced from the cDNA sequence was calculated with the Kyte-Doolittle hydropathy algorithm (18). Potential *N*-linked glycosylation sites were predicted with Genetyx-Win 5.1. (Software Development Co., Ltd., Tokyo, Japan)

Polyclonal antibody against ABCC11

The cytoplasmic domain (aa 746–804) close to the first ATP-binding cassette, located between transmembrane domains 6 and 7 (TM6 and TM7) of the ABCC11 protein, was selected as an epitope for development of the polyclonal antibody. The epitope-encoding cDNA (2236-2412) was inserted into the pGEX-2T vector (GE Healthcare, Little Chalfont, UK), and then *Escherichia coli* strain BL21 [F^-, *ompT*, *hsd*S(r_B^-, m_B^-), gal] cells were transformed with the vector to produce a glutathione-*S*-transferase (GST)-like fusion peptide. The expressed fusion peptide was isolated by using a glutathione-affinity gel column (Pierce, Rockford, IL, USA) according to the manufacturer's protocol. The epitope peptide was then recovered by PreScission Protease digestion (GE Healthcare). The anti-ABCC11 polyclonal antibody was produced by immunization of rabbits with the epitope peptide. The antiserum was purified through an affinity column and used for immunohistochemistry and immunoblotting studies.

Immunohistochemical detection of ABCC11 expressed in the ceruminous apocrine gland

External auditory canal tissues were obtained, with informed consent, from a patient undergoing surgical excision of a squamous cell carcinoma around the ear at Nagasaki University. The tissue samples were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS, pH 7.4) at 4°C overnight, embedded in paraffin, and sectioned at 4 μm in thickness. After antigen retrieval by microwave treatment in 0.01 M citrate buffer (pH 6.0), deparaffinized sections were preincubated with 10% normal goat serum. Thereafter, tissues were reacted with anti-ABCC11 rabbit polyclonal antibody at 1:500 dilution. The slides were subsequently incubated with Alexa Fluor 488-conjugated goat anti-rabbit IgG antibody (Invitrogen Co., Carlsbad, CA, USA). Specimens were counterstained with 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI-I; Vysis Inc., Downers Grove, IL, USA), visualized, and photographed with a fluorescence microscope (Zeiss Axioplan2; Carl Zeiss Japan, Tokyo, Japan) equipped with a charge-coupled device camera (19).

Expression of ABCC11 in Sf9 insect cells

The cDNA of human ABCC11 WT was inserted into the pFastBac1 vector (Invitrogen) between the restriction enzyme sites of *Xba*I and *Hind*III. To express ABCC11 in insect cells, recombinant baculoviruses were generated with the Bac-to-Bac baculovirus expression systems (Invitrogen) according to the manufacturer's protocols. Insect *Spodoptera frugiperda* Sf9 cells (1.0×10^6 cells/ml) were infected with the recombinant baculovirus and cultured in NIM-Ex Insect serum-free medium (Nosan Co., Yokohama, Japan) supplemented with penicillin (100 U/ml), and streptomycin (100 μg/ml) (Invitrogen) with gentle shaking at 28°C. Three days after infection, cells were harvested by centrifugation. To prepare the whole-cell lysate, cells were subsequently washed with PBS and collected by centrifugation.

Generation of ABCC11 variant forms

The human ABCC11 WT or G180R cDNA was inserted into the pcDNA3.1/Hygro(−) vector (Invitrogen) between the restriction enzyme sites of *Xho*I/*Sal*I and *Hind*III, respectively. The resulting expression construct [ABCC11 WT-pcDNA3.1/Hygro(−)] was used as the template for site-directed mutagenesis to obtain ABCC11 variants, *i.e.*, N838Q, N844Q, and N838Q/N844Q. To generate Asn-to-Gln variant forms of ABCC11, a codon (AAT) encoding asparagine was converted to CAA by using *Pfu Turbo* DNA polymerase and the QuikChange site-directed mutagensis kit (Stratagene, La Jolla, CA, USA) according to the manufacturer's protocol. The sequences of internal complementary PCR primers used for the site-directed mutagenesis, as well as the corresponding PCR conditions, are summarized in Supplemental Table 1. To generate the N838Q/N844Q variant, ABCC11 N838Q-pcDNA3.1/Hygro(−) was used as the template. Each variant cDNA generated in the pcDNA3.1/Hyg(−) plasmid was subjected to nucleotide sequence analysis (Shimadzu Co., Kyoto, Japan). To substitute Gly180 to Arg, Lys, His, Asp, Glu, AL, or Pro in the ABCC11 WT protein, the codon (GGG) encoding the Gly residue in TM1 was changed by site-directed mutagenesis, as described above. The PCR primers used for the site-directed mutagenesis are shown in Supplemental Table 2. Furthermore, hitherto known nonsynonymous SNP variants of ABCC11 were generated in the same manner as described above (see Supplemental Table 3 for the PCR primers used for site-directed mutagenesis).

Expression of ABCC11 WT and variants in Flp-In-293 cells

Flp-In-293 cells (Invitrogen) were maintained in high-glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) heat-inactivated fetal calf serum (FCS), 2 mM L-glutamine, penicillin (100 U/ml), and streptomycin (100 μg/ml) in a humidified atmosphere at 37°C with 5% CO_2 in air, as described previously (20-22). Flp-In-293 cells were replated onto 35-mm dishes at a concentration of 1.0×10^6 cells/dish. After 24 h, the cells were transfected with the ABCC11-pcDNA3.1/Hyg(−) vector by using Lipofectamine 2000 (Invitrogen). The amount of plasmid DNA used for transfection was adjusted to be the same among ABCC11 WT and its variants. To inhibit proteasomal degradation when needed, cells at 24 h after transfection were cultivated in the presence of 2 μM MG132 for a further 24 h.

At 48 h after the transfection, cells were washed with PBS and then treated with lysis buffer A containing 50 mM Tris/HCl (pH 7.4), 1% (w/v) Triton X-100, 1 mM dithiothreitol, and protease inhibitor mixture (CompleteTM Mini; Nacalai Tesque, Inc., Kyoto, Japan). The cell suspension sample was homogenized by passage through a 27-gauge needle and then centrifuged at 800 *g* at 4°C for 10 min. The resulting supernatant fraction (whole-cell lysate) was transferred to a new tube, and the protein concentration was quantified by using the BCA Protein Assay Kit (Pierce, Rockford, IL, USA).

Detection of ABCC11 mRNA by reverse transcription-PCR

Total RNA was extracted from the ABCC11-transfected cells with the High-Pure RNA Isolation Kit (Roche Ltd., Mannheim, Germany) according to the manufacturer's protocol. cDNA was prepared from the extracted RNA in a reverse transcriptase reaction by using the High Capacity cDNA Archive kit (Applied Biosystems Inc.) and random hexamers, according to the manufacturer's instructions. The mRNA levels of ABCC11 and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were determined by PCR in an iCycler™ thermal cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA) with the following specific primers: ABCC11, 5′-AGCTTACGGAGTCGCTTAGATGA-3′ and 5′-TGGCTGTGCGTTGGTTGA-3′; and GAPDH, 5′-AATTCCATGGCACCGTCAA-3′ and 5′-CATGAGTCCTTCCACGATACCA-3′. The PCR reaction consisted of a hot-start incubation at 95°C for 2 min, followed by 95°C for 30 s, 56°C for 30 s,and 72°C for 30 s for 25 cycles. The PCR products were then analyzed by 2.0%

(w/v) agarose gel electrophoresis and detected under UV light by using ethidium bromide.

Analysis of *N*-linked glycosylation of ABCC11

The whole-cell lysates (150 μg of protein) were incubated with 190 U of either peptide *N*-glycosidase F (PNGase F; New England Biolabs, Ipswich, MA, USA) or endoglycosidase H (Endo H; New England Biolabs) at 37°C for 10 min before immunoblot analysis.

Immunoblotting detection of ABCC11 expressed in Flp-In-293 cells

The ABCC11 protein expressed in Flp-In-293 cells was detected by immunoblotting with the anti-ABCC11 polyclonal antibody. Briefly, samples were separated by SDS-PAGE (7.5% polyacrylamide slab gel) and transferred to Hy-bond ECL nitrocellulose membrane (GE Healthcare) by electroblotting at 15 V for 70 min. One hour after blocking in Tris-buffered saline containing 0.05% (v/v) Tween 20 (TTBS) and 5% (w/v) skim milk, immunoblotting was performed by using anti-human ABCC11 antibody (1:1000 dilution) as the first antibody and a goat anti-rabbit IgG-horseradish peroxidase (HRP)-conjugate (1:2000; Zymed, San Francisco, CA, USA) as the second antibody. HRP-dependent luminescence was developed by using Western Lightning Chemiluminescence Reagent Plus (PerkinElmer Life Sciences, Boston, MA, USA) and detected with a Lumino Imaging Analyzer FAS-1000 (Toyobo, Osaka, Japan). To detect GAPDH, used as an internal loading control, immunoblot detection was carried out in the same manner as described above, except for the use of mouse monoclonal anti-GAPDH antibody (1:2000; American Research Products, Belmont, MA, USA) as the primary antibody and HRP-conjugated horse anti-mouse IgG (1:3000; Cell Signaling Technology, Inc., Beverly, MA, USA) as the second antibody.

