



- (51) **International Patent Classification:**
A61K 8/89 (2006.01) *A61Q 5/00* (2006.01)
- (21) **International Application Number:**
PCT/US2013/043519
- (22) **International Filing Date:**
31 May 2013 (31.05.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/654,073 31 May 2012 (31.05.2012) US
- (71) **Applicant:** STIEFEL LABORATORIES, INC. [US/US];
20 TW Alexander Drive, Research Triangle Park, NC
27709 (US).
- (72) **Inventors:** HOFLAND, Hans; 160 Tramore Drive, Chapel
Hill, NC 27516-7749 (US). MOOSER, Vincent, Eugene;
Centre Hospitalier Universitaire Vaudois (CHUV), Rue Du
Bugnon 21, CH-1011 Lausanne, Vaud (CH). SONG, Ki-
joun; 709 Swedeland Road, King Of Prussia, PA 19406
(US). WATERWORTH, Dawn, Marie; 709 Swedeland
Road, King Of Prussia, PA 19406 (US).
- (74) **Agents:** TOMASKA, Margaret, M. et al.; GLAXOS-
MITHKLINE, Corporate Intellectual Property, Uw2220,
709 Swedeland Road, P.O. Box 1539, King Of Prussia, PA
19406-0939 (US).

(81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title:** METHOD OF TREATING

(57) **Abstract:** The invention relates to methods for treating alopecia, e.g. androgenic alopecia, with an HDAC4 inhibitor, or a pharmaceutically acceptable salt or composition thereof.



WO 2013/181487 A2

METHOD OF TREATING

Field of the Invention

The invention relates to methods for treating alopecia, e.g. androgenic alopecia, with an HDAC4 inhibitor, or a pharmaceutically acceptable salt or composition thereof. The invention further relates to methods of treating alopecia, e.g. androgenic alopecia with an HDAC4 inhibitor, or a pharmaceutically acceptable salt or composition thereof, in combination with one or more of finasteride, minoxidil, dutasteride, spironolactone, latanoprost, or oral contraceptives.

Summary of the Invention

Methods are provided for treating alopecia in a mammal in need thereof comprising administering an HDAC (histone deacetylase) inhibitor, or a pharmaceutically acceptable salt or composition thereof. In certain embodiments, the HDAC inhibitor is a class IIa HDAC inhibitor. In other embodiments, the HDAC inhibitor is an HDAC4 inhibitor. In certain embodiments, the HDAC4 inhibitor is a selective HDAC4 inhibitor.

Detailed Description

HDACs

Histone deacetylases (HDACs) are involved in the deacetylation of histones, transcription factors, and other proteins in eukaryotic cells. HDAC inhibitors cause hyperacetylation of histones, opening chromatin and leading to gene transcription and leading to cell-growth arrest, differentiation, and apoptosis.

There are currently 18 known mammalian HDACs that fall into four classes as determined by sequence similarity. Class I, II, and IV HDACs are zinc-dependent enzymes, whereas class III HDACs (SIRT1-SIRT7) are structurally unrelated. Class I HDACs (HDAC1, 2, 3, and 8) are usually present in the nucleus, while the class II HDACs exhibit shuttling. Class II HDACs are subdivided into class IIa (HDAC4, 5, 7, and 9) and IIb (HDAC6 and 10).

Class IIa enzymes have an N-terminal extension of about 5-600 residues whereas the class IIb enzymes contain two catalytic domains. Class IV comprises only HDAC11, which exhibits sequence similarity to the class I enzymes.

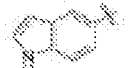
5



Structures 1 and 2

In other embodiments, the HDAC inhibitor is selected from the table below:

15

6i		>1,000	>1,000	540±8.0	>1,000
6j		910±100	54±7.7	160±9.0	15±0.7
6k		>10,000	6,200±130	>10,000	>5,000
6l		>10,000	1,400±250	>10,000	37±1.7
10		3,700±430	190±15	130±10	210±16
2		2,700±300 (5)	780±48 (5)	24±2.6 (5)	230±23 (5)
11		2,200±100	520±85	30±2.9	240±52
12		6,600±43	680±57	60±8.5	700±56
13		3,900±350	1,000±48	35±2.3	520±99

As used herein, "treatment" means any manner in which one or more symptoms associated with the disorder are beneficially altered. Accordingly, the term includes healing or amelioration of a symptom or side effect of the disorder or a decrease in the rate of advancement of the disorder.

