EDA2R Is Associated with Androgenetic Alopecia

Dionigio Antonio Prodi[1,4], Nicola Pirastu[1,2,4], Giuseppe Maninchedda[1], Alessandro Sassu[1], Andrea Picciau[1], Maria Antonietta Palmas[1], Alessandra Mossa[1], Ivana Persico[1], Mauro Adamo[1], Andrea Angius[1,3] and Mario Pirastu[1,3]

Androgenetic alopecia (AGA) is a common heritable polygenic disorder whose genetics is not fully understood, even though it seems to be X-linked. We carried out an epidemiological survey for AGA on 9,000 people from 8 isolated villages of a secluded region of Sardinia (Ogliastra), and identified a large cohort of affected individuals. We genotyped 200 cases and 200 controls (mean kinship 0.001) with the 500k chip array and conducted case–control association analysis on the X chromosome. We identified Xq11-q12 as strongly associated with AGA. In particular, we found that rs1352015 located 8 kb from the *EDA2R* gene showed the best result ($P = 7.77e^{-7}$). This region also contains the *AR* gene, hence we tested both genes in 492 cases and 492 controls. We found that the non-synonymous SNP rs1385699 on *EDA2R* gave the best result ($P = 3.9e^{-19}$) whereas rs6152 on the AR gene is less significant ($P = 4.17e^{-12}$). Further statistical analysis carried out by conditioning each gene to the presence of the other showed that the association with *EDA2R* is independent while the association with AR seems to be the result of linkage disequilibrium. These results give insight into the pathways involved in AGA etiology.

Journal of Investigative Dermatology (2008) **128,** 2268–2270; doi:10.1038/jid.2008.60; published online 3 April 2008

Androgenetic alopecia (AGA) is characterized by hair loss that affects up to 50% of all males (Hamilton, 1951). Although it is generally accepted that it is a polygenic heritable trait, a clearly responsible gene has yet to be identified (Kuster and Happle, 1984; Ellis *et al.*, 1998). We carried out an epidemiological survey for AGA in eight villages of Ogliastra, a secluded area of central Sardinia (Angius *et al.*, 2001). The population of each village is characterized by high endogamy, little immigration, and slow population growth, and there have been few marriage exchanges among the villages during the centuries. This was proven both by genealogical reconstructions and through genetic studies, that is, of mitochondrial DNA. For this reason, each village can be considered independently of the others (Fraumene *et al.*, 2003). To calculate prevalence, we selected males older than 18 years and with a grade higher than IIv on the Norwood–Hamilton hair-loss scale. We found that the average mean prevalence in men was 47%, varying from 39% in the village of Seui to 56% in Talana. Given the large number of collected samples (9,000), it was possible to select the most severe cases of AGA. We picked men who had a baldness grade of at least IV on the Norwood–Hamilton scale and had onset before 30 years of age. For controls, we selected men who were at least 40 years old at the time of the visit and had no evidence of AGA. Using these parameters, we selected 200 cases and 200 controls; samples were selected in equal number (25 cases and 25 controls) from each of the eight villages to avoid population stratification. Within each village we selected cases and controls so that they were the most distantly related as possible, and so that cases did not have greater kinship than controls (averaged mean kinship in cases = 0.0016, SD = 0.01; average mean kinship in controls = 0.0011, SD = 0.01). The algorithm used, involving sampling from the many pairwise relationships present in extended genealogies, is described by Falchi *et al.* (2004). We genotyped these samples with the Affymetrix-GeneChip human mapping 500 k array. The minimum call rate for each individual was 93%; for each single nucleotide polymorphism (SNP) it was 90%. All individuals participating in the study signed informed consent forms, and all the samples were taken in accordance with the Declaration of Helsinki Principles (http://www.wma.net/e/policy/17-c_e.html). Institutional approval was not required for experiments.

We decided to investigate the association first on the X chromosome, not only because previous studies have associated this chromosome with AGA but also because, on a first genome-wide scan, the most positive signal was on this chromosome. We tested 7,093 SNPs on the whole X chromosome using Quasi-likelihood score test for case-contol (CCQLS) (Bourgain *et al.*, 2003), a software that permits the correction of the association analysis results by the kinship matrix. We used this method because the population is inbred, and, although samples were selected to avoid spurious association due to kinship, we cannot exclude cryptic relatedness not accounted for. We used markers with

[1]*Shardna Life Sciences, Pula, Italy;* [2]*Dipartimento di Scienze Applicate ai Biosistemi, Università degli studi di Cagliari, Cagliari, Italy and* [3]*Istituto di Genetica delle Popolazioni CNR Alghero, Alghero, Italy*

