A replication study confirmed the *EDAR* gene to be a major contributor to population differentiation regarding head hair thickness in Asia

Akihiro Fujimoto[1, 2], Jun Ohashi[1, 3*], Nao Nishida[1], Taku Miyagawa[1], Yasuyuki Morishita[4], Tatsuhiko Tsunoda[2], Ryosuke Kimura[5], and Katsushi Tokunaga[1]

[1]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Hongo, Tokyo, Japan

[2]Center for Genomic Medicine, RIKEN, Yokohama, Japan

[3]Doctoral Program in Life System Medical Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Ibaraki, Japan

[4]Department of Molecular Pathology, Graduate School of Medicine, The University of Tokyo, Hongo, Tokyo, Japan

[5]Department of Forensic Medicine, Tokai University School of Medicine, Isehara, Kanagawa, Japan

*Corresponding author:

Dr. Jun Ohashi

Doctoral Program in Life System Medical Sciences

Graduate School of Comprehensive Human Sciences

University of Tsukuba

1-1-1 Tenno-dai, Tsukuba,

Ibaraki 305-8575, Japan

Tel: +81-29-853-5600 (ext. 91554)

E-mail: juno-tky@umin.ac.jp

Abstract

Hair morphology is a highly divergent phenotype among human populations. We recently reported that a nonsynonymous SNP in the *ectodysplasin A receptor* (*EDAR* 1540T/C) is associated with head hair fiber thickness in an ethnic group in Thailand (Thai-Mai) and an Indonesian population. However, these Southeast Asian populations are genetically and geographically close, and thus the genetic contribution of *EDAR* to hair morphological variation in the other Asian populations has remained unclear. In this study, we examined the association of 1540T/C with hair morphology in a Japanese population (Northeast Asian). As observed in our previous study, 1540T/C showed a significant association with hair cross-sectional area ($P = 2.7\times10^{-6}$) in Japanese. When all populations (Thai-Mai, Indonesian, and Japanese) were combined, the association of 1540T/C was stronger ($P = 3.8\times10^{-10}$) than those of age, sex, and population. These results indicate that *EDAR* is the genetic determinant of hair thickness as well as a strong contributor to hair fiber thickness variation among Asian populations.

Introduction

The rapid expansion and spread of the human species and resulting adaptation to various environments have produced several phenotypic differences among human populations. The genetic foundation of these traits and their evolutionary histories is one of the most interesting problems in human evolution. Although many human genes or polymorphisms have been suggested as candidates for positive selection (Akey et al. 2004; Carlson et al. 2005; The International HapMap Consortium 2005; Voight et al. 2006; Kimura et al. 2007; Barreiro et al. 2008), the genetic bases of physiological and morphological traits with diverged phenotypes have just begun to be identified (Lamason et al. 2005; Duffy et al. 2007; Stokowski et al. 2007; Sulem et al. 2007; Fujimoto et al. 2008).

Head hair morphology is one of the most differentiated traits among human populations. African individuals tend to have frizzier hair than European and Asian individuals, and Asian hair has a more circular and thicker cross-section than European and African hair (Hrdy 1973; Khumalo et al. 2000; Franbourg et al. 2003). However, the genetic determinants of human hair variation as well as its evolutionary history had remained to be elucidated. Recently, we demonstrated that an Asian specific single nucleotide polymorphism (SNP) rs3827760 in the *ectodysplasin A receptor* (*EDAR*), *EDAR* 1540T/C (370Val/Ala), is strongly associated with head hair fiber thickness in an ethnic group, Thai-Mai, in Thailand (THM) and an Indonesian population (IDN) (Fujimoto et al. 2008). In addition, using a reporter gene assay, we revealed that 1540T/C affects the activity of the downstream transcription factor NF-κB (Fujimoto et al. 2008). Taken together, we concluded that *EDAR* 1540T/C is one of the genetic determinants of Asian hair fiber thickness.

The chromosomal region of the *EDAR* gene has been identified as a strong candidate of recent positive selection in Northeast Asian populations (i.e., HapMap-CHB+JPT populations) (Carlson et al. 2005; Voight et al. 2006; Kimura et al. 2007; Sabeti et al. 2007; Barreiro et al. 2008; Bryk et al. 2008; Fujimoto et al. 2008), thus suggesting that Asian hair fiber thickness has evolved by local selection. Bryk et al. examined the world–wide distribution of *EDAR* 1540C and suggested that *EDAR* 1540C reached high frequency in East Asia prior to 10,000 years ago (Bryk et al. 2008). Since our previous association study was applied to genetically and geographically close Southeast Asian populations, the genetic contribution of 1540T/C to hair morphological variation in the other Asian populations has remained unclear. It is therefore important to assess the effect of *EDAR* 1540T/C on hair fiber thickness in Northeast Asian populations, such as Japanese and Chinese, to achieve a better understanding of the evolutionary impact of 1540C.

In the present study, we implemented a population-based association study in the Japanese population in order to answer the following questions: (1) Is there any difference in hair morphology, especially in hair fiber thickness, between the Northeast Asian and the Southeast Asian populations? (2) Is 1540T/C associated with hair thickness in the Northeast Asian population? (3) Is the effect of 1540T/C in the Northeast Asian population similar to that in the Southeast Asian population? (4) Does 1540T/C account for the largest amount of variation between the Northeast Asian and the Southeast Asian populations? (5) Does population stratification have any effect on the association with 1540T/C in Asian populations?