As used herein, the term "pharmaceutically acceptable salts" may comprise acid addition salts derived from a nitrogen on a substituent in the compound of formula (I).

Representative salts include the following salts: acetate, benzenesulfonate, benzoate, bicarbonate, bisulfate, bitartrate, borate, bromide, calcium edetate, camsylate, carbonate, chloride, clavulanate, citrate, dihydrochloride, edetate, edisylate, estolate, esylate, fumarate, gluceptate, gluconate, glutamate, glycolylarsanilate, hexylresorcinate, hydrabamine, hydrobromide, hydrochloride, hydroxynaphthoate, iodide, isethionate, lactate, lactobionate, laurate, malate, maleate, mandelate, mesylate, methylbromide, methylnitrate, methylsulfate, monopotassium maleate, mucate, napsylate, nitrate, N-methylglucamine, oxalate, pamoate (embonate), palmitate, pantothenate,

phosphate/diphosphate, polygalacturonate, potassium, salicylate, sodium, stearate, subacetate, succinate, tannate, tartrate, teoclate, tosylate, triethiodide, trimethylammonium and valerate.

- 5 Pharmaceutical formulations may be presented in unit dose forms containing a predetermined amount of active ingredient per unit dose. Such a unit may contain, for example, 0.5 mg to 1g, preferably 1mg to 800mg, of a compound of the formula (I) depending on the condition being treated, the route of administration and the age, weight and condition of the patient. Preferred unit dosage formulations are those containing a
- 10 daily dose or sub-dose, as herein above recited, or an appropriate fraction thereof, of an active ingredient. Furthermore, such pharmaceutical formulations may be prepared by any of the methods well known in the pharmacy art.

- Pharmaceutical formulations may be adapted for administration by any appropriate route,
- 15 for example by the oral (including buccal or sublingual), rectal, nasal, topical (including buccal, sublingual or transdermal), vaginal or parenteral (including subcutaneous, intramuscular, intravenous or intradermal) route. Such formulations may be prepared by any method known in the art of pharmacy, for example by bringing into association the active ingredient with the carrier(s) or excipient(s).

- 20 Pharmaceutical formulations adapted for oral administration may be presented as discrete units such as capsules or tablets; powders or granules; solutions or suspensions in aqueous or non-aqueous liquids; edible foams or whips; or oil-in-water liquid emulsions or water-in-oil liquid emulsions.

- 25 For instance, for oral administration in the form of a tablet or capsule, the active drug component can be combined with an oral, non-toxic pharmaceutically acceptable inert carrier such as ethanol, glycerol, water and the like. Powders are prepared by comminuting the compound to a suitable fine size and mixing with a similarly comminuted
- 30 pharmaceutical carrier such as an edible carbohydrate, as, for example, starch or mannitol. Flavoring, preservative, dispersing and coloring agent can also be present.

Capsules are made by preparing a powder mixture as described above, and filling formed gelatin sheaths. Glidants and lubricants such as colloidal silica, talc, magnesium stearate, calcium stearate or solid polyethylene glycol can be added to the powder mixture before the filling operation. A disintegrating or solubilizing agent such as agar-agar, calcium carbonate or sodium carbonate can also be added to improve the availability of the medicament when the capsule is ingested.