Immunofluoresence microscopy

Flp-In-293 cells were replated onto 35-mm dishes at a concentration of 1.0×10^6 cells/dish. After 24 h, the cells were transfected with the ABCC11-pcDNA3.1/Hyg(−) vector by using Lipofectamine 2000 (Invitrogen), as described above. At 36 h after transfection, ABCC11-expressing Flp-In-293 cells were seeded onto collagen type I-coated coverslips and incubated under the above-mentioned culture conditions for a further 36 h. Cells were fixed with 4% paraformaldehyde in PBS at room temperature for 20 min. Thereafter, cell membranes were permeabilized by incubation with 0.02% Triton X-100 in PBS at room temperature for 5 min. To block the free aldehyde groups of formaldehyde, cells were treated with 10 mg/ml glycine in PBS at room temperature for 10 min, followed by a further incubation with 0.5% (w/v) bovine serum albumin (BSA) in PBS at room temperature for 1 h. To detect the ABCC11 protein, cells were treated with the anti-ABCC11 antibody (1:500 dilution) as the primary antibody and subsequently with the Alexa Fluor 488-conjugated goat anti-rabbit IgG antibody (1:1000; Invitrogen). In the same preparations, nuclear DNA was stained with 4 μg/ml propidium iodide in PBS containing 0.5% (w/v) BSA. The immunofluorescence of Flp-In-293 cells was detected with a laser-scanning confocal fluorescence microscope (IX70/Fluoview; Olympus, Tokyo, Japan).

Detection of SNP 538G>A in ABCC11 gene by SmartAmp method

Templates used for the SmartAmp (K.K. DNAFORM, Yokohama, Japan) detection of the SNP 538G>A in the *ABCC11* gene were prepared from genomic DNA samples that were incubated at 98°C for 3 min. After chilling on ice, the sample preparation (20 ng of genomic DNA) was added directly into the reaction mixture (total volume of 25 μl) containing 2.0 μM folding primer (FP), 2.0 μM turn-back primer (TP), 1.0 μM boost primer (BP), 0.25 μM of each outer primer (OP1 and OP2), 20 μM competitive probe (CP), 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, 1:100,000 SYBR® Green I (Takara Bio Inc., Shiga, Japan), and 0.24 U/μl *Aac* DNA polymerase (K.K. DNAFORM). SmartAmp reaction mixtures were incubated at 60°C for 60 min under an isothermal condition in a real-time PCR system (Mx3000P; Stratagene), where changes in the fluorescence intensity of SYBR Green I dye indicating DNA amplification were monitored during the reaction.

Statistical tests

Experimental data were analyzed by determining the statistical significance according to Student's *t* test.

RESULTS

Localization of ABCC11 in secretory cells of human ceruminous apocrine glands

Fig. 1*A* displays photomicrographs of wet- and dry-type ceruminous apocrine glands stained with hematoxylin-eosin. In this study, we obtained the samples from Japanese subjects carrying the heterozygous (538G/A) and SNP homozygous (538A/A) alleles. Wet-type ceruminous apocrine glands in the subjects carrying heterozygous (538G/A) alleles are well developed, and they exhibited large luminal cavities as compared with dry-type ceruminous apocrine glands (homozygous 538A/A). Such morphological differences were consistent with previously reported observations for wet and dry types of human ceruminous glands (23).

To elucidate the expression and cellular localization of the ABCC11 protein in the ceruminous apocrine gland, we developed an ABCC11-specific polyclonal antibody raised against the epitope encoding amino acid residues 746–804 in the ABCC11 protein. The specificity and immunological activity of this antibody were confirmed as described below. Immunofluorescence staining of human tissue specimens containing cerumen glands revealed the expression of ABCC11 in the luminal domain of the plasma membrane of secretory cells in both wet- and dry-type ceruminal glands (Fig. 1*B*). It is noteworthy, however, that ABCC11 was predominantly localized in intracellular granules and large vacuoles in the secretory cells of wet-type ceruminous glands (Fig. 1*B*). In contrast, such granular and vacuolar localization of ABCC11 was not detected in the dry-type ceruminous glands (Fig. 1*B*). It is likely that the nonsynonymous SNP 538G>A (Gly180Arg) greatly affects the cellular localization of the ABCC11 protein in secretory cells.

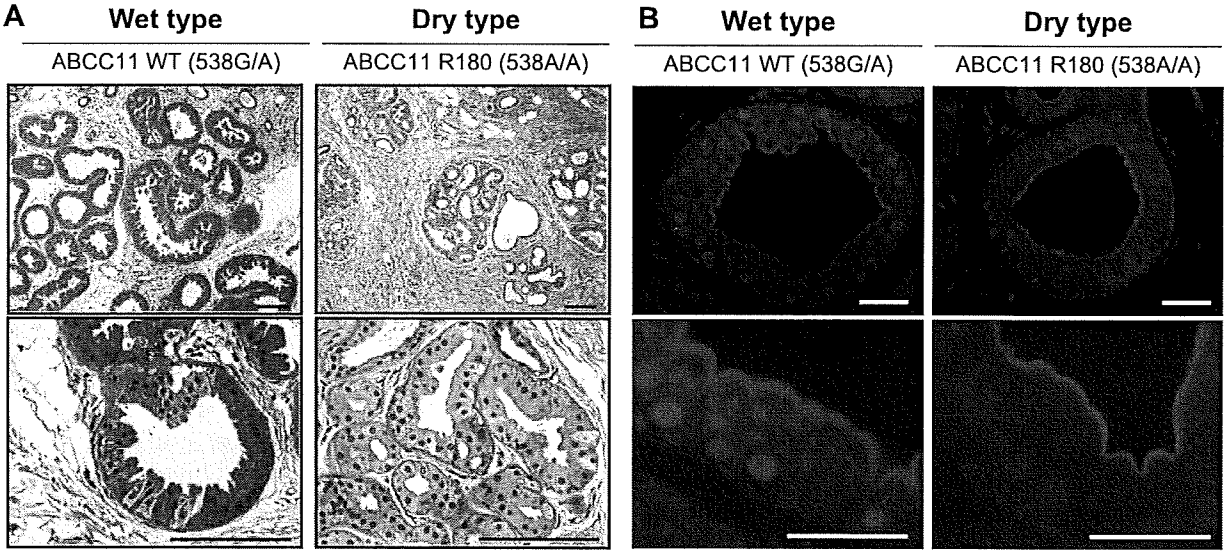


Figure 1. Microscopic pictures of wet- and dry-type ceruminous apocrine glands after hematoxylin-eosin staining and immunofluorescence staining of human ABCC11 protein. *A*) External auditory canal tissues were fixed with 4% paraformaldehyde in PBS (pH 7.4) overnight at 4°C, embedded in paraffin, and sectioned at 4 μm in thickness. Tissue sections were then stained with hematoxylin and eosin. *B*) Immunofluorescence staining of human ABCC11 protein in tissue specimens containing ceruminous apocrine glands, from human subjects carrying heterozygous (538G/A) and SNP homozygous (538A/A) alleles. ABCC11 protein expressed in secretory cells of wet- and dry-type ceruminous apocrine glands, 538G/A and 538A/A, was immunologically detected with ABCC11-specific polyclonal antibody and Alexa Fluor 488 (green). Cellular nuclei were stained with Hoechst 33342 (blue). Scale bars = 100 μm (*A*); 40 μm (*B*).

Validation of the ABCC11-specific antibody

Fig. 2 provides evidence that the polyclonal antibody is specific to human ABCC11. We expressed ABCC11 WT in Sf9 insect cells and Flp-In-293 human embryonic kidney cells by using pFastBac1 and pcDNA3.1/Hygro(−) vectors, respectively. Immunoblotting performed with the polyclonal antibody revealed one immunologically positive band (Fig. 2*C*, left column, arrowhead) in the ABCC11-expressing Sf9 cells. The molecular weight of the protein band was estimated to be ~150,000 (Fig. 2*C*). On the contrary, in the case of Flp-In-293 cells, two immunologically positive bands were observed at molecular weights of 180,000 and 150,000 (Fig. 2*C*, right column). After the sample was treated with PNGase F, the larger band (MW=180,000) disappeared, and one single im-

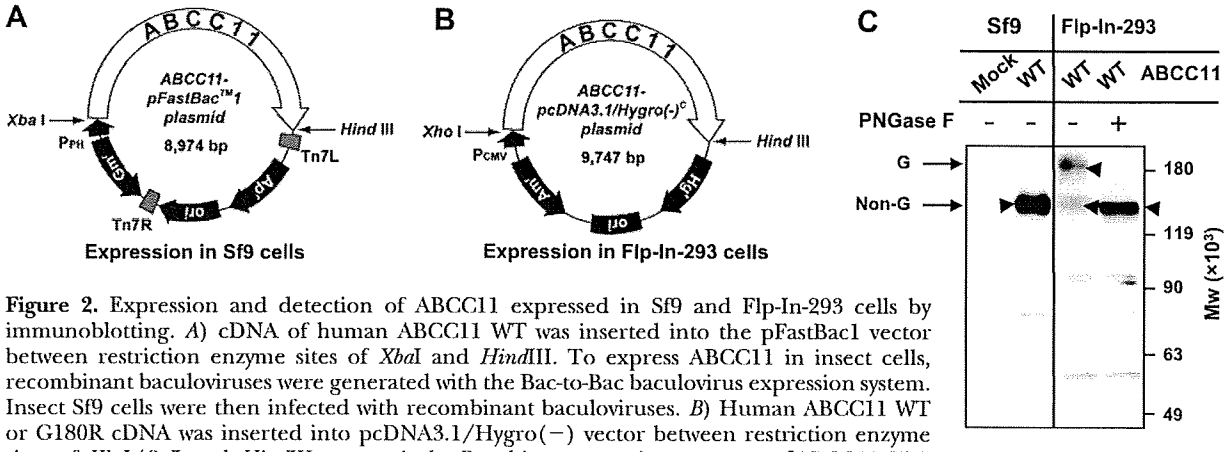


Figure 2. Expression and detection of ABCC11 expressed in Sf9 and Flp-In-293 cells by immunoblotting. *A*) cDNA of human ABCC11 WT was inserted into the pFastBac1 vector between restriction enzyme sites of *XbaI* and *Hind*III. To express ABCC11 in insect cells, recombinant baculoviruses were generated with the Bac-to-Bac baculovirus expression system. Insect Sf9 cells were then infected with recombinant baculoviruses. *B*) Human ABCC11 WT or G180R cDNA was inserted into pcDNA3.1/Hygro(−) vector between restriction enzyme sites of *XhoI*/*SalI* and *Hind*III, respectively. Resulting expression construct [ABCC11 WT-pcDNA3.1/Hygro(−)] was used as template for site-directed mutagenesis with corresponding primers (Supplemental Tables 1 to 3). Human Flp-In-293 cells were transfected with expression vector by using Lipofectamine 2000 to express ABCC11 in cells. *C*) ABCC11 protein expressed in Sf9 and Flp-In-293 cells was detected by immunoblotting with anti-ABCC11 polyclonal antibody. ABCC11 protein expressed in Sf9 cells was not glycosylated, whereas it was *N*-link glycosylated in Flp-In-293 cells. The *N*-linked glycosylated form of ABCC11 (MW=180,000) was changed to the nonglycosylated form (MW=150,000) by PNGase F treatment.

munologically positive band remained to be detected at the molecular weight of 150,000 (Fig. 2*C*). This molecular weight is consistent with that estimated from the ABCC11 cDNA. Thus, these observations strongly suggest that the ABCC11 WT protein is *N*-glycosylated in mammalian cells.