Materials and methods

Samples

Hair and DNA samples were gathered from 189 unrelated individuals (92 males and 97 females) in Tokyo, Japan, consisting of 109 individuals in their 20s, 68 in their 30s, 10 in their 40s, and 2 in their 50s, which are designated as JPN here. Ethical approval was obtained from the ethics committee of the medical faculty at the University of Tokyo, and informed consent was obtained from all subjects. DNA samples were extracted from mouth swabs. Hair samples were gathered from the back of the head of each individual. Hair samples were dissected and large diameter, small diameter and cross-sectional area were measured under microscopy (Fujimoto et al. 2008). Hair index, which is the ratio of small diameter to large diameter, was used to evaluate the shape of the cross-section (Fujimoto et al. 2008). We used the average value of five hairs from each individual for the following statistical analyses.

We also utilized DNA samples and hair morphological data from our previous study (Fujimoto et al. 2008) which came from 121 unrelated Indonesian (IDN) individuals in the west part of Java island, Indonesia and from 65 unrelated Thai-Mai (THM) individuals in the Rawai village of Phuket, Thailand. As a combined Southeast Asian population sample, they are referred to as SEA.

SNP selection and genotyping

The *EDAR* 1540T/C, a nonsynonymous SNP in exon12, and -1430A/G, a representative SNP in the promoter region, were genotyped by PCR-direct sequencing. Twenty three genomic control SNPs were selected based on the genome-wide SNP data of three human populations obtained from the International HapMap project (60 from Yorba in Ibadan, Nigeria, YRI; 60 from the CEPH population of northern and western European ancestry in Utah, CEU; 90 from Han Chinese in Beijing and Japanese in Tokyo, CHB+JPT) (The International HapMap Consortium 2005). To select

genomic control SNPs, we calculated the maximum value of F_{ST} (mF_{ST}) across all SNPs in a 50kbp window (Fujimoto et al. 2008). SNPs showing higher population differentiation than the 95th percentile of mF_{ST} under the empirical distribution between HapMap-CHB+JPT and other populations were employed as the genomic control SNPs. These SNPs were sufficiently distant so that they were thought not to be in linkage disequilibrium with each other. Genomic control SNPs were genotyped in SEA and JPN populations by means of DigTag2 method (Nishida et al. 2007).

Statistical analyses

The allele frequencies were estimated by gene counting. Deviation from Hardy-Weinberg equilibrium was examined by chi-square test. Here, we analyzed the combined SEA group, since our previous regression analysis did not show there to be significant population differentiation between the IDN and THM groups (Fujimoto et al. 2008). Comparisons of the cross-sectional area, small diameter, large diameter, and hair index between JPN and SEA were carried out by t-test. Associations between SNPs in *EDAR* (*EDAR* 1540T/C and -1430G/A) and hair morphology were evaluated by ANOVA. The effects of the copy number of *EDAR* 1540C allele (i.e., 0, 1, or 2), age, and sex on hair morphology in JPN were examined by using a stepwise multiple regression analysis, where the criteria for variable selection (F_{IN}) and rejection (F_{OUT}) were set at 4.0. Furthermore, the effect of population (SEA vs. JPN) on hair cross-sectional area was also evaluated by using a stepwise multiple regression analysis with the copy number of *EDAR* 1540C allele (i.e., 0, 1, or 2), age, sex, and population (SEA or JPN) as independent variables. To examine the possibility of population stratification, a multiple regression analysis on hair cross-sectional area was performed

using the target allele's copy number (i.e., 0, 1, or 2) of the genomic control SNP, age, sex, and population (SEA or JPN) as independent variables.

Results

Association between EDAR and hair morphology in the Japanese population

We gathered DNA and hair samples from 189 Japanese individuals (JPN) and measured their hair cross-sectional area, small diameter, and large diameter under microscopy. For the following analyses, we also examined 121 Indonesian (IDN) and 65 Thai-Mai individuals (THM) as a combined Southeast Asian (SEA) population. The genotype frequencies of *EDAR* 1540T/C and -1430G/A were consistent with expectations under Hardy-Weinberg equilibrium. The allele frequency of 1540T/C in JPN (78.6%) agreed with that in the HapMap-JPT population (79.5%) (Table 1). To the best of our knowledge, this research provides the first comparative analysis of head hair cross-sectional area and diameters for a number of individuals from Asian populations. As shown in Table 2, JPN individuals had significantly greater values for cross-sectional area, small diameter, large diameter, and hair index (i.e., the ratio of small diameter to large diameter) than SEA individuals (t-test: cross-sectional area $P < 2\times10^{-16}$, small diameter $P < 2\times10^{-16}$, large diameter $P = 3.4\times10^{-16}$, hair index $P = 0.0029$).

In JPN, 1540T/C showed significant associations with cross-sectional area, small diameter, and large diameter (ANOVA: cross-sectional area $P = 1.4\times10^{-5}$, small diameter $P = 4.9\times10^{-5}$, and large diameter $P = 0.0010$) but not with hair index (Fig. 1). The Asian specific 1540C allele increases hair cross-sectional area and appears to have a codominant effect (Fig. 1). On the other hand, -1430G/A showed no significant association with any of the variables (data not shown).