Moreover, when desired or necessary, suitable binders, lubricants, disintegrating agents and coloring agents can also be incorporated into the mixture. Suitable binders include starch, gelatin, natural sugars such as glucose or beta-lactose, corn sweeteners, natural and synthetic gums such as acacia, tragacanth or sodium alginate, carboxymethylcellulose, polyethylene glycol, waxes and the like. Lubricants used in these dosage forms include sodium oleate, sodium stearate, magnesium stearate, sodium benzoate, sodium acetate, sodium chloride and the like. Disintegrators include, without limitation, starch, methyl cellulose, agar, bentonite, xanthan gum and the like. Tablets are formulated, for example, by preparing a powder mixture, granulating or slugging, adding a lubricant and disintegrant and pressing into tablets. A powder mixture is prepared by mixing the compound, suitably comminuted, with a diluent or base as described above, and optionally, with a binder such as carboxymethylcellulose, an aliginat, gelatin, or polyvinyl pyrrolidone, a solution retardant such as paraffin, a resorption accelerator such as a quaternary salt and/or an absorption agent such as bentonite, kaolin or dicalcium phosphate. The powder mixture can be granulated by wetting with a binder such as syrup, starch paste, acadia mucilage or solutions of cellulosic or polymeric materials and forcing through a screen. As an alternative to granulating, the powder mixture can be run through the tablet machine and the result is imperfectly formed slugs broken into granules. The granules can be lubricated to prevent sticking to the tablet forming dies by means of the addition of stearic acid, a stearate salt, talc or mineral oil. The lubricated mixture is then compressed into tablets. The compounds of the present invention can also be combined with a free flowing inert carrier and compressed into tablets directly without going through the granulating or slugging steps. A clear or opaque protective coating consisting of a sealing coat of shellac, a coating of sugar or polymeric material and a polish coating of wax can be provided. Dyestuffs can be added to these coatings to distinguish different unit dosages.

Oral fluids such as solution, syrups and elixirs can be prepared in dosage unit form so that a given quantity contains a predetermined amount of the compound. Syrups can be prepared by dissolving the compound in a suitably flavored aqueous solution, while elixirs are prepared through the use of a non-toxic alcoholic vehicle. Suspensions can be formulated by dispersing the compound in a non-toxic vehicle. Solubilizers and emulsifiers such as ethoxylated isostearyl alcohols and polyoxy ethylene sorbitol ethers, preservatives, flavor additives such as peppermint oil or natural sweeteners or saccharin or other artificial sweeteners, and the like can also be added.

Where appropriate, dosage unit formulations for oral administration can be microencapsulated. The formulation can also be prepared to prolong or sustain the release as for example by coating or embedding particulate material in polymers, wax or the like.

Dosage unit forms can also be in the form of liposome delivery systems, such as small unilamellar vesicles, large unilamellar vesicles and multilamellar vesicles. Liposomes can be formed from a variety of phospholipids, such as cholesterol, stearylamine or phosphatidylcholines.

Pharmaceutical formulations adapted for transdermal administration may be presented as discrete patches intended to remain in intimate contact with the epidermis of the recipient for a prolonged period of time. For example, the active ingredient may be delivered from the patch by iontophoresis as generally described in Pharmaceutical Research, 3(6), 318 (1986).

Pharmaceutical formulations adapted for topical administration may be formulated as ointments, creams, suspensions, lotions, powders, solutions, pastes, gels, sprays, aerosols or oils.

Pharmaceutical formulations adapted for topical administration in the mouth include lozenges, pastilles and mouth washes.

Pharmaceutical formulations adapted for rectal administration may be presented as suppositories or as enemas.

Pharmaceutical formulations adapted for nasal administration wherein the carrier is a solid include a coarse powder having a particle size for example in the range 20 to 500 microns which is administered in the manner in which snuff is taken, i.e., by rapid inhalation
5 through the nasal passage from a container of the powder held close up to the nose. Suitable formulations wherein the carrier is a liquid, for administration as a nasal spray or as nasal drops, include aqueous or oil solutions of the active ingredient.

Pharmaceutical formulations adapted for administration by inhalation include fine particle
10 dusts or mists, which may be generated by means of various types of metered, dose pressurised aerosols, nebulizers or insufflators.

Pharmaceutical formulations adapted for vaginal administration may be presented as pessaries, tampons, creams, gels, pastes, foams or spray formulations.

15 Pharmaceutical formulations adapted for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include
20 suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example sealed ampules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets.

25 It should be understood that in addition to the ingredients particularly mentioned above, the formulations may include other agents conventional in the art having regard to the type of formulation in question, for example those suitable for oral administration may include flavouring agents.