Expression of ABCC11 WT and SNP variant in Flp-In-293 cells

To understand the molecular mechanism underlying the difference in cellular localization of the ABCC11 WT and the SNP variant (Fig. 1*B*), we expressed both types (WT and R180) in Flp-In-293 cells *in vitro*. **Figure 3** shows the differential interference and immunofluorescence images of Flp-In-293 cells expressing ABCC11 WT and R180 variant proteins, as well as Flp-In-293 cells transfected with a mock vector. These ABCC11 proteins were probed with the polyclonal antibody and then labeled with green fluorescence dye (Alexa Fluor 488), whereas DNA in the nuclei was stained with propydium iodide (red fluorescence). Strong green fluorescence was observed at the plasma membrane and within intracellular compartments in Flp-In-293 cells expressing ABCC11 WT. In the case of the SNP variant, however, expression of the ABCC11 R180 variant at the plasma membrane was negligibly low. In contrast, the variant protein was detected within intracellular compartments proximal to the cellular nuclei in Flp-In-293 cells.

Protein expression levels of ABCC11 WT and SNP variant and their *N*-linked glycosylation status

In accordance with the results shown in Fig. 3, immunoblotting experiments revealed a significant difference in the protein expression level between the WT (G180) and the SNP variant (R180) of ABCC11 (**Fig. 4***A*). Despite having similar mRNA levels, the expression level of the R180 variant protein was much lower than that of the WT (Fig. 4*A*). Interestingly, the R180 protein level was remarkably (14-fold) enhanced by treatment of the cells with the proteasome inhibitor MG132 (Fig. 4*B*). In contrast, the protein level of the WT was only moderately affected by MG132 treatment (Fig. 4*B*), suggesting that the R180 protein is more susceptible to proteasomal degradation than the WT. Furthermore, the R180 protein was not the *N*-linked glycosylated form, whereas the WT was expressed as both *N*-linked glycosylated and nonglycosylated forms (Fig. 4*C*). The *N*-linked glycan was digested by PNGase F, but not by Endo H (Fig. 4*D*).

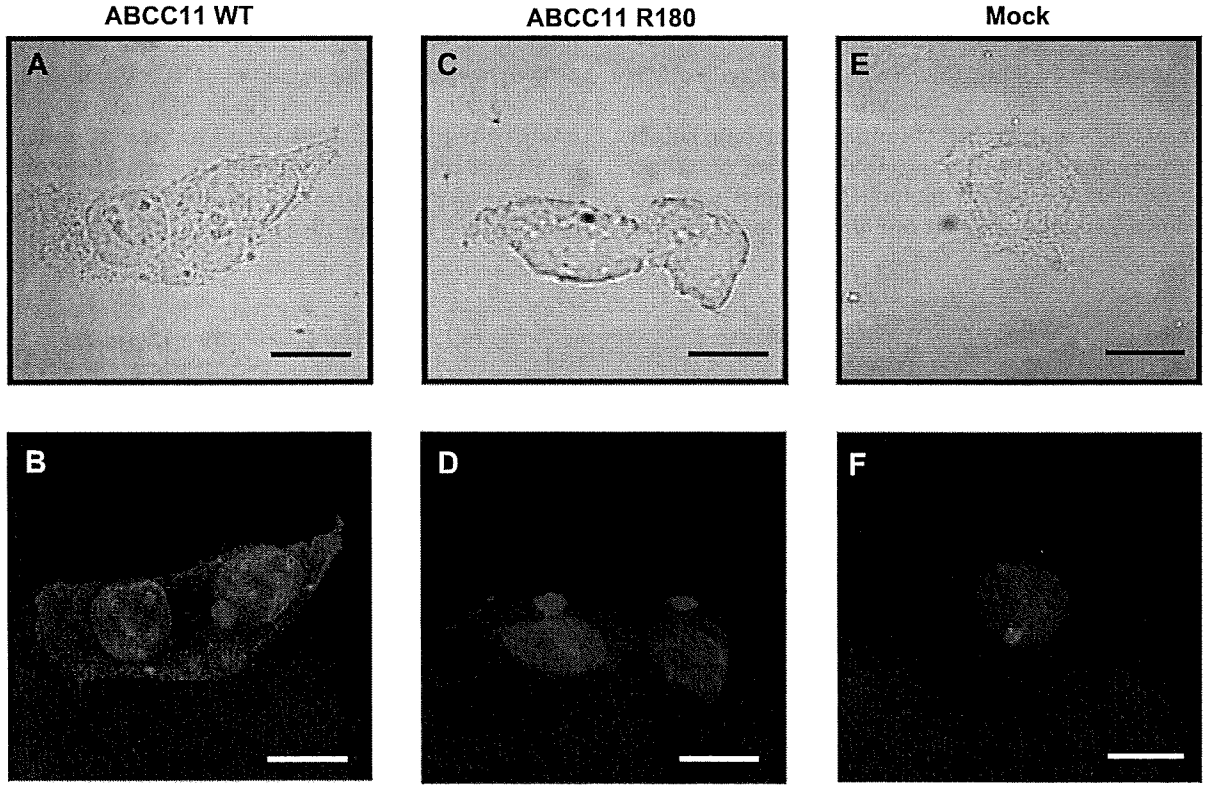


Figure 3. Immunocytochemical staining of Flp-In-293 cells expressing ABCC11 WT and R180 variant proteins. Differential interference (*A, C, E*) and immunofluorescence images (*B, D, F*). ABCC11 proteins were immunologically detected with ABCC11-specific polyclonal antibody and Alexa Fluor 488 (green). Cellular nuclei were stained by propydium iodide (red). Antibody interacts with epitope peptide residing in cytoplasmic domain (aa 746–804) between TM6 and TM7. Since Gly180 or Arg180 residue resides in TM1, this amino acid alteration (Gly180Arg) does not affect immunoreactivity. Scale bars = 10 μm.

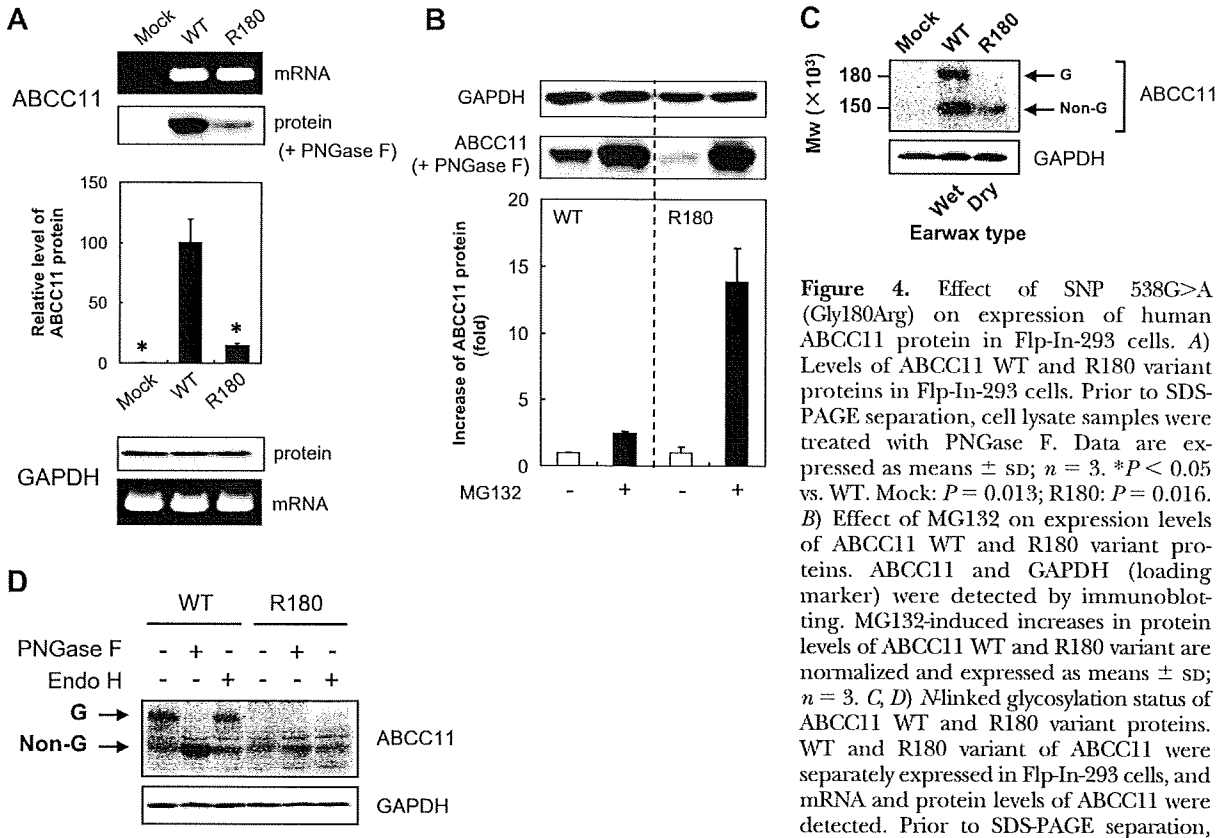


Figure 4. Effect of SNP 538G>A (Gly180Arg) on expression of human ABCC11 protein in Flp-In-293 cells. *A*) Levels of ABCC11 WT and R180 variant proteins in Flp-In-293 cells. Prior to SDS-PAGE separation, cell lysate samples were treated with PNGase F. Data are expressed as means ± SD; $n = 3$. $^*P < 0.05$ vs. WT. Mock: $P = 0.013$; R180: $P = 0.016$. *B*) Effect of MG132 on expression levels of ABCC11 WT and R180 variant proteins. ABCC11 and GAPDH (loading marker) were detected by immunoblotting. MG132-induced increases in protein levels of ABCC11 WT and R180 variant are normalized and expressed as means ± SD; $n = 3$. *C, D*) *N*-linked glycosylation status of ABCC11 WT and R180 variant proteins. WT and R180 variant of ABCC11 were separately expressed in Flp-In-293 cells, and mRNA and protein levels of ABCC11 were detected. Prior to SDS-PAGE separation, cell lysate samples were treated with PNGase F or Endo H. *N*-linked glycosylated (G) and nonglycosylated (non-G) ABCC11 proteins were detected by immunoblotting. GAPDH, a control for equal loading, was detected in a similar manner.