To evaluate other factors, stepwise multiple regression analysis using the copy number of 1540C (i.e., 0, 1, or 2), age, and sex as independent variables was performed for JPN. The 1540T/C and age were finally included in the model. The copy number of 1540C was positively associated with cross-sectional area, small diameter and large diameter, and increasing age was significantly associated with cross-sectional area, large diameter, and hair index (Table 3). In particular, cross-sectional area exhibited the strongest association ($P = 2.7\times10^{-6}$) (Table 3). The effect of one copy of 1540C on the cross-sectional area in the JPN population (667μm^2) seemed to be a little larger than that in the SEA population (491μm^2) (Fujimoto et al. 2008).

The effects of population and 1540T/C on hair cross-sectional area

To assess whether hair fiber thickness is affected by populations (i.e., SEA vs. JPN), we further carried out multiple regression analysis on hair cross-sectional area with the following independent variables: the copy number of 1540C (i.e., 0, 1, or 2), age, sex, and population (SEA or JPN). Interestingly, the effect of population (i.e., SEA vs. JPN) was highly significant ($P = 4.3\times10^{-6}$) even when adjusting for 1540T/C (Table 4). The standardized regression coefficient (SRC) of 1540T/C was larger than that of the others, and 1540T/C showed the lowest *P* value ($P = 3.8\times10^{-10}$) in comparison to the other factors (Table 4)

The possibility of population stratification

Since hair morphology is a highly divergent phenotype among populations (Hrdy 1973; Franbourg et al. 2003), population stratification might result in a false positive association. Although this possibility was explored using one SNP with high population differentiation in our previous study

(Fujimoto et al. 2008), we reviewed the possibility by using 23 additional genomic control SNPs. We selected these “genomic control SNPs” based on the same criteria of the previous candidate gene selection; SNPs with higher population differentiation under the empirical distribution between HapMap-JPT+CHB and other populations were employed. The selected SNPs had a minor allele frequency of more than 0.05 in the JPN and SEA populations and were consistent with expectations under Hardy-Weinberg equilibrium (Supplementary Table S1). No significant association with hair cross-sectional area was observed for all the genomic control SNPs after correcting for multiple testing of 25 SNPs (1540T/C, -1430G/A, and 23 control SNPs) (Fig. 2 and Supplementary Table S2).

Discussion

Previously, we reported that *EDAR* 1540T/C is a genetic determinant of Asian hair fiber thickness and evolved by recent positive selection. To confirm the association in Southeast Asians and to evaluate the contribution of 1540T/C to the difference in hair fiber thickness among Asian populations, we carried out an association study in a set of newly collected Japanese samples. First, we compared hair morphology between Japanese and the Southeast Asian (SEA) populations. Among four traits (cross-sectional area, small diameter, large diameter, and hair index), cross-sectional area was found to be most divergent; JPN individuals have more than 30% larger mean cross-sectional area (6518μm^2) than SEA (4957μm^2) (Table 2) and more than 50% larger than Africans (4274μm^2) and Caucasians (3857μm^2) (12).

As observed in our previous study in the Southeast Asian populations, 1540T/C showed the strongest association with cross-sectional area in the JPN population. The multiple regression

analyses showed that increasing age, 1540T/C, and population differentiation have an influence on hair fiber thickness (Table 4). Among these factors, 1540T/C showed the lowest *P* value, thus suggesting that 1540T/C has a relatively stronger effect on hair thickness than the other factors (Table 4). However, since the population differentiation between JPN and SEA was also significant, 1540T/C by itself cannot explain all the differentiation of hair fiber thickness between JPN and SEA. In other words, other genetic and/or environmental factors may play a role in hair fiber thickness variation among Asian populations. We speculate that in Northeast Asians there may be other highly frequent genetic variants associated with the increase in hair fiber thickness.

In this study, 23 genomic control SNPs were genotyped to examine the possibility of population stratification. Although the number of genomic control SNPs is not large, we expected these SNPs to be sensitive to population stratification because of high levels of population differentiation between the JPN and SEA populations (see Materials and methods). Since no SNPs other than 1540T/C showed significant association with hair fiber thickness (Fig. 2), we can say that although the population history or genetic structure of the studied populations is not fully understood, the association of 1540T/C with hair cross-sectional area is not a false positive finding stemming from population stratification. These results lead us to conclude that *EDAR* 1540T/C is the major genetic determinant of hair fiber thickness and can explain a large part of the variation of hair fiber thickness in Asians.

Recently, Bryk et al. and Mou et al. reported that the 1540C allele enhances NF-κB signaling *in vitro* (Bryk et al. 2008; Mou et al. 2008). These results are opposite to our previous finding (Fujimoto et al. 2008). The difference may be caused by the different experimental time-courses. Since we measured luciferace activity 48 hours after transfection but Bryk et al. and

Mou et al. did so after 18 hours, feedback-regulation might have influenced *EDAR* signaling in our previous result. Further investigation of various luciferase assay time-courses along with measurement of the expression level of EDAR target molecules would likely improve our understanding of functional differences involving the 1540T/C polymorphism. In addition to the *in vitro* study, Mou et al. showed that elevation of Edar activity in transgenic mouse leads to increased hair fiber thickness and circularizes hair cross-sectional area (Mou et al. 2008). In our reports, 1540T/C is associated with hair fiber thickness, but not with the shape of hair cross-section.