[4]*These authors contributed equally to this work*

Correspondence: Dr Mario Pirastu, Shardna Spa, Edificio 3, Località Piscinamanna, Pula 9010, Italy. E-mail: pirastu@shardna.it

Abbreviations: CCQLS, Quasi-likelihood score test for case–control; MAF, minimum allele frequency; STR, short tandem repeat

Received 11 October 2007; revised 18 January 2008; accepted 9 February 2008; published online 3 April 2008

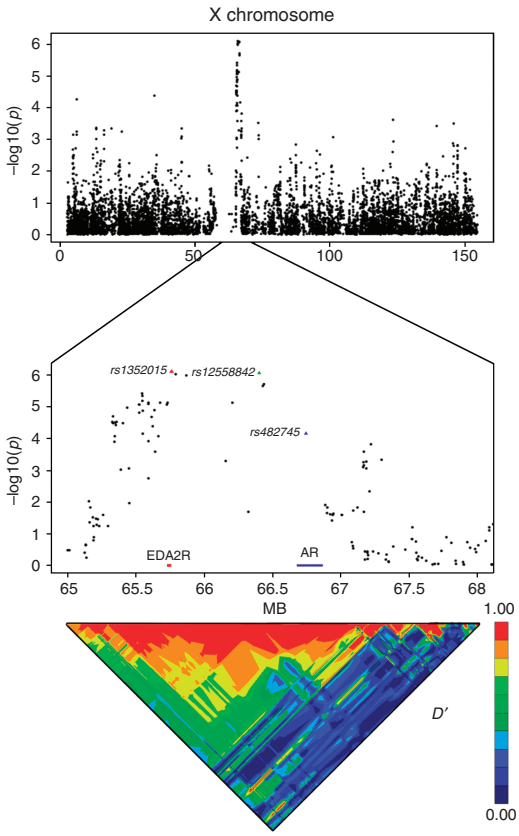

Figure 1. Association results on X chromosome. The upper figure shows the −log10(p) for all the SNPs tested on the whole chromosome. The lower figure shows enlargement of the most significant region. The LD (D′) plot is shown below. The solid bars represent the genes.

minimum allele frequency (MAF) >0.025 in the whole sample because we did not have enough statistical power to detect association with rarer variants. We found that several markers in the Xq11–q12 region (Figure 1, Table S1) were strongly associated with AGA: in particular, rs1352015 gave the best result ($P=7.77e^{-7}$). This result is still significant at the 5% level when adjusted for multiple testing using Bonferroni correction, giving a corrected *P*-value of 0.0144, although it does not reach genome-wide significance (genome-wide corrected *P*-value 0.17). The rs1352015 SNP is located 8 kb outside the 5′ end of the *EDA2R* (*EDA-A2* receptor) gene. This region is close to the androgen receptor gene (*AR*), whose intragenic variants (in particular, rs6152) have been associated with the AGA phenotype (Ellis *et al.*, 2001; Hayes *et al.*, 2005; Hillmer *et al.*, 2005; Levy-Nissenbaum *et al.*, 2005). The single *AR* informative intragenic marker included in the Affymetrix 500K array was rs4827545, and it gave a strongly significative *P*-value ($P=6.49e^{-5}$). Because the association close to the *EDA2R* gene was stronger than the one on the *AR* gene, we decided to investigate the possible role of these two genes in the etiology of AGA in our population. In addition, we also tested rs12558842 because its *P*-value was almost as low as that of rs1352015.

To test our results, we selected 492 cases (mean onset age = 24 years, mean hair-loss grade = VI, mean age = 54 years) and 492 controls (mean age = 56 years). This new set included 127 cases and 138 controls already used in the first step of our study. We initially tested the new set for two short tandem repeats (STRs) already reported as being associated with AGA–polyglutamine-encoding CAG repeat and the polyglycine-encoding GGN repeat (Ellis *et al.*, 2001)—and for rs6152 (A/G) as described by La Spada *et al.* (1991) and Hillmer *et al.* (2005). We detected 16 alleles for CAG and 10 for GGN. AGA had not been associated with a particular allele but with groups of alleles on the basis of repeat number. We divided alleles into long and short classes on the basis of previously published findings (Ellis *et al.*, 2001).

Case–control analysis performed on the whole sample (492 cases and 492 controls) showed a very strong association with polymorphism rs6152 ($P=4.17e^{-12}$; G-allele kinship-corrected frequencies cases = 0.92, controls = 0.76), whereas on CAG and GGN repeats, the association with AGA was weaker but still significant ($P=0.01$ and $P=0.0004$, respectively). These results seem to contrast with those recently published by Ellis *et al.* (2007); however, we believe that none of these triplets is actually causative but that the association could be due to the higher linkage disequilibrium (LD) present in our population.