Identification of *N*-linked glycosylation sites in ABCC11 WT protein

Based on the amino acid sequence of ABCC11, a total of 8 potential *N*-linked glycosylation sites (Asn-*X*-Thr/Ser) was predicted (**Fig. 5*A***). The ABCC11 protein has 12 transmembrane helices, and its hydropathy profile indicates that only Asn838 and Asn844 reside in an extracellular loop between transmembrane helices TM7 and TM8 (Fig. 5*A*). Thus, we changed these Asn residues to Gln by site-directed mutagenesis and then expressed the protein in Flp-In-293 cells. Substitution of both Asn838 and Asn844 to Gln residues completely diminished *N*-linked glycosylation of the ABCC11 WT (Fig. 5*B*), demonstrating that these two Asn residues are *N*-linked glycosylation sites in the ABCC11 WT protein.

We hypothesized that amino acid substitution at 180 in the first transmembrane helix (TM1) might affect *N*-linked glycosylation of ABCC11. To examine our hypothesis, we substituted Gly180 to Arg, Lys, His, Asp, Glu, AL, and Pro in the ABCC11 WT protein. The substitution of Gly180 to a positively or negatively charged amino acid (*i.e.*, Arg, Lys, His, Asp, or Glu) diminished *N*-linked glycosylation of ABCC11, whereas substitution to a neutral amino acid (*i.e.*, AL or Pro) had no great effect (Fig. 5*C*). It is suggested that the electrostatic charge (either positive or negative) at aa 180 in the TM1 interferes with correct folding of the *de novo* synthesized ABCC11 protein in the ER.

Effect of nonsynonymous SNPs and a Δ27 mutation on the *N*-linked glycosylation of ABCC11

Furthermore, we created previously known nonsynonymous SNP variants and the Δ27 mutant of ABCC11 (**Fig. 6*A***) and expressed them in Flp-In-293 cells to examine their *N*-linked glycosylation status. The rare deletion mutation Δ27 results from the removal of a 9-aa stretch (Asp1313–Arg1321) from the C-terminal intracellular region of the ABCC11 protein (6). Both the rare mutation and G180R provide the dry type of earwax (6). Among the variants tested, only G180R and Δ27 diminished *N*-linked glycosylation of ABCC11 (Fig. 6*B*). Like the SNP variant R180 (Fig. 4*A, B*), the expression level of the Δ27 mutant protein was low (Fig. 6*C*), whereas it was significantly enhanced by MG132 treatments (Fig. 6*D*). Thus, lack of *N*-linked glycans of ABCC11 appears to be related with the dry type of human earwax.

Clinical genotyping of the SNP 538G>A (Gly180Arg) in the ABCC11 gene by the SmartAmp method

We tried to create a clinical method to genotype the SNP 538G>A in the heman *ABCC11* gene. For rapid genotyping, we have recently developed the SmartAmp method

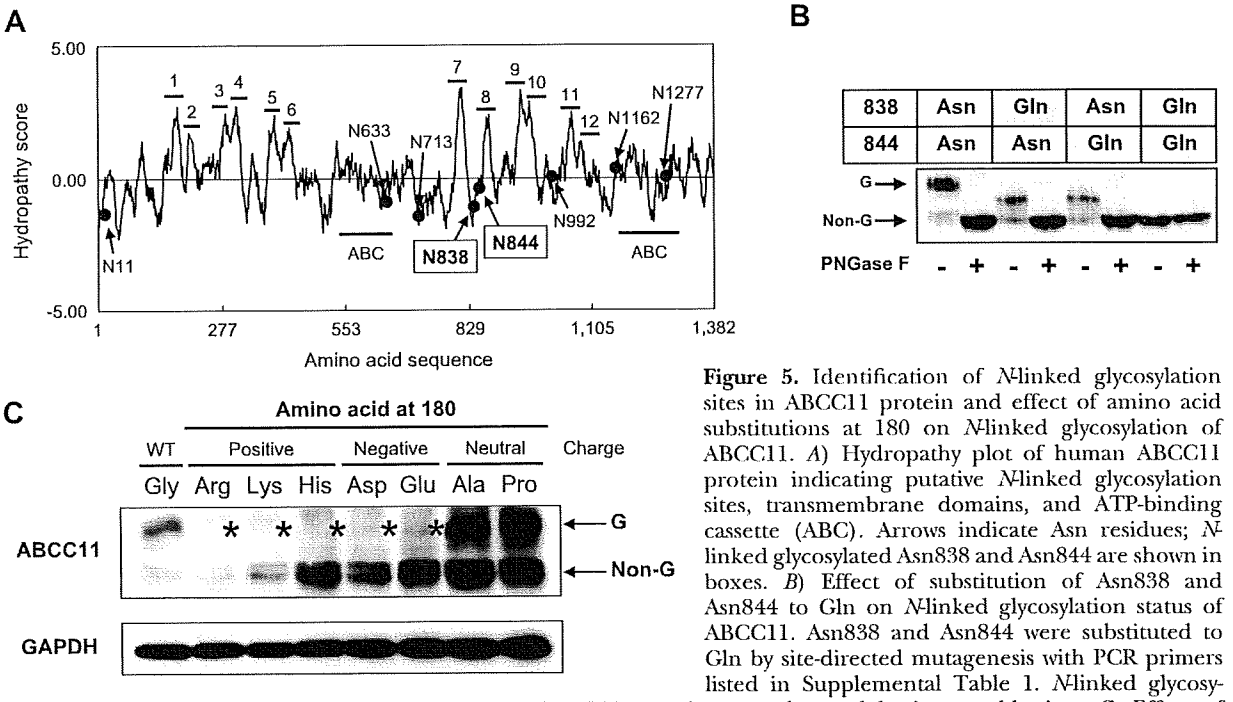


Figure 5. Identification of *N*-linked glycosylation sites in ABCC11 protein and effect of amino acid substitutions at 180 on *N*-linked glycosylation of ABCC11. *A*) Hydropathy plot of human ABCC11 protein indicating putative *N*-linked glycosylation sites, transmembrane domains, and ATP-binding cassette (ABC). Arrows indicate Asn residues; *N*-linked glycosylated Asn838 and Asn844 are shown in boxes. *B*) Effect of substitution of Asn838 and Asn844 to Gln on *N*-linked glycosylation status of ABCC11. Asn838 and Asn844 were substituted to Gln by site-directed mutagenesis with PCR primers listed in Supplemental Table 1. *N*-linked glycosylated (G) and nonglycosylated (non-G) forms of ABCC11 protein were detected by immunoblotting. *C*) Effect of electrostatic charges at amino acid residue 180 on *N*-linked glycosylation status of ABCC11. Gly180 residue in ABCC11 WT protein was substituted to Arg, Lys, His, Asp, Glu, AL, and Pro by site-directed mutagenesis with PCR primers listed in Supplemental Table 2. ABCC11 protein was detected by immunoblotting; glycosylation forms are indicated. Asterisk (*) indicates no *N*-linked glycosylation of ABCC11.

that enables us to detect genetic polymorphisms or mutations in ~30 min under isothermal conditions (24). **Figure 7***A* schematically illustrates the strategy of SNP detection by the SmartAmp method. To determine the SNP 538G>A (Gly180Arg) in the *ABCC11* gene, we prepared one set of primers designated TP, FP, BP, OP, and CP (Fig. 7*B*). The TPs discriminate the polymorphism 538G or 538A in the *ABCC11* gene, and the CPs