Although 1540T/C is the major genetic contributor to hair fiber thickness variation in Asian populations, the possible impact of *EDAR* on other human phenotypes still remains to be elucidated. Interestingly, recent evolutionary studies reported *EDA*, which encodes the ligand of EDAR, to be clearly related to the body armor plate morphology of stickleback (Colosimo et al. 2005) and that the evolutionary pattern of the EDA-pathway corresponds to the morphological diversification of vertebrate skin appendages (Pantalacci et al. 2008). Moreover, *EDAR* is a promising candidate as a genetic determinant of human tooth and/or gland variations (Schmidt-Ullrich and Paus 2005). Therefore, *EDAR* could prove to be a key locus that has led to a number of morphological changes in Asians.

ACKNOWLEDGEMENTS We are deeply grateful to the people participating in this study. We also gratefully acknowledge Mr. Yusuke Ohnishi, Mr. Chiaki Miura, Mr. Masahiko Kumagai, Mr. Kazuo Ebine, Ms. Mihoko Shimada and Ms. Takako Fujimoto for their valuable cooperation in our sample collection. We deeply appreciate to Mr. Todd A. Johnson and two anonymous referees

for helpful comments on the manuscript. This study was partly supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to J.O.).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

References

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L (2004) Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol 2: e286

Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L (2008) Natural selection has driven population differentiation in modern humans. Nat Genet

Bryk J, Hardouin E, Pugach I, Hughes D, Strotmann R, Stoneking M, Myles S (2008) Positive selection in East Asians for an EDAR allele that enhances NF-kappaB activation. PLoS ONE 3: e2209

Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ, Nickerson DA (2005) Genomic regions exhibiting positive selection identified from dense genotype data. Genome Res 15: 1553-65

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Jr., Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science 307: 1928-33

Duffy DL, Montgomery GW, Chen W, Zhao ZZ, Le L, James MR, Hayward NK, Martin NG, Sturm RA (2007) A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. Am J Hum Genet 80: 241-52

Franbourg A, Hallegot P, Baltenneck F, Toutain C, Leroy F (2003) Current research on ethnic hair. J Am Acad Dermatol 48: S115-9

Fujimoto A, Kimura R, Ohashi J, Omi K, Yuliwulandari R, Batubara L, Mustofa MS, Samakkarn U, Settheetham-Ishida W, Ishida T, Morishita Y, Furusawa T, Nakazawa M,

Ohtsuka R, Tokunaga K (2008) A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. Hum Mol Genet 17: 835-43

Hrdy D (1973) Quantitative hair form variation in seven populations. Am J Phys Anthropol 39: 7-17

Khumalo NP, Doe PT, Dawber RP, Ferguson DJ (2000) What is normal black African hair? A light and scanning electron-microscopic study. J Am Acad Dermatol 43: 814-20

Kimura R, Fujimoto A, Tokunaga K, Ohashi J (2007) A practical genome scan for population-specific strong selective sweeps that have reached fixation. PLoS ONE 2: e286

Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynec MJ, Mao X, Humphreville VR, Humbert JE, Sinha S, Moore JL, Jagadeeswaran P, Zhao W, Ning G, Makalowska I, McKeigue PM, O'Donnell D, Kittles R, Parra EJ, Mangini NJ, Grunwald DJ, Shriver MD, Canfield VA, Cheng KC (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science 310: 1782-6

Mou C, Thomason HA, Willan PM, Clowes C, Harris WE, Drew CF, Dixon J, Dixon MJ, Headon DJ (2008) Enhanced ectodysplasin-A receptor (EDAR) signaling alters multiple fiber characteristics to produce the East Asian hair form. Hum Mutat (DOI: 10.1002/humu.20795)

Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 364: 78-85

Pantalacci S, Chaumot A, Benoit G, Sadier A, Delsuc F, Douzery EJ, Laudet V (2008) Conserved features and evolutionary shifts of the EDA signaling pathway involved in vertebrate skin appendage development. Mol Biol Evol 25: 912-28

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L,

Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, et al. (2007) Genome-wide detection and characterization of positive selection in human populations. Nature 449: 913-8

Schmidt-Ullrich R, Paus R (2005) Molecular principles of hair follicle induction and morphogenesis. Bioessays 27: 247-61

Stokowski RP, Pant PV, Dadd T, Fereday A, Hinds DA, Jarman C, Filsell W, Ginger RS, Green MR, van der Ouderaa FJ, Cox DR (2007) A genomewide association study of skin pigmentation in a South Asian population. Am J Hum Genet 81: 1119-32

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, Jakobsdottir M, Steinberg S, Palsson S, Jonasson F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Kiemeney LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K (2007) Genetic determinants of hair, eye and skin pigmentation in Europeans. Nat Genet 39: 1443-52

The International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437: 1299-320

Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS Biol 4: e72

Figure legends

Figure1. *EDAR* 1540T/C and hair morphology in JPN. (A) Cross-sectional area ANOVA $P = 1.4\times10^{-5}$ (B) Small diameter ANOVA $P = 4.9\times10^{-5}$ (C) Large diameter ANOVA $P = 0.0010$ (D) Hair index.

Figure 2. Association of *EDAR* and control SNPs with hair cross-sectional area. *P* values were calculated for *EDAR* 1540T/C, -1430G/A and the 23 genomic control SNPs based on a multiple regression analysis using the target allele's copy number (i.e., 0, 1, or 2), age, sex, and population (SEA or JPN) as independent variables.