30 A therapeutically effective amount of a compound or a pharmaceutically acceptable salt or solvate thereof will depend upon a number of factors including, for example, the age and weight of the animal, the precise condition requiring treatment and its severity, the nature

of the formulation, and the route of administration, and will ultimately be at the discretion of the attendant physician or veterinarian.

Examples

5 A main advantage of genome-wide association (GWA) studies is that their hypothesis-free scan of the genome for genes associated with common disease enables identification of novel associations between different diseases and pathways. For example, a pioneering GWA study demonstrated that susceptibility alleles for age-related macular degeneration were found in the complement pathway[1], giving rise to novel insights into the etiology
10 and treatment of this condition[2].

Through previous GWA studies, we have identified genetic determinants of androgenetic alopecia (AGA, male-pattern baldness) that highlight the importance of the androgen pathway in this condition[3,4]. However, these studies did not identify readily apparent
15 novel pathways that link AGA to other conditions.

AGA is the most common form of hair loss in humans, affecting 80% of men by age 80[5]. Its etiologic factors are androgen dependency and genetic predisposition[5]. While largely a cosmetic condition, the mechanisms influencing its etiology may also impact
20 upon important medical conditions such as coronary heart disease, metabolic syndrome, and prostate cancer[6,7].

Prostate cancer is the most frequently diagnosed cancer and ranks second as a cancer killer among men in the United States[8]. Androgens play a key role in stimulation of normal
25 prostate growth and are essential in prostate cancer initiation and progression[9]. The main genetic determinant of AGA is the androgen receptor (*AR*)[3,4] and prostate cancer susceptibility loci identified through recent GWA studies overlap with androgen receptor binding sites[10,11] demonstrating a shared etiologic factor in these two conditions. However, while evidence suggests that AGA and prostate cancer are strongly influenced
30 by androgen sensitivity[12,13], attempts to examine the relationship between AGA and prostate cancer through case-control studies have yielded inconsistent results[7] [14-18].

Since previous GWA studies for AGA explained only 13.7% of the variance in this condition[3] we aimed to identify novel determinants of AGA and test their association with common diseases, by undertaking a large-scale meta-analysis of GWA studies involving 12,806 Europeans from seven cohorts included in the Meta-Analysis for Androgenetic Alopecia Novel Determinants (MAAN) consortium. The current analysis comprised 3,891 cases and 8,915 controls of European ancestry from seven independent studies: Bonn, CoLaus, TwinsUK, Nijmegen Biomedical, 23andMe, Icelandic, and an Australian population based twin study (Table 1). Briefly, we used an extreme discordant case-control design to contrast individuals with early-onset AGA to older individuals without alopecia as assessed by questionnaire, clinical visit or photographs evaluated by a dermatologist, where available. Genome-wide genotyping, using standard platforms, and imputation using the CEU panel of Phase II HapMap were performed. After the quality control criteria were applied, 2,391,230 SNPs remained for genotype-phenotype association analysis.

The sample size of the current meta-analysis ($n = 12,806$) was more than fourfold that of the earlier GWA studies of AGA[3,4]. The genomic inflation factors of the individual studies (Table S1) and overall meta-analysis ($\lambda_{gc} = 1.02$) were low, indicating that the observed GWA results were not due to population stratification. The fixed-effect meta-analytic data from all seven cohorts demonstrated a substantial excess of significant associations with AGA at the tail of the QQ plot (Figure S1). We identified a total of 645 SNPs that achieved genome-wide significance ($p < 5 \times 10^{-8}$, Table 2, Table S2). The fixed-effect meta-analytic results showed no evidence of substantial heterogeneity across populations (Tables S1 and S2).

Consistent with our previous reports[3,4], a susceptibility locus for AGA on chromosome 20p11 was confirmed over a ~253 kb interval with the strongest signal arising at rs6047844 ($p = 1.71 \times 10^{-39}$, odds ratio (OR) = 1.60, 95% confidence interval (CI) = 1.49 - 1.72) (Figure 1). We also showed confirmed association with AGA at the *AR* locus through the identification of a highly significant signal spanning the gene region (OR = 2.20, 95% CI = 2.04 - 2.37, $p = 2.40 \times 10^{-91}$ for risk allele T of the top SNP rs2497938). Even after removing SNPs in these two established AGA loci, the QQ-plot demonstrated an excess of SNPs associated with AGA (Figure S2).