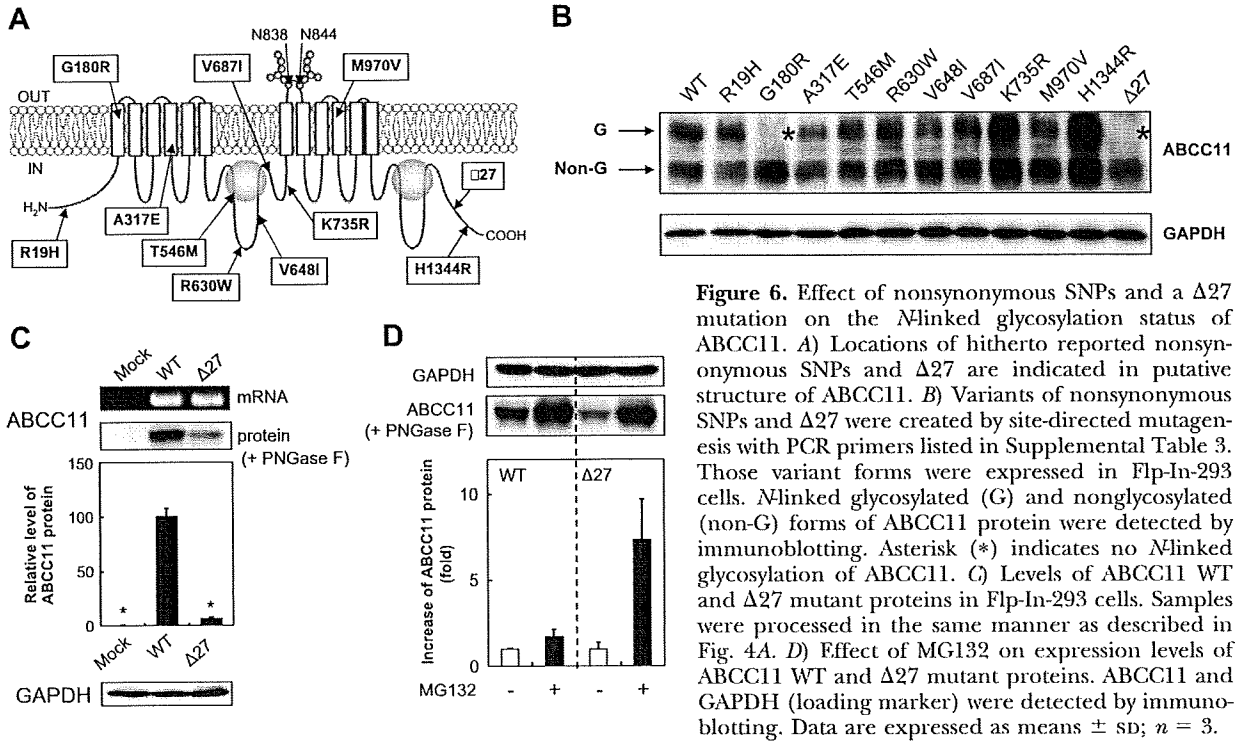


Figure 6. Effect of nonsynonymous SNPs and a Δ27 mutation on the *N*-linked glycosylation status of ABCC11. *A*) Locations of hitherto reported nonsynonymous SNPs and Δ27 are indicated in putative structure of ABCC11. *B*) Variants of nonsynonymous SNPs and Δ27 were created by site-directed mutagenesis with PCR primers listed in Supplemental Table 3. Those variant forms were expressed in Flp-In-293 cells. *N*-linked glycosylated (G) and nonglycosylated (non-G) forms of ABCC11 protein were detected by immunoblotting. Asterisk (*) indicates no *N*-linked glycosylation of ABCC11. *C*) Levels of ABCC11 WT and Δ27 mutant proteins in Flp-In-293 cells. Samples were processed in the same manner as described in Fig. 4*A*. *D*) Effect of MG132 on expression levels of ABCC11 WT and Δ27 mutant proteins. ABCC11 and GAPDH (loading marker) were detected by immunoblotting. Data are expressed as means ± SD; $n = 3$.

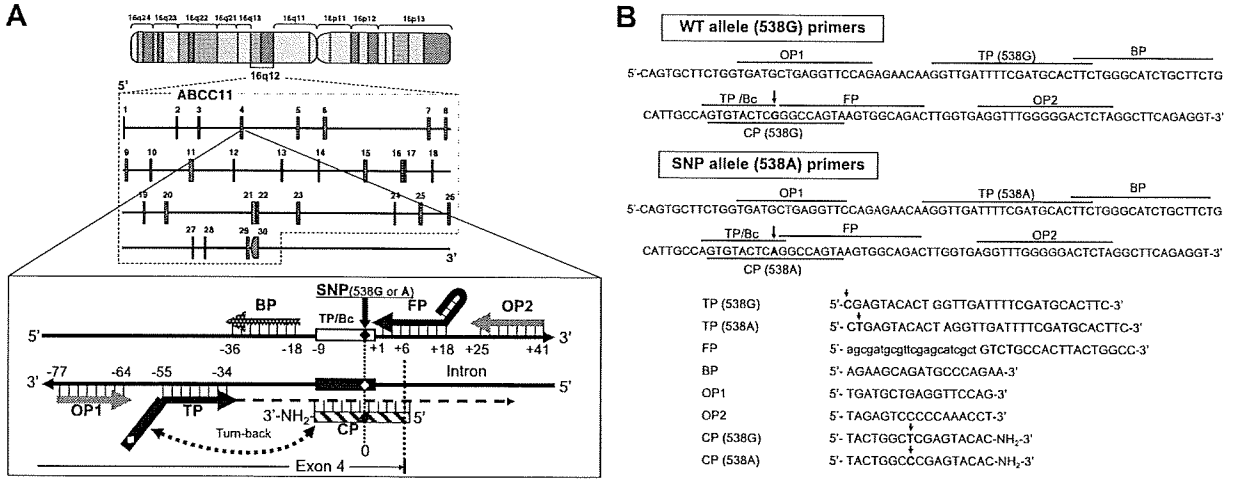


Figure 7. Strategy for SmartAmp-based detection of SNP 538G>A in *ABCC11* gene. *A*) Schematic illustration of SmartAmp-based SNP detection. SNP 538G>A resides in exon 4 of ABCC11 gene on chromosome 16q12. *B*) Partial genomic DNA sequences of *ABCC11* gene carrying WT (538G) and SNP (538A) alleles and sequences of primers used for SmartAmp assay.

inhibit the background amplification from mismatch sequence pairs (25, 26) (Fig. 7*A*). These primers selectively recognized the SNP 538G>A of the *ABCC11* gene to discriminate homozygous 538G/G (wet type), heterozygous 538G/A (wet type), and homozygous 538A/A (dry type) in genomic DNA (**Fig. 8***A*). By this new method, we typed the SNP in genomic DNA samples from a total of 124 healthy volunteers and compared the results obtained by DNA sequencing. Neither false positives nor false negatives were observed (Fig. 8*B*). Thus, the perfect matching validated the SmartAmp-based SNP typing method for use in clinical practice.

Since previous phenotype-based studies have suggested a positive association between the wet-type earwax and axillary osmidrosis (1), we analyzed the SNP in the *ABCC11* gene of Japanese axillary osmidrosis patients. In this clinical study, axillary osmidrosis was diagnosed basically from both the self-declaration of patients and the odor-smelling test carried out by authorized medical doctors. The earwax type was determined from wax samples obtained by stirring a cotton swab in the external auditory canal of the patient. **Table 1** summarizes the clinical results from this study with respect to the *ABCC11* genotype as well as axillary osmidrosis and earwax types. It is notable that all axillary osmidrosis patients carrying the 538 alleles as either WT homozygous (538G/G) or heterozygous (538G/A) have the wet type of earwax, without exception. Although the population size was rather small in this study, the *ABCC11* WT allele is suggested to be a genetic biomarker for clinical diagnosis of axillary osmidrosis.

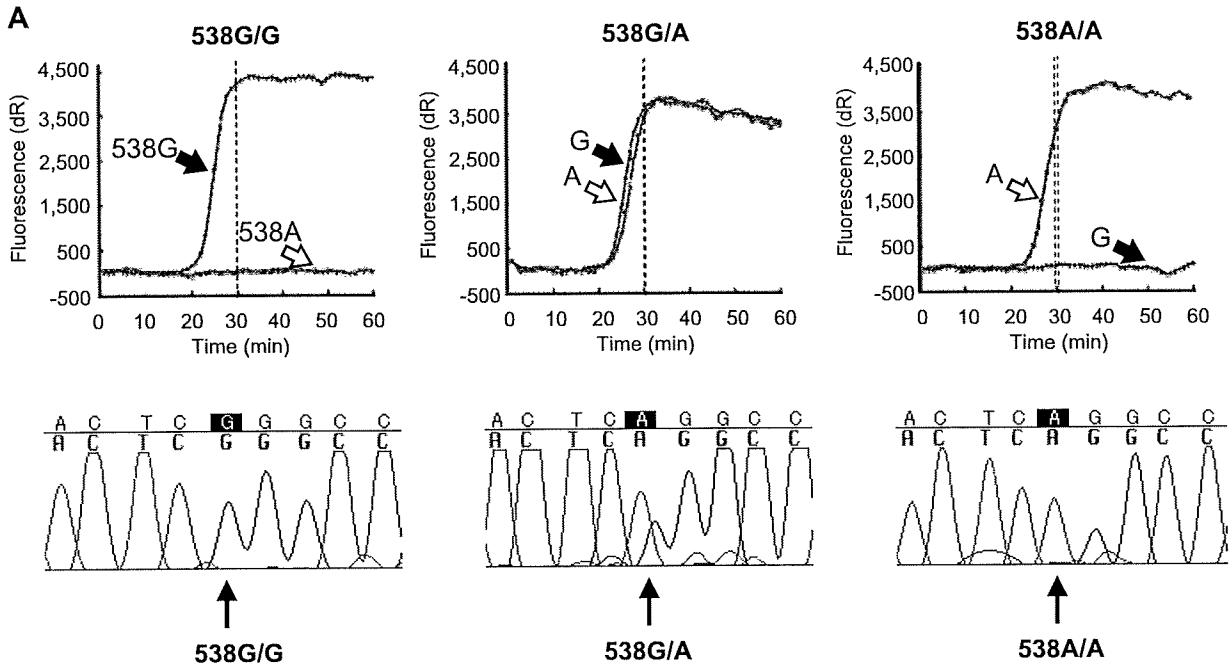
DISCUSSION

Proteasomal degradation of ABCC11 R180 is the molecular mechanism determining the dry-type earwax

The present study provides direct evidence that human ABCC11 WT is an *N*-linked glycosylated protein, which is localized in intracellular granules and large vacuoles as well as at the luminal membrane of secretory cells in the cerumen apocrine gland. *N*-linked glycosylation occurs at both Asn838 and Asn844 in the extracellular loop between TM7 and 8 TM8 of the ABCC11 WT protein. In contrast, the SNP variant R180 lacks *N*-linked glycosylation and readily undergoes proteasomal degradation, most probably *via* ubiquitination. As a consequence, granular or vacuolar localization was not detected in the cerumen apocrine gland for the SNP variant.