To study *EDA2R*, we sequenced all its exons and the 5′- and 3′-UTR (untranslated region) regions for 20 cases and 20 controls randomly chosen from the whole sample. We found only an informative nsSNP on exon 2: rs1385699 (C/T), which causes the substitution of arginine with lysine on amino acid 57, and four additional informative SNPs in the 5′-UTR (rs4827380, rs12855916, rs11093958, and rs1485682) that were all in complete LD with rs1385699 in all 40 samples. Genotypes for this SNP were obtained using Applied Biosystems TaqMan SNP Genotyping Assays. rs1385699 was revealed to have the strongest association, with a *P*-value of $3.9e^{-19}$ (T-allele kinship-corrected frequencies cases = 0.92, controls = 0.7), odds ratio 4.65 (95% confidence intervals; 3.15–6.87). rs12558842 resulted in strong association with AGA (*P*-value $7.6e^{-14}$); it is, however, in very strong LD with rs1385699. The role of the amino acid in position 57 in *EDA2R* protein is not defined, but it is located in a cysteine-rich domain (Yan *et al.*, 2000). Both amino acids have a polar basic chain, but the N-terminal group on Arg is bigger and is more basic and could influence the protein's activity. Two receptors for *EDA* were found that are specific for the two isoforms *EDA-A1* and *EDA-A2*: *EDAR* and *EDA2R*, respectively. *EDA-A1* and its receptor *EDAR* are capable of activating the *NF-κB* pathway and are implicated in hair growth (Botchkarev and Fessing, 2005). *EDA2R* is capable of activating the *NF-κB* pathway and also through *TRAF3,6*, *JNK* (c-Jun N-terminal kinase) (Sinha *et al.*, 2002), which activates *c-Jun*. Mutations in *EDA* and *EDAR* give rise to ectodermal dysplasia, a clinical syndrome characterized by loss of hair, sweat glands, and teeth, whereas mutations in *EDA2R* do not (Monreal *et al.*, 1999; Naito *et al.*, 2002; Newton *et al.*, 2004). Recently, a preliminary report suggested that *EDAR* may influence hair thickness in Asians (A. Fujimoto, R. Kimura, J. Ohashi, U. Samakkarn,

W. Settheetham-Ishida, T. Ishida, Y. Morishita, T. Furusawa, M. Nakazawa, R. Ohtsuka, R. Yuliwulandari, L. Batubara, M.S. Mustofa, K. Tokunaga, A scan for genetic determinants of human hair morphology: *EDAR* is associated with Asian hair thickness, ASHJ Meeting 2007). *EDA2R* could influence the onset of AGA through the activation of the *NF-κB* pathway or by *c-Jun*, which has been shown to be critical for *AR* transactivation (Bubulya *et al*., 1996). Moreover, in adult mice, *EDA2R* is also expressed in the hair bulb and in differentiating hair matrix (Botchkarev and Fessing, 2005). Looking at the human expression data from the UniGene database (http://www.ncbi.nlm.nih.gov/sites/entrez), we noticed that it is expressed during embryonic life and, especially, in the first weeks after birth. Expression then seems to be absent until the 17th year of age, when it recurs in different tissues, including skin. This expression pattern fits very well with the course of AGA, with its onset around puberty.

Our study shows that *AR* and *EDA2R* are significantly associated with AGA. However, there is some LD between the two most associated markers for each gene (rs6152, rs1385699: $D' = 0.74$, $r^2 = 0.43$). To test if they are independently associated, we conditioned the analysis of each gene to the other one. We used the UNPHASED software (Dudbridge, 2003), which permits the association of a marker to be conditioned to the presence of another marker. The analysis of rs1385699 conditioned to the presence of rs6152 gave a very significant *P*-value of $6.136e^{-9}$, whereas when we conditioned the analysis of rs6152 to the presence of rs1385699 the *P*-value was 0.04. Again, rs1385699 conditioned to the presence of rs12558842 gave a very significant result (*P*-value 0.007), whereas rs12558842 conditioned to the presence of the *EDA2R* variant did not give a significant result (*P*-value 0.06). These results show that in our population, the *EDA2R* gene variation causes susceptibility to AGA. The conditioned analysis suggests that markers on the *AR* gene could be associated because of LD. However, we cannot exclude that other variants in LD with both genes (that is, regulatory elements of either or both genes) could be associated with AGA. Moreover, the functional importance of *AR* has already been proven by many means, and its involvement in this pathology cannot be excluded. Further functional and genetic studies are needed to clarify the role of these two genes and their possible interactions in the etiology of AGA.