FOXA2 (forkhead box A2) is worth noting among the flanking genes at the chromosome 20 locus. Foxa2, the transcription factor encoded by *FOXA2* in mouse, interacts with AR, especially through DNA binding domain, to regulate gene expression[19]. It is expressed in prostate tissue and plays a pivotal role in neuroendocrine prostate tumors, a form of metastatic prostate tumors, by inducing development from androgen-dependent tumors to androgen-independent tumors[20-22]. Its expression was associated with the invasive phenotype in the primary prostate cancer[22]. Furthermore, the association of metastatic prostate cancer and AGA through *AR* polymorphisms was previously described in a well-defined case-control study[14].

Multiple SNPs in strong linkage disequilibrium (LD) ($r^2 > 0.6$) revealed a newly discovered association interval encompassing several genes on chromosome 17q21.31. The most significant signal of this region was detected for the synonymous (His649His) SNP rs12373124 ($P = 5.07 \times 10^{-10}$, OR = 1.33, 95%CI = 1.21 - 1.45 for risk allele T) at the gene *IMP5* (intramembrane protease 5). However, as the genome-wide significant SNPs are highly-correlated with each other in this region, the most proximal gene is not necessarily the gene functionally affected by causal SNPs. The *MAPT* (microtubule-associated protein tau) gene, which encodes the tau protein, is of particular interest in this region because its expression was detected in hair follicles ($p = 9.22 \times 10^{-3}$) but not in the other tissues examined in our tissue expression analysis (Table S3).

The SNP with the smallest p value, apart from the AGA locus on chromosome 20, (rs9287638[A]: OR = 1.31, 95% CI = 1.21 - 1.41, $P = 1.01 \times 10^{-12}$), is located on chromosome 2q37, with seven other genome-wide significant SNPs in the same region. These SNPs lie 558 kb downstream of *HDAC4* (histone deacetylase 4). The next locus mapped within *HDAC9* (histone deacetylase 9) on chromosome 7p21.1 (rs2073963[G]: OR = 1.29, 95% CI = 1.20 - 1.38, $P = 1.08 \times 10^{-12}$). Our expression analysis revealed that *HDAC9* and *HDAC4* were well expressed in hair follicles ($p = 6.59 \times 10^{-3}$ and 2.64×10^{-3} respectively), and *HDAC9* was not expressed in skin or scalp tissues (Table S3).

Thus, two independent signals arose from HDACs (*HDAC4* and *HDAC9*) which act as transcriptional corepressors by deacetylating nucleosomal histones[23]. By interaction with transcription factors ARR19 and CRIF1, HDAC4 plays a critical role in inhibition of

AR transactivation[24,25] and its accumulation coincides with loss of androgen sensitivity in prostate cancer[26]. Furthermore, HDAC9 and HDAC4 share conserved residues and their tissue specific expression pattern overlaps[27]. Very recently, *ZNF652* which has been shown to be involved in transcriptional repression effect of HDACs[28] was identified to be a prostate cancer susceptibility locus[29]. All together, our findings give rise to the possibility that *HDAC4* and *HDAC9* might influence pathogenesis of AGA through dysregulation of the androgen pathway, highlighting a shared etiologic factor in this condition and prostate cancer.

The *AUTS2* (autism susceptibility candidate 2) gene was also identified (lead SNP rs6945541[C]: OR = 1.27, 95%CI = 1.18 - 1.38, $p = 1.71 \times 10^{-9}$) at the chromosome 7q11.22 locus. The mechanism underlying the association between the *AUTS2* locus and AGA is currently unknown, but its expression profile reveals abundance of transcript in hair, skin, and scalp ($p = 2.64 \times 10^{-3}$, $< 1.00 \times 10^{-3}$, and $< 1.00 \times 10^{-3}$ respectively, Table S3), but not blood. Additionally, variants at *AUTS2* were recently associated with the regulation of alcohol consumption through a GWA study[30]. The function of *AUTS2* is currently unknown. It was previously associated with autism and mental retardation[30]. How the *AUTS2* variants affect AGA outcomes deserves further investigation.