Morphological differences were previously reported between the secretory cells of wet and dry types of human ceruminous glands (23). In the wet-type glands, the Golgi apparatus was reportedly well developed, whereas it was generally small in the corresponding cells of the dry type. Furthermore, intracellular granules were abundantly observed in the wet-type gland in close relationship to their well-developed Golgi apparatus, whereas intracellular granules were rare in the dry-type gland.

The ER and Golgi apparatus are the sites of synthesis and maturation of proteins destined for the plasma membrane, for the secretory and endocytic organelles, and for secretion (27, 28). Efficient quality-control systems have evolved to prevent incompletely folded proteins from moving along the secretory pathway. Accumulation of misfolded proteins in the ER would detrimentally affect cellular functions. Therefore, misfolded proteins may be removed from the ER by retrotranslocation to the cytosol compartment, where they are degraded by the ubiquitin-proteasome system. This process is known as endoplasmic reticulum-associated degradation (ERAD) (29-32). It is likely that both the SNP variant R180 and the Δ27 mutant are recognized as misfolded proteins in the ER and readily undergo the proteasomal degradation. Indeed, the protein levels of both R180 (Fig. 4*B*) and Δ27 (Fig. 6*D*) were greatly enhanced by treatment of cells with the



B

	SMAP			
Sequence	Geno - type	G/G	G/A	A/A
	G/G	4	0	0
	G/A	0	33	0
	A/A	0	0	87

N =124

Wet type Dry type

Figure 8. Diagnosis of SNP 538G>A by SmartAmp assay and DNA sequencing. *A*) Time course of SmartAmp assay reaction with *ABCC11* allele-specific primers. Results of SmartAmp assay (top panels) and DNA sequence analysis (bottom panels) are shown for 3 diploid genotypes of ABCC11. *B*) Comparison of SNP analysis data obtained by SmartAmp method and DNA sequencing.

proteasome inhibitor MG132. This protein processing may influence the activity of ceruminous apocrine glands and determine the type of human earwax. In **Fig. 9**, we schematically illustrate the effects of the SNP on the cellular localization and function of ABCC11 in secretory cells of the apocrine gland.

TABLE 1. *Genotyping of the ABCC11 gene in axillary osmidrosis patients and comparison with earwax type*

	ABCC11 genotype (538G>A)		
Diagnosis	G/G	G/A	A/A
Axillary osmidosis			
Positive	1	11	0
Negative	0	0	2
Earwax type			
Wet	1	11	0
Dry	0	0	2

Diagnosis of axillary osmidrosis and earwax type as well as genotyping of *ABCC11* were performed at Fujita Health University School of Medicine according to the protocol approved by the Ethical Review Board. The SNP 538G>A in the *ABCC11* gene was analyzed by both DNA sequencing and the SmartAmp method.

We have recently demonstrated that the intramolecular disulfide bond formation and *N*-linked glycosylation in the extracellular loop are important for stability of human ABC transporter ABCG2 in the ER (33-36). During *de novo* synthesis in the ER, cysteine disulfide bonds are formed, and oligosaccharides are added to asparagine (*N*-linked glycosylation) or serine residues (*O*-glycosylation) of glycoproteins. In general, *N*-linked glycans are added *en block* to proteins as "core oligosaccharides" ($Glc_3Man_9GlcNAc_2$) (27, 29). As described above, Asn838 and Asn844 are glycosylation target sites in human ABCC11. The *N*-linked glycans are thought to be subjected to extensive modification as glycoproteins mature and move through the ER *via* the Golgi apparatus to their final destination, for example, intracellular granules and large vacuoles of secretory cells in the apocrine gland. It is puzzling, however, how amino acid alteration at 180 in TM1 affected the *N*-linked glycosylation occurring in the extracellular loop between TM7 and TM8. To answer this question, we assume that the electrostatic charge at aa 180 in the TM1 could interfere with correct folding of the *de novo* synthesized ABCC11 protein in the ER, as shown in Fig.

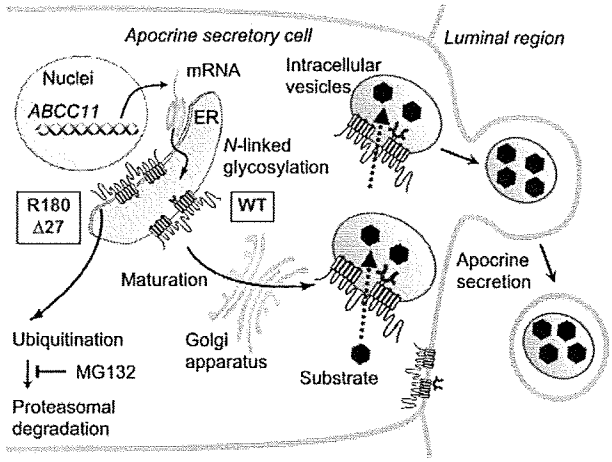


Figure 9. Schematic illustration of intracellular sorting of ABCC11 WT and proteasomal degradation of R180 and Δ27 variants in secretory cells in ceruminous apocrine gland. *De novo* synthesized ABCC11 WT is *N*-link glycosylated at Asn838 and Asn844 in the ER, further processed in the Golgi apparatus, and destined for membrane of intracellular granules and vacuoles. Ceruminous components are thought to be transported by ABCC11 WT and sequestrated in intracellular granules and vacuoles. SNP variant R180 and Δ27 mutant lacking *N*-linked glycosylation are recognized as misfolded proteins in the ER and readily undergo ubiquitination and proteasomal degradation.

5C. It is important to note, however, that *N*-linked glycosylation may affect in part the protein stability, but it is not the principal factor for correctly folding and/or enhancing the stability of the ABCC11 protein in the ER. As exemplified by our recent study, substitutions of one single amino acid due to nonsynonymous polymorphisms greatly affected the protein stability of SNP variants of human ABCG2 (33), whereas deletion of *N*-linked glycosylation at Asn596 reduced the protein expression level of ABCG2 WT only by half, without any change in its plasma membrane localization (unpublished results). Thus, molecular mechanisms underlying the misfolding and degradation of membrane proteins appear to be rather complex, as they involve interactions with different chaperone proteins.

At present, it remains unclear how misfolded membrane proteins are selected and destroyed during ERAD. Chaperones are considered to solubilize aggregation-prone motifs. In the case of the yeast ABC transporter Ste6p, a 12-transmembrane protein, it has recently been shown that Hsp70/40s act before ubiquitination and facilitate Ste6p association with an E3 ubiquitin ligase (37). Furthermore, polyubiquitination was a prerequisite for retrotranslocation, which required the Cdc48 complex and ATP (37). In this regard, Wang *et al.* (38) and Younger *et al.* (39) recently provided new insights into the complexity of protein networks that govern the fate of the cystic fibrosis transmembrane conductance regulator (CFTR), an apical membrane ABC transporter. The loss of phenylalanine residue at position 508 in the first nucleotide-binding domain disrupts the folding pathway of CFTR protein in the ER, and subsequently misfolded CFTR is targeted for ERAD (38-42). In the case of human bile salt export pump (BSEP/ABCB11), at least two *N*-linked glycans are reportedly important for protein stability, intracellular trafficking, and function in the apical membrane (43). In this context, it is increasingly important to identify and to characterize multiple chaperone proteins that control the folding and degradation of ABC transporter proteins, including ABCC11.

Clinical relevance of genetic polymorphisms of human ABCC11

Genetic polymorphisms of human ABC transporter genes are reportedly related to the risk of disease and patient response to medication (44, 45). Phenotype-based studies have suggested positive associations among the wet-type earwax, axillary osmidrosis (1), colostrum secretion from the mammary gland (2), and the potential susceptibility of breast cancer (4). Our genotyping study has revealed that the *ABCC11* WT allele is intimately associated with axillary osmidrosis (**Table 1**). Since the frequency of the *ABCC11* WT allele is low in Japan, axillary osmidrosis is recognized as a disease that is covered by the national health insurance system. Axillary osmidrosis is often perceived, especially by young women, as a distressing and troublesome problem.

Sweat produced by axillary glands is odorless. Secretions from the apocrine glands, however, can be converted to odoriferous compounds by bacteria (*Corynebacteria*), which results in the formation of the unique "human axillary odor" (46). In axillary osmidrosis patients (G/G homozygote or G/A heterozygote), significantly numerous and larger-sized axillary apocrine glands were observed as compared with the subjects carrying the A/A homozygote. Our results (**Table 1**) suggest that the 538G allele of the *ABCC11* gene is associated with axillary osmidrosis and that ABCC11 WT is responsible for the secretion of preodoriferous compounds from the axillary apocrine gland.

In primates, the axillary odors may play a role in olfactory communication, although no documented behavioral or endocrine changes by volatiles produced in the axillae have been reported to occur in humans. Previous studies have shown that androgen steroids were present in the axillary area. Androsterone sulfate (AS) and DHEAS were detected in the extract of axillary hairs, in addition to high levels of cholesterol (47). It was also demonstrated, following injection of radioactive pregnenolone or progesterone, that steroid secretion was concentrated in the axillary area (48). The axillary sweat collected in these studies from the skin surface, however, represents a mixture of materials from apocrine, eccrine, and sebaceous glands, in addition to desquamating epidermal cells. In this respect, Lebows *et al.* (49) conclusively demonstrated that at least two androgen steroids, AS and DHEAS, in addi-

tion to cholesterol, did exist in pure apocrine secretions. It has more recently been reported that DHAES is transported by ABCC11 WT (13-15). Further studies are needed to define more precisely the substrates for ABCC11 in apocrine glands.