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

We express our gratitude to the population of Ogliastra for their collaboration and to the municipal administrations for its logistic support. This work was supported by grants from the Italian Ministry of Education, University and Research (5571/DSPAR/2002) and Regione Autonoma Sardegna (7 del 28/02/2004).

SUPPLEMENTARY MATERIAL

Table S1. Association results of the SNPs in the EDA2R/AR region in 200 cases and 200 controls.

REFERENCES

Angius A, Melis PM, Morelli L, Petretto E, Casu G, Maestrale GB *et al.* (2001) Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits. *Hum Genet* 109: 198–209

Botchkarev VA, Fessing MY (2005) Edar signaling in the control of hair follicle development. *J Investig Dermatol Symp Proc* 10:247–51

Bourgain C, Hoffjan S, Nicolae R, Newman D, Steiner L, Walker K *et al.* (2003) Novel case–control test in a founder population identifies P-selectin as an atopy-susceptibility locus. *Am J Hum Genet* 73:612–26

Bubulya A, Wise SC, Shen XQ, Burmeister LA, Shemshedini L (1996) c-Jun can mediate androgen receptor-induced transactivation. *J Biol Chem* 271:24583–9

Dudbridge F (2003) Pedigree disequilibrium tests for multilocus haplotypes. *Genet Epidemiol* 25:115–21

Ellis JA, Scurrah KJ, Cobb JE, Zaloumis SG, Duncan AE, Harrap SB (2007) Baldness and the androgen receptor: the AR polyglycine repeat polymorphism does not confer susceptibility to androgenetic alopecia. *Hum Genet* 121:451–7

Ellis JA, Stebbing M, Harrap SB (1998) Genetic analysis of male pattern baldness and the 5a-reductase genes. *J Invest Dermatol* 110:849–53

Ellis JA, Stebbing M, Harrap SB (2001) Polymorphism of the androgen receptor gene is associated with male pattern baldness. *J Invest Dermatol* 116:452–5

Falchi M, Forabosco P, Mocci E, Borlino CC, Picciau A, Virdis E *et al.* (2004) A genomewide search using an original pairwise sampling approach for large genealogies identifies a new locus for total and low-density lipoprotein cholesterol in two genetically differentiated isolates of Sardinia. *Am J Hum Genet* 75:1015–31

Fraumene C, Petretto E, Angius A, Pirastu M (2003) Striking differentiation of sub-populations within a genetically homogeneous isolate (Ogliastra) in Sardinia as revealed by mtDNA analysis. *Hum Genet* 114:1–10

Hamilton JB (1951) Patterned loss of hair in man; types and incidence. *Ann N Y Acad Sci* 53:708–28

Hayes VM, Severi G, Eggleton SA, Padilla EJ, Southey MC, Sutherland RL *et al.* (2005) The E211 G>A androgen receptor polymorphism is associated with a decreased risk of metastatic prostate cancer and androgenetic alopecia. *Cancer Epidemiol Biomarkers Prev* 14:993–6

Hillmer AM, Hanneken S, Ritzmann S, Becker T, Freudenberg J, Brockschmidt FF *et al.* (2005) Genetic variation in the human androgen receptor gene is the major determinant of common early-onset androgenetic alopecia. *Am J Hum Genet* 77:140–8

Kuster W, Happle R (1984) The inheritance of baldness: two B or not two B? *J Am Acad Dermatol* 11:921–6

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352:77–9

Levy-Nissenbaum E, Bar-Natan M, Frydman M, Pras E (2005) Confirmation of the association between male pattern baldness and the androgen receptor gene. *Eur J Dermatol* 15:339–40

Monreal AW, Ferguson BM, Headon DJ, Street SL, Overbeek PA, Zonana J (1999) Mutations in the human homologue of mouse dl cause autosomal recessive and dominant hypohidrotic ectodermal dysplasia. *Nat Genet* 22:366–9

Naito A, Yoshida H, Nishioka E, Satoh M, Azuma S, Yamamoto T *et al.* (2002) TRAF6-deficient mice display hypohidrotic ectodermal dysplasia. *Proc Natl Acad Sci USA* 99:8766–71

Newton K, French DM, Yan M, Frantz GD, Dixit VM (2004) Myodegeneration in EDA-A2 transgenic mice is prevented by EDA2R deficiency. *Mol Cell Biol* 24:1608–13

Sinha SK, Zachariah S, Quinones HI, Shindo M, Chaudhary PM (2002) Role of TRAF3 and -6 in the activation of the NF-kappa B and JNK pathways by X-linked ectodermal dysplasia receptor. *J Biol Chem* 277:44953–61

Yan M, Wang LC, Hymowitz SG, Schilbach S, Lee J, Goddard A *et al.* (2000) Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* 290:523–7