Remaining novel findings include one locus on chromosome 1 which is near the genes *TARDBP* (TAR DNA binding protein), *PEX14* (peroxisomal biogenesis factor 14), *MASP2* (mannan-binding lectin serine peptidase 2), and *SRM* (spermidine synthase) (lowest $p = 9.07 \times 10^{-11}$ for SNP rs12565727[A], OR = 1.33, 95%CI = 1.22 - 1.45) and another locus on chromosome 18q21.1, near the 3' end of *SETBP1* (SET binding protein 1), of which the most significant was rs10502861[C] ($P = 2.62 \times 10^{-9}$, OR = 1.28, 95%CI = 1.18 - 1.39). *SETBP1* again demonstrated expression in hair, skin, and scalp (detection p values are 2.64×10^{-3} , $< 1.00 \times 10^{-3}$, and $< 1.00 \times 10^{-3}$ respectively, Table S3), but not blood.

Regional association plots for seven loci associated with AGA are shown in Figure 1. All these loci were associated with AGA in random-effect results except the locus on chromosome 18 (Table S4).

While any of the above susceptibility loci may impart a small risk, examining the combined effect of these loci in individuals harboring more than one risk allele may improve the ability to identify individuals at high risk of AGA. Using the top SNPs identified in this study from both the 6 novel loci and the 2 confirmed regions on chromosome 20p11 and *AR* gene, we constructed a genotype risk score based on the weighted number of susceptibility alleles in an independent replication sample from 23andMe study. As shown in Table 3, there was an increased risk for AGA across each quartile of the genotypic risk score. For the individuals with a genotype score in the highest quartile, we observed a substantially increased odds for AGA (OR = 5.78, 95% CI = 4.86 - 6.87, $p = 1.4 \times 10^{-88}$), compared to individuals at the lowest quartile.

Reasoning that *AR* locus is the most important genetic factor contributing to AGA, we explored the relationship between *AR* locus and other AGA susceptibility loci by several complimentary approaches. First, we examined the effect of *AR* locus on other AGA loci by adding an interaction term in the logistic regression model. We found no evidence for interaction (Table S5), which is consistent with our previous reports[4,31]. Second, we sought clues of interaction between the *AR* locus and other AGA loci mediated by *AR* binding sites since *AR* functions as a transcription factor. Although a large number of genes were found to be targets of *AR*, no *AR* binding sites have been identified in any of the candidate genes at the AGA loci[32,33]. However, as aforementioned, among these candidate genes, interaction with *AR* and consequent regulation on an epididymis-specific gene has been shown on *FOXA2* (chromosome 20 locus), and the candidate gene at chromosome 2 locus, *HDAC4*, was reported to act as an inhibitor of *AR*, especially in prostate cancer cells, suggesting that these two loci may have a role in androgen-dependent pathway.

To discover potential association between AGA and other common diseases, we identified that the 17q21.31 locus overlaps with a widely replicated locus that is strongly associated with Parkinson's disease[34-38]. Interestingly, *MAPT* on this locus contained identical genome-wide significant risk alleles that were shared between AGA (rs2942168 [G], [OR = 1.25, 95% CI = 1.15 - 1.36, $p = 1.95 \times 10^{-7}$]) and Parkinson's disease (OR = 1.32, 95%CI = 1.23-1.39, $p = 1.62 \times 10^{-18}$) reported in a recently published large GWA

study[38]. The lead AGA SNP (rs12373124) at the 17q21.31 locus is in high linkage disequilibrium ($r^2 = 0.87$) with SNPs that have been associated with Parkinson's disease in recent GWA studies[34,35].