Table 1. Genotype and allele frequencies for *EDAR* -1430G/A and 1540T/C in JPN

SNP	Allele A	Allele B	AA	AB	BB	A	B
EDAR-1430	G	A	134 (71.7%)	46 (24.6%)	7 (3.7%)	84.0%	16.0%
*EDAR*1540	T	C	14 (7.4%)	53 (28.0%)	122 (64.6%)	21.4%	78.6%

Allele A and B represent the ancestral and derived allele, respectively.

Table 2. Summary statistics of hair morphology

	Sex		Age		Area (μm^2)		Small diameter (μm)		Large diameter (μm)		Hair index	
Population	M	F	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
JPN	92	97	-	-	6518	1290	76.99	13.03	102.49	8.66	0.765	0.092
SEA	108	78	38.23	12.79	4957	1228	65.93	8.80	90.93	13.22	0.738	0.0823

SEA: Southeast Asian population (Indonesian and Thai-Mai ethnic group in Thailand).

JPN consists of 109 individuals in their 20s, 68 in their 30s, 10 in their 40s and 2 in their 50s.

Table 3. Multiple regression analyses on hair morphology in JPN (n = 189)

	Area (μm^2)				Small diameter (μm)			
Explanatory variables	RC	SRC	t	*P* value	RC	SRC	t	*P* value
Age	-44.3	-0.22	-3.32	0.0011	-	-	-	-
Sex (M:0, F:1)	-	-	-	-	-	-	-	-
EDAR 1540T/C (TT:0, TC:1, CC:2)	667.6	0.33	4.84	2.7×10^{-6}	4.6	0.32	4.59	8.2×10^{-6}

	Large diameter (μm)				Hair Index			
Explanatory variables	RC	SRC	t	*P* value	RC	SRC	t	*P* value
Age	-0.5	-0.26	-3.75	2.4×10^{-4}	0.0024	0.17	2.31	0.022
Sex (M:0, F:1)	-	-	-	-	-	-	-	-
EDAR 1540T/C (TT:0, TC:1, CC:2)	5.3	0.25	3.73	2.5×10^{-4}	-	-	-	-

A stepwise method (F_{IN} and F_{OUT}: 4) was used in the multiple regression analyses.

RC: regression coefficient.

SRC: Standardized regression coefficient.

-: a variable excluded in the stepwise procedure.

Table 4. Multiple regression analysis on hair cross-sectional area in JPN and SEA populations (n = 375)

Explanatory variables	RC	SRC	t	*P* value
Age	-34.7	-0.26	-5.86	1.0×10^{-8}
Sex (M:0, F:1)	-	-	-	-
Population (SEA:0, JPN:1)	704.2	0.24	4.67	4.3×10^{-6}
EDAR 1540T/C (TT:0, TC:1, CC:2)	571.4	0.31	6.43	3.8×10^{-10}

SEA: Southeast Asian populations (Indonesian and Thai-Mai ethnic group in Thailand).

A stepwise method (F_{IN} and F_{OUT}: 4) was used in the multiple regression analyses.

RC: regression coefficient.

SRC: Standardized regression coefficient.

-: A variable excluded in the stepwise procedure.

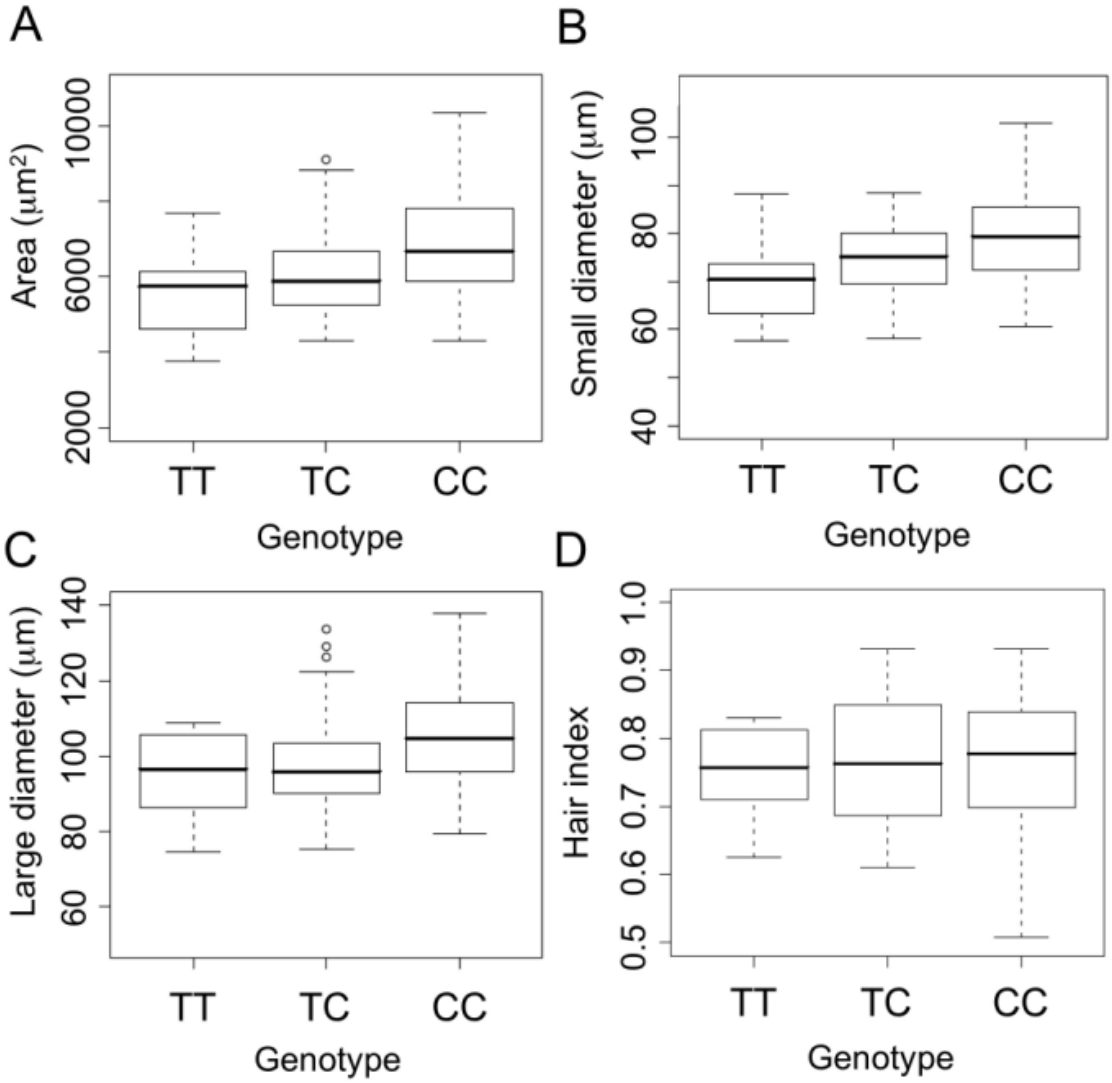
A
Area (μm²)
2000
6000
10000
TT
TC
CC
Genotype
B
Small diameter (μm)
40
60
80
100
TT
TC
CC
Genotype
C
Large diameter (μm)
60
80
100
120
140
TT
TC
CC
Genotype
D
Hair index
0.5
0.6
0.7
0.8
0.9
1.0
TT
TC
CC
Genotype