CONCLUSIONS

Based on the previous phenotype analysis, mortality and frequency rates for breast cancer were reported to be associated with the frequency of the allele for wet-type earwax (4), although the association with breast cancer was controversial (50). Our preliminary genotyping study, however, revealed that the allele of 538G (wild type) in the *ABCC11* gene is associated with risk of breast cancer among a total of 416 Japanese premenopausal women as compared with the 538A (SNP type) allele. The odds ratio was found to be 1.86 (unpublished results). The SNP typing method presented in this study would provide a practical tool to examine a latent genetic link between wet-type earwax and the potential risk of breast cancer. FJ

This study was supported by the Japan Science and Technology Agency (JST) research project, "Development of the world's fastest SNP detection system," to T.I.; by Grant-in-Aid for Scientific Research (A) 18201041 to T.I.; in part by a JST Solution Oriented Research for Science and Technology (SORST) grant to N.N.; by Japanese Society for the Promotion of Science (JSPS) Grant-in-Aid for Exploratory Research 19659136 to T.I.; by Grant-in-Aid for Young Scientists (B) 19791361 to H.N.; by Grant-in-Aid for Scientific Research on Priority Areas 18059027 to K.Y.; and by a Ministry of Education, Culture, Sports, Science and Technology (MEXT) research grant for RIKEN Omics Science Center to Y.H. Y.T. is a JSPS fellow.

REFERENCES

1. Yoo, W. M., Pae, N. S., Lee, S. J., Roh, T. S., Chung, S., and Tark, K. C. (2006) Endoscopy-assisted ultrasonic surgical aspiration of axillary osmidrosis: a retrospective review of 896 consecutive patients from 1998 to 2004. *J. Plast. Reconstr. Aesthet. Surg.* **59,** 978–982
2. Miura, K., Yoshiura, K., Miura, S., Shimada, T., Yamasaki, K., Yoshida, A., Nakayama, D., Shibata, Y., Niikawa, N., and Masuzaki, H. (2007) A strong association between human earwax-type and apocrine colostrum secretion from the mammary gland. *Hum. Genet.* **121,** 631–633
3. Petrakis, N. L. (1983) Cerumen phenotype and epithelial dysplasia in nipple aspirates of breast fluid. *Am. J. Phys. Anthropol.* **62,** 115–118
4. Petrakis, N. L. (1971) Cerumen genetics and human breast cancer. *Science* **173,** 347–349
5. Gesase, A. P., and Satoh, Y. (2003) Apocrine secretory mechanism: recent findings and unresolved problems. *Histol. Histopathol.* **18,** 597–608
6. Yoshiura, K., Kinoshita, A., Ishida, T., Ninokata, A., Ishikawa, T., Kaname, T., Bannai, M., Tokunaga, K., Sonoda, S., Komaki, R., Ihara, M., Saenko, V., Alipov, G., Sekine, I. O., Komatsu, K., Takahashi, H., Nakashima, M., Sosonkina, N., Mapendano, C., Ghadami, M., Nomura, M., Liang, D.-S., Miwa, N., Dae-Kwang, K., Garidkhuu, A., Natsume, N., Ohta, T., Tomita, H., Kaneko, A., Kikuchi, M., Russomando, G., Hirayama, K., Ishibashi, M., Takahashi, A., Saitou, N., Murray, J. C., Saito, S., Nakamura, Y., and Nikawa, N. (2006) A SNP in the ABCC11 gene is the determinant of human earwax type. *Nat. Genet.* **38,** 324–330
7. Yabuuchi, H., Shimizu, H., Takayanagi, S., and Ishikawa, T. (2001) Multiple splicing variants of two new human ATP-binding cassette transporters, ABCC11 and ABCC12. *Biochem. Biophys. Res. Commun.* **288,** 933–939
8. Tammur, J., Prades, C., Arnould, I., Rzhetsky, A., Hutchinson, A., Adachi, M., Schuetz, J. D., Swoboda, K. J., Ptacek, L. J., Rosier, M., Dean, M., and Allikmets, R. (2001) Two new genes from the human ATP-binding cassette transporter superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome 16q12. *Gene* **273,** 89–96
9. Bera, T. K., Lee, S., Salvatore, G., Lee, B., and Pastan, I. (2001) MRP8, a new member of ABC transporter superfamily, identified by EST database mining and gene prediction program, is highly expressed in breast cancer. *Mol. Med.* **7,** 509–516
10. Ono, N., Van der Heijden, I., Scheffer, G. L., Van de Wetering, K., Van Deemter, E., De Haas, M., Boerke, A., Gadella, B. M., De Rooij, D. G., Neefjes, J. J., Groothuis, T. A. M., Oomen, L., Brocks, L., Ishikawa, T., and Borst, P. (2007) Multidrug resistance-associated protein 9 (ABCC12) is present in mouse and boar sperm. *Biochem. J.* **406,** 31–40
11. Shimizu, H., Taniguchi, H., Hippo, Y., Hayashizaki, Y., Aburatani, H., and Ishikawa, T. (2003) Characterization of the mouse Abcc12 gene and its transcript encoding an ATP-binding cassette transporter, an orthologue of human ABCC12. *Gene* **310,** 17–28
12. Toyoda, Y., Hagiya, Y., Adachi, T., Hoshijima, K., Kuo, M. T., and Ishikawa, T. (2008) MRP class of human ATP binding cassette (ABC) transporters: historical background and new research directions. *Xenobiotica* **38,** 833–862
13. Kruh, G. D., Guo, Y., Hopper-Borge, E., Belinsky, M. G., and Chen, Z. S. (2007) ABCC10, ABCC11, and ABCC12. *Pflügers Arch.* **453,** 675–684
14. Bortfeld, M., Rius, M., Konig, J., Herold-Mende, C., Nies, A. T., and Keppler, D. (2006) Human multidrug resistance protein 8 (MRP8/ABCC11), an apical efflux pump for steroid sulfates, is an axonal protein of the CNS and peripheral nervous system. *Neuroscience* **137,** 1247–1257
15. Chen, Z. S., Guo, Y., Belinsky, M. G., Kotova, E., and Kruh, G. D. (2005) Transport of bile acids, sulfated steroids, estradiol 17-beta-D-glucuronide, and leukotriene C4 by human multidrug resistance protein 8 (ABCC11). *Mol. Pharmacol.* **67,** 545–557
16. Guo, Y., Kotova, E., Chen, Z. S., Lee, K., Hopper-Borge, E., Belinsky, M. G., and Kruh, G. D. (2003) MRP8, ATP-binding cassette C11 (ABCC11), is a cyclic nucleotide efflux pump and a resistance factor for fluoropyrimidines 2′,3′-dideoxycytidine and 9′-(2′-phosphonylmethoxyethyl)adenine. *J. Biol. Chem.* **278,** 29509–29514
17. Yamaguchi-Kabata, Y., Nakazono, K., Takahashi, A., Saito, S., Hosono, N., Kubo, M., Nakamura, Y., and Kamatani, N. (2008) Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83,** 445–456
18. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132
19. Nakashima, M., Meirmanov, S., Matsufuji, R., Hayashida, M., Fukuda, E., Naito, S., Matsuu, M., Shichijo, K., Kondo, H., Ito, M., Yamashita, S., and Sekine, I. (2002) Altered expression of beta-catenin during radiation-induced colonic carcinogenesis. *Pathol. Res. Pract.* **198,** 717–724
20. Tamura, A., Wakabayashi, K., Onishi, Y., Takeda, M., Ikegami, Y., Sawada, S., Tsuji, M., Matsuda, Y., and Ishikawa, T. (2007) Re-evaluation and functional classification of non-synonymous single nucleotide polymorphisms of the human ATP-binding cassette transporter ABCG2. *Cancer Sci.* **98,** 231–239
21. Wakabayashi, K., Nakagawa, H., Adachi, T., Kii, I., Kobatake, E., Kudo, A., and Ishikawa, T. (2006) Identification of cysteine residues critically involved in homodimer formation and protein expression of human ATP-binding cassette transporter ABCG2: a new approach using the flp recombinase system. *J. Exp. Ther. Oncol.* **5,** 205–222
22. Tamura, A., Wakabayashi, K., Onishi, Y., Nakagawa, H., Tsuji, M., Matsuda, Y., and Ishikawa, T. (2006) Genetic polymorphisms of human ABC transporter ABCG2: development of the standard method for functional validation of SNPs by using the Flp recombinase system. *J. Exp. Ther. Oncol.* **6,** 1–11