Figure 1.

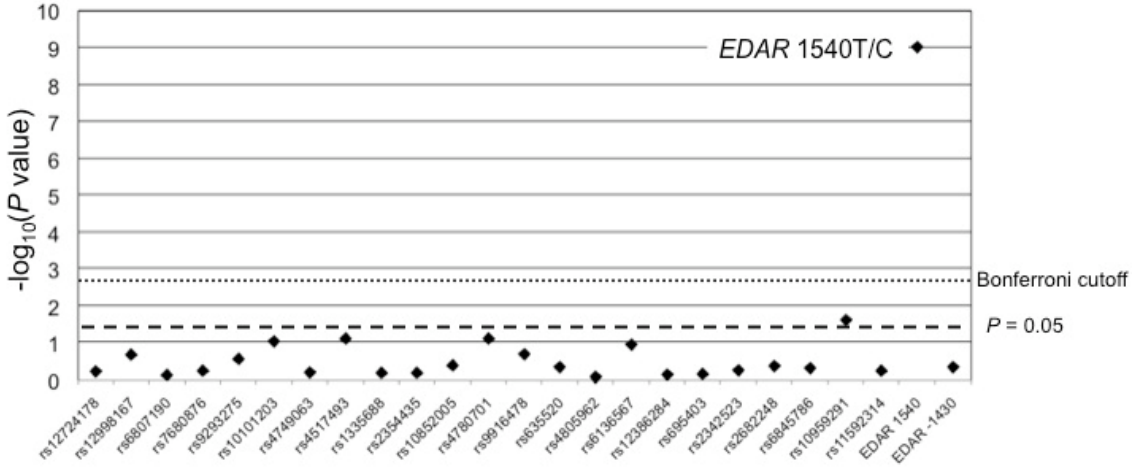
10
9
8
7
6
5
4
3
2
1
0
-log10(P value)
EDAR 1540T/C
Bonferroni cutoff
P = 0.05
rs12724178
rs12998167
rs6807190
rs7680876
rs9293275
rs10101203
rs4749063
rs4517493
rs1335688
rs2354435
rs10852005
rs4780701
rs9916478
rs635520
rs4805962
rs6136567
rs12386284
rs695403
rs2342523
rs2682248
rs6845786
rs10959291
rs11592314
EDAR 1540
EDAR -1430

Figure 2.

A list of abbreviations

EDAR; *ectodysplasin A receptor*

JPN; 189 participants from Tokyo, Japan

THM; 65 participants from Thai-Mai, in Thailand

IDN; 121 participants from Indonesia

SEA; Southeast Asians (THM and IDN)

TableS1. Genotype and allele frequencies of 23 genomic control SNPs, *EDAR*-1430G/A and 1540T/C

			HapMap						JPN						IDN						THM					
	Allele		Frequency of allele A			F_{ST} value			Genotype frequency			Allele frequency		HWE	Genotype frequency			Allele frequency		HWE	Genotype frequency			Allele frequency		HWE
	A	B	CEU	YRI	CHB+JPT	CY[a]	YA[b]	CA[c]	AA	AB	BB	A	B	*P*	AA	AB	BB	A	B	*P*	AA	AB	BB	A	B	*P*
rs12724178	A	G	0.69	0.05	0.87	0.44	0.68	0.05	154	35	1	0.90	0.10	0.80	91	26	0	0.89	0.11	0.40	53	12	0	0.91	0.09	0.71
rs12998167	A	G	0.24	0.02	0.90	0.11	0.79	0.44	141	27	1	0.91	0.09	0.97	87	28	3	0.86	0.14	0.92	44	19	1	0.84	0.16	0.80
rs6807190	A	G	0.68	0.98	0.18	0.15	0.66	0.26	10	69	112	0.23	0.77	0.99	1	30	87	0.14	0.86	0.66	2	17	46	0.16	0.84	0.96
rs7680876	A	T	0.40	0.97	0.14	0.37	0.70	0.09	3	28	150	0.09	0.91	0.47	9	42	65	0.26	0.74	0.83	5	27	32	0.29	0.71	0.98
rs9293275	A	G	0.59	1.00	0.20	0.26	0.67	0.16	8	59	124	0.20	0.80	0.96	11	48	59	0.30	0.70	0.96	2	14	49	0.14	0.86	0.74
rs10101203	A	G	0.69	0.00	0.81	0.53	0.68	0.02	131	52	8	0.82	0.18	0.63	67	42	9	0.75	0.25	0.80	37	21	7	0.73	0.27	0.35
rs4749063	C	T	0.69	0.06	0.88	0.43	0.68	0.05	156	35	0	0.91	0.09	0.38	79	35	4	0.82	0.18	1.00	48	16	1	0.86	0.14	0.97
rs4517493	A	G	0.45	0.98	0.11	0.34	0.76	0.15	3	59	129	0.17	0.83	0.43	1	13	104	0.06	0.94	0.72	0	4	60	0.03	0.97	0.97
rs1335688	C	T	0.16	0.00	0.84	0.09	0.72	0.46	104	78	9	0.75	0.25	0.50	44	49	25	0.58	0.42	0.28	26	31	8	0.64	0.36	0.97
rs2354435	C	T	0.38	1.00	0.20	0.45	0.67	0.04	2	45	143	0.13	0.87	0.76	10	62	46	0.35	0.65	0.23	6	26	29	0.31	0.69	1.00
rs10852005	A	C	0.45	0.08	0.88	0.18	0.64	0.20	150	37	4	0.88	0.12	0.64	99	17	1	0.92	0.08	0.96	46	18	0	0.86	0.14	0.42
rs4780701	C	T	0.33	0.00	0.81	0.20	0.68	0.23	133	49	9	0.82	0.18	0.29	57	42	19	0.66	0.34	0.08	35	20	7	0.73	0.27	0.33
rs9916478	A	G	0.97	0.83	0.33	0.05	0.26	0.45	14	74	103	0.27	0.73	0.99	16	57	45	0.38	0.62	0.95	10	30	21	0.41	0.59	0.99
rs635520	A	T	0.19	0.68	0.84	0.24	0.04	0.42	135	54	2	0.85	0.15	0.40	83	32	2	0.85	0.15	0.86	40	13	7	0.78	0.23	0.01
rs4805962	C	T	0.08	0.36	0.90	0.11	0.31	0.67	146	27	0	0.92	0.08	0.54	66	43	4	0.77	0.23	0.64	19	20	10	0.59	0.41	0.55
rs6136567	C	G	0.91	0.76	0.25	0.04	0.26	0.45	14	87	90	0.30	0.70	0.52	4	25	89	0.14	0.86	0.43	1	12	52	0.11	0.89	0.95
rs12386284	G	T	0.78	0.43	0.14	0.13	0.10	0.42	8	46	137	0.16	0.84	0.29	4	27	87	0.15	0.85	0.59	0	13	52	0.10	0.90	0.67
rs695403	A	G	0.18	0.76	0.82	0.34	0.01	0.42	113	67	11	0.77	0.23	0.97	61	43	12	0.71	0.29	0.58	20	33	12	0.56	0.44	0.97
rs2342523	A	T	0.14	0.32	0.81	0.05	0.24	0.44	118	61	11	0.78	0.22	0.71	78	36	3	0.82	0.18	0.89	47	16	2	0.85	0.15	0.91
rs2682248	G	T	0.94	0.30	0.11	0.44	0.06	0.70	6	47	137	0.16	0.84	0.73	8	45	65	0.26	0.74	1.00	6	25	32	0.29	0.71	0.94
rs6845786	C	G	0.16	0.95	0.83	0.64	0.04	0.45	130	55	6	0.82	0.18	1.00	79	33	6	0.81	0.19	0.59	54	10	0	0.92	0.08	0.79
rs10959291	C	T	0.14	0.48	0.82	0.14	0.13	0.47	135	44	10	0.83	0.17	0.06	4	34	80	0.18	0.82	0.99	15	18	32	0.37	0.63	0.00
rs11592314	G	T	0.93	0.93	0.31	0.00	0.40	0.41	38	78	74	0.41	0.59	0.12	19	56	43	0.40	0.60	0.99	42	20	3	0.80	0.20	0.95
EDAR 1540	C	T	0.00	0.00	0.88	0.00	0.78	0.78	122	53	14	0.79	0.21	0.07	18	46	57	0.34	0.66	0.25	7	25	33	0.30	0.70	0.79
EDAR -1430	G	A	-	-	-	-	-	-	134	46	7	0.16	0.84	0.49	49	40	32	0.57	0.43	1.6×10^{-3}	28	27	10	0.64	0.36	0.72

[a]mF_{ST} value CEU vs. YRI

[b]mF_{ST} value Asia (CHB+JPT) vs. YRI

[c]mF_{ST} value Asia (CHB+JPT) vs. CEU

95th percentile values of mF_{ST} in the empirical distribution are 0.63 and 0.40 in YA and CA, respectively.