23. Shugyo, Y., Sudo, N., Kanai, K., Yamashita, T., Kumazawa, T., and Kanamura, S. (1988) Morphological differences between secretory cells of wet and dry types of human ceruminous glands. *Am. J. Anat.* **181,** 377–384
24. Mitani, Y., Lezhava, A., Kawai, Y., Kikuchi, T., Oguchi-Katayama, A., Kogo, Y., Itoh, M., Miyagi, T., Takakura, H., Hoshi, K., Kato, C., Arakawa, T., Shibata, K., Fukui, K., Masui, R., Kuramitsu, S., Kiyotani, K., Chalk, A., Tsunekawa, K., Murakami, M., Kamataki, T., Oka, T., Shimada, H., Cizdziel, P. E., and Hayashizaki, Y. (2007) Rapid SNP diagnostics using asymmetric isothermal amplification and a new mismatch-suppression technology. *Nat. Methods* **4,** 257–262
25. Kawai, Y., Kikuchi, T., Mitani, Y., Kogo, Y., Itoh, M., Usui, K., Kanamori, H., Kaiho, A., Takakura, H., Hoshi, K., Cizdziel, P. E., and Hayashizaki, Y. (2008) Sensitive detection of EGFR mutations using a competitive probe to suppress background in the SMart Amplification Process. *Biologicals* **36,** 234–238
26. Watanabe, J., Mitani, Y., Kawai, Y., Kikuchi, T., Kogo, Y., Oguchi-Katayama, A., Kanamori, H., Usui, K., Itoh, M., Cizdziel, P. E., Lezhava A., Tatsumi, K., Ichikawa, Y., Togo, S., Shimada, H., and Hayashizaki, Y. (2007) Use of a competitive probe in assay design for genotyping of the UGT1A1 *28 microsatellite polymorphism by the smart amplification process. *BioTechniques* **43,** 479–484
27. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73,** 1019–1049
28. Ellgaard, L., Molinari, M., and Helenius, A. (1999) Setting the standards: quality control in the secretory pathway. *Science* **286,** 1882–1888
29. Kleizen, B., and Braakman, I. (2004) Protein folding and quality control in the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **16,** 343–349
30. Hampton, R. Y. (2002) ER-associated degradation in protein quality control and cellular regulation. *Curr. Opin. Cell Biol.* **14,** 476–482
31. Ellgaard, L., and Helenius, A. (2001) ER quality control: towards an understanding at the molecular level. *Curr. Opin. Cell Biol.* **13,** 431–437
32. Mori, K. (2000) Tripartite management of unfolded proteins in the endoplasmic reticulum. *Cell* **101,** 451–454
33. Nakagawa, H., Tamura, A., Wakabayashi, K., Hoshijima, K., Komada, M., Yoshida, T., Kometani, S., Matsubara, T., Mikuriya, K., and Ishikawa, T. (2008) Ubiquitin-mediated proteasomal degradation of non-synonymous SNP variants of human ABC transporter ABCG2. *Biochem. J.* **411,** 623–631
34. Wakabayashi, K., Nakagawa, H., Tamura, A., Koshiba, S., Hoshijima, K., Komada, M., and Ishikawa, T. (2007) Intramolecular disulfide bond is a critical check point determining degradative fates of ATP-binding cassette (ABC) transporter ABCG2 protein. *J. Biol. Chem.* **282,** 27841–27846
35. Furukawa, T., Wakabayashi, K., Tamura, A., Nakagawa, H., Morishima, Y., Osawa, Y., and Ishikawa, T. (2009) Major SNP (Q141K) variant of human ABC transporter ABCG2 undergoes lysosomal and proteasomal degradations. *Pharm. Res.* **26,** 469–479
36. Wakabayashi-Nakao, K., Tamura, A., Furukawa, T., Nakagawa, H., and Ishikawa, T. (2009) Quality control of human ABCG2 protein in the endoplasmic reticulum: Ubiquitination and proteasomal degradation. *Adv. Drug Deliv. Rev.* **61,** 66–72
37. Nakatsukasa, K., Huyer, G., Michaelis, S., and Brodsky, J. L. (2008) Dissecting the ER-associated degradation of a misfolded polytopic membrane protein. *Cell* **132,** 101–112
38. Wang, X., Venable, J., LaPointe, P., Hutt, D. M., Koulov, A. V., Coppinger, J., Gurkan, C., Kellner, W., Matteson, J., Plutner, H., Riordan, J. R., Kelly, J. W., Yates, J. R., and Balch, W. E. (2006) Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis. *Cell* **127,** 803–815
39. Younger, J. M., Chen, L., Ren, H. Y., Rosser, M. F., Turnbull, E. L., Fan, C. Y., Patterson, C., and Cyr, D. M. (2006) Sequential quality-control checkpoints triage misfolded cystic fibrosis transmembrane conductance regulator. *Cell* **126,** 571–582
40. Skach, W. R. (2006) CFTR: new members join the fold. *Cell* **127,** 673–675
41. Riordan, J. R. (2005) Assembly of functional CFTR chloride channels. *Annu. Rev. Physiol.* **67,** 701–718
42. Qu, B. H., Strickland, E. H., and Thomas, P. J. (1997) Localization and suppression of a kinetic defect in cystic fibrosis transmembrane conductance regulator folding. *J. Biol. Chem.* **272,** 15739–15744
43. Mochizuki, K., Kagawa, T., Numari, A., Harris, M. J., Itoh, J., Watanabe, N., Mine, T., and Arias, I. M. (2007) Two N-linked glycans are required to maintain the transport activity of the bile salt export pump (ABCB11) in MDCK II cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **292,** G818–G828
44. Ishikawa, T. (2003) Multidrug resistance in cancer: Genetics of ABC transporters. In *Nature Encyclopedia of the Human Genome,* Vol. 4 (Cooper, D. N., ed) pp. 154-160, Nature Publishing Group, London
45. Borst, P., and Elferink, R. O. (2002) Mammalian ABC transporters in health and disease. *Annu. Rev. Biochem.* **71,** 537–592
46. Shehadeh, N. H., and Kligman, A. M. (1963) The effect of topical antibacterial agents on the bacterial flora of the axilla. *J. Invest. Dermatol.* **40,** 61–71
47. Julesz, M. (1968) New advances in the field of androgenic steroidogenesis of the human skin. *Acta Med. Acad. Sci. Hung.* **25,** 273–285
48. Brooksbank, B. W. (1970) Labelling of steroids in axillary sweat after administration of 3H-delta-5-pregnenolone and 14C-progesterone to a healthy man. *Experientia* **26,** 1012–1014
49. Labows, J. N., Preti, G., Hoelzle, E., Leyden, J., and Kligman, A. (1979) Steroid analysis of human apocrine secretion. *Steroids* **34,** 249–258
50. Ing, R., Petrakis, L., and Ho, H. C. (1973) Evidence against association between wet cerumen and breast cancer. *Lancet* **1,** 41

Received for publication February 24, 2009.
Accepted for publication April 9, 2009.

Gene 432 (2009) 97–101

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Developmentally dynamic changes of DNA methylation in the mouse *Snurf/Snrpn* gene

Kazumi Miyazaki [a], Christophe K. Mapendano [b], Tomokazu Fuchigami [c], Shinji Kondo [a], Tohru Ohta [d], Akira Kinoshita [b], Kazuhiro Tsukamoto [c], Ko-ichiro Yoshiura [b], Norio Niikawa [b], Tatsuya Kishino [a,*]

[a] Division of Functional Genomics, Center for Frontier Life Sciences, Nagasaki University, Sakamoto 1-12-4, Nagasaki 852-8523, Japan
[b] Department of Human Genetics, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, 852-8523, Japan
[c] Department of Pharmacotherapeutics, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, 852-8523, Japan
[d] Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, 061-0293, Japan

ARTICLE INFO

Article history:
Received 29 August 2008
Received in revised form 9 November 2008
Accepted 16 November 2008
Available online 27 November 2008

Received by T. Sekiya

Keywords:
Imprinting
Snrpn
DNA methylation
DMR
Enhancer

ABSTRACT

The mouse *Snurf/Snrpn* gene has two differentially methylated regions (DMRs), the maternally methylated region at the 5′ end (DMR1) and the paternally methylated region at the 3′ end (DMR2). DMR1, a region that includes the *Snrpn* promoter and the entire intron 1, has been thought to be a germline DMR, which inherits the parental-specific methylation profile from the gametes. DMR1 is not only associated with imprinted *Snrpn* expression, but implicated in imprinting control of other genes in the region. We have now characterized the highly conserved activator sequence (CAS) in the *Snrpn* intron 1 among human and rodents and demonstrate that the mouse CAS is not a germline DMR but shows developmentally dynamic changes of DNA methylation and has methylation-sensitive enhancer activity. The tissue-specific methylation of the mouse CAS and its methylation-sensitive enhancer activity may control tissue-specific expression of IC transcripts, resulting in the establishment and/or maintenance of imprinting in the *Snrpn* locus.

1. Introduction

Genomic imprinting is an important mechanism of gene regulation, which causes genetic nonequivalence in expression between maternal and paternal genomes in mammals. Such parent-of-origin specific gene regulation is caused by epigenetic modifications, which initially occur during gametogenesis without any nucleic acids changes (Surani, 1998; Tilghman, 1999). Epigenetic modifications in gametes continue to differentiate alleles of parental origin even after zygote formation, so that one parentally-derived allele eventually becomes preferentially expressed. One of the epigenetic modifications in imprinting is DNA methylation. DNA methylation can be stably inherited in somatic cells and reset in gametes. In the imprinted loci, differentially methylated regions (DMRs) between the maternal and paternal alleles are often found and associated with parent-allele-specific expression. For some imprinted loci, DMRs are gamete-derived methylated regions (germline DMRs), where DNA methylation in the gamete is maintained throughout development in all somatic lineages. However, there exist secondary DMRs, which are acquired during development and associated with primary imprints in the gamete (Constancia et al. 1998). Although the primary DMRs are essential for establishment and maintenance of imprinting, which are associated with the imprinting center (IC) (Bourc'his et al. 2001; Hata et al. 2002), it is unknown whether the secondary DMRs directly control the imprinted expression or exist only as the consequence of an epigenetic event.

The mouse chromosome 7C is a large imprinted domain orthologous to the Prader-Willi syndrome (PWS)/Angelman syndrome (AS) critical region at human chromosome 15q11-q13. The imprinted domain 7C contains paternally expressed genes, *Snurf/Snrpn* (hereafter termed *Snrpn*), *Ndn*, *Magel2*, *Mkrn3* and C/D-box small nucleolar RNAs (snoRNAs), and the maternally expressed gene, *Ube3a* (Nicholls and Knepper, 2001). Imprinted expression within this large domain is coordinated by a bipartite cis-acting IC located upstream from the *Snrpn* gene. In the large imprinted domain, several DMRs have been identified. One of them is in the *Snrpn* locus, which has two DMRs (Fig. 1A), the 5′ end methylated on the maternal allele (DMR1) and the 3′ end methylated on the paternal allele (DMR2) (Shemer et al. 1997; Gabriel et al. 1998). DMR1 is a ~6 kb region containing the 5′ end of the *Snrpn* gene and the entire *Snrpn* intron 1 (Fig. 1A). DMR2 is a 3.5-kb region spanning exons 7–10. Both DMR1 and DMR2 are thought as germline DMRs, which inherit the parental-specific methylation profile from the

Abbreviations: AS, Angelman syndrome; CAS, conserved activator sequence; ChIP, Chromatin immunoprecipitation; DHS, nuclease hypersensitive sites; DMR, differentially methylated region; IC, imprinting center; PWS, Prader-Willi syndrome; SMP, *Snrpn* minimal promoter; UPD, uniparental disomy.

* Corresponding author. Fax: +81 95 819 7178.
E-mail address: kishino@nagasaki-u.ac.jp (T. Kishino).

0378-1119/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2008.11.019