-: *EDAR* -1430 was genotyped in this study.

Table S2. Results of multiple regression analysis on hair cross-sectional area in JPN+SEA (n=375)

	Age			Sex (M:0, F:1)			Population (SEA:0, JPN:1)			Genotype		
	RC	t	*P* value	RC	t	*P* value	RC	t	*P* value	RC	t	*P* value
rs12724178	-39.3	-6.25	1.1×10^{-9}	-274.5	-2.19	0.030	1240.5	9.04	$<2.0\times10^{-16}$	75.9	0.5	0.62
rs12998167	-38.0	-6.00	5.1×10^{-9}	-314.6	-2.42	0.016	1299.0	9.07	$<2.0\times10^{-16}$	-174.2	-1.2	0.22
rs6807190	-38.6	-6.12	2.4×10^{-9}	-275.1	-2.20	0.029	1227.0	8.79	$<2.0\times10^{-16}$	32.3	0.3	0.78
rs7680876	-37.8	-5.98	5.5×10^{-9}	-309.2	-2.44	0.015	1268.2	8.78	$<2.0\times10^{-16}$	-62.3	-0.5	0.59
rs9293275	-38.4	-6.13	2.2×10^{-9}	-266.7	-2.13	0.034	1242.9	9.05	$<2.0\times10^{-16}$	111.1	1.1	0.29
rs10101203	-37.7	-6.04	3.9×10^{-9}	-274.1	-2.20	0.029	1212.1	8.82	$<2.0\times10^{-16}$	171.9	1.7	0.094
rs4749063	-38.3	-6.12	2.4×10^{-9}	-275.9	-2.21	0.028	1243.7	8.96	$<2.0\times10^{-16}$	61.1	0.5	0.65
rs4517493	-39.5	-6.35	6.3×10^{-10}	-290.6	-2.34	0.020	1301.6	9.29	$<2.0\times10^{-16}$	-253.4	-1.8	0.080
rs1335688	-38.3	-6.11	2.6×10^{-9}	-274.2	-2.19	0.029	1223.3	8.75	$<2.0\times10^{-16}$	-38.7	-0.4	0.68
rs2354435	-37.7	-5.93	7.1×10^{-9}	-307.3	-2.46	0.014	1274.0	8.82	$<2.0\times10^{-16}$	-47.7	-0.4	0.67
rs10852005	-38.3	-6.09	2.9×10^{-9}	-269.7	-2.15	0.033	1240.6	8.99	$<2.0\times10^{-16}$	113.9	0.8	0.42
rs4780701	-37.6	-5.88	9.4×10^{-9}	-280.0	-2.23	0.026	1195.7	8.56	3.2×10^{-16}	-168.9	-1.8	0.080
rs9916478	-38.6	-6.13	2.4×10^{-9}	-299.6	-2.40	0.017	1210.4	8.72	$<2.0\times10^{-16}$	-118.2	-1.2	0.21
rs635520	-38.3	-5.70	2.5×10^{-8}	-261.2	-2.06	0.040	1211.9	8.73	$<2.0\times10^{-16}$	86.3	0.7	0.47
rs4805962	-39.7	-6.07	3.5×10^{-9}	-338.6	-2.56	0.011	1285.5	8.50	6.9×10^{-16}	21.3	0.2	0.87
rs6136567	-38.7	-6.20	1.6×10^{-9}	-282.8	-2.27	0.024	1292.0	9.12	$<2.0\times10^{-16}$	173.2	1.6	0.12
rs12386284	-38.2	-6.09	2.9×10^{-9}	-276.8	-2.21	0.028	1233.2	8.98	$<2.0\times10^{-16}$	-37.3	-0.3	0.76
rs695403	-38.5	-6.10	2.7×10^{-9}	-282.9	-2.24	0.026	1237.9	8.82	$<2.0\times10^{-16}$	-33.3	-0.3	0.73
rs2342523	-37.8	-6.06	3.4×10^{-9}	-288.7	-2.32	0.021	1221.4	8.89	$<2.0\times10^{-16}$	60.9	0.6	0.58
rs2682248	-37.8	-6.08	3.0×10^{-9}	-240.2	-1.93	0.054	1225.3	8.87	$<2.0\times10^{-16}$	-81.7	-0.8	0.44
rs6845786	-38.0	-6.06	3.4×10^{-9}	-278.5	-2.22	0.027	1238.1	9.00	$<2.0\times10^{-16}$	-77.0	-0.7	0.51
rs10959291	-37.5	-6.00	4.7×10^{-9}	-273.5	-2.19	0.029	1497.6	8.26	2.7×10^{-15}	-223.2	-2.3	0.025
rs11592314	-38.4	-6.12	2.4×10^{-9}	-294.7	-2.35	0.019	1236.4	8.96	$<2.0\times10^{-16}$	-44.9	-0.5	0.59
EDAR 1540	-34.8	-5.90	8.5×10^{-9}	-226.9	-1.91	0.057	737.7	4.87	1.6×10^{-6}	557.5	6.3	9.6×10^{-10}
EDAR -1430	-38.7	-6.28	9.9×10^{-10}	-272.4	-2.19	0.030	1164.7	8.01	1.50×10^{-14}	-66.8	-0.7	0.47

RC: regression coefficient.

In the multiple regression analyses, the copy number of target allele (i.e., 0, 1, or 2), age, sex, and population were utilized as independent variables.