To explore this unexpected relationship between Parkinson's disease and early-onset AGA we next sought to understand whether AGA itself was a risk factor for Parkinson's disease. From among a set of the early-onset AGA cases and controls using the same definition for AGA as in the meta-analysis, we identified 568 self-reported, physician diagnosed Parkinson's disease cases, and 7,664 population controls[39]. This definition of Parkinson's disease has been used previously to replicate genetic loci for this condition in the 23andMe cohort[39]. We found that AGA cases had significantly higher odds of Parkinson's disease (OR=1.28, 95% CI = 1.06 - 1.55, $p = 8.9 \times 10^{-3}$). Restricting the analysis to 714 individuals with current age 70 or higher, the association was stronger (OR=1.94, 95% CI = 1.31 - 2.88, $p = 6.5 \times 10^{-4}$).

To investigate whether the shared genetic association at 17q21.31 explained the association between Parkinson's disease and AGA, we evaluated a regression model with an additional term for an individual's genotype at rs12185268, which is the variant in this region most strongly associated with Parkinson's disease in the 23andMe study. The OR for association with early-onset AGA and its significance were essentially unchanged (OR=1.96, 95% CI = 1.32 - 2.90, $p = 6.0 \times 10^{-4}$). In the individuals in the age 70 or higher group, we also tested the association between Parkinson's disease and AGA stratified by rs12185268 genotypes. The AG and GG genotypes were pooled due to the small proportion of GG homozygotes. There was no important difference between the odds ratios for individuals with rs12185268 AA genotypes (OR = 1.93, 95% CI = 1.16 - 3.21, $p = 8.7 \times 10^{-3}$) and AG+GG genotypes (OR = 2.07, 95% CI = 1.09 - 3.93, $p = 2.2 \times 10^{-2}$), indicating that rs12185268 does not modify the association between Parkinson's disease and early-onset AGA.

We next looked for evidence that any other loci associated with Parkinson's disease were also associated with early-onset AGA, or vice versa. We identified the lead SNPs from 27 loci with $p < 1.0 \times 10^{-5}$ for association with AGA, and 31 loci with $p < 1.0 \times 10^{-5}$ for association with Parkinson's disease. We tested the AGA loci for association with

Parkinson's disease, and the Parkinson's disease loci for association with AGA. The 17q21.31 locus was the only locus demonstrating convincing evidence for association across both phenotypes (Tables S6 and S7).

- 5 Parkinson's disease is the second most common neurodegenerative disorder with a prevalence of one percent in individuals that are over 60 years old[40]. Despite the often-reported higher prevalence of Parkinson's disease in men, as compared to women[40], there are no previous reports investigating the relationship between AGA and Parkinson's disease. This novel association between Parkinson's disease and early-onset AGA
10 indicated that there could be a shared genetic or environmental cause for both conditions

- Our data specifically identify genetic variation in the tau protein (encoded by *MAPT*) as shared genetic risk factors for these two conditions. To date, only drug-induced hair loss has been described in patients having Parkinson's disease after use of dopamine
15 agonist[41], although most of the patients affected by drug-induced hair loss are females[41]. As noted above, a greater incidence of Parkinson's disease has been reported in elderly men than in women[40] and androgen mediated neurotoxicity has been proposed to contribute to the gender bias in Parkinson's disease[42]. Since this association is entirely novel, it is unlikely that our results have arisen due to recall bias. At present we
20 are unaware of any prospective Parkinson's disease cohorts having collected AGA data to further explore this relationship, and our evidence provides rationale to undertake such studies.

- Furthermore, the 17q21.31 locus harbors an inversion polymorphism that has previously
25 been described to be under negative selection pressure and has been demonstrated to be associated with decreased fertility in women[43]. We found a genome-wide significant SNP for early-onset AGA, whose risk allele (rs1800547 [A], $p = 2.85 \times 10^{-8}$, OR = 1.27, 95%CI = 1.16-1.39) represents the H1 haplotype of the 17q21.31 inversion. This H1 haplotype is under negative selection pressure in Europeans[44] and
30 Icelandic female carriers of the H1 lineage have fewer children than non-carriers, while men sharing this haplotype have a trend towards decreased fertility[43]. We note that previous studies have identified an association between polycystic ovarian syndrome (PCOS), which is the most common cause of anovulatory infertility in women[45], and

early-onset AGA in their male relatives[46,47]. And increased androgen levels strongly affect both traits. Our findings, therefore, provide rationale to explore the androgen pathway as a possible explanation for the decreased fertility associated with the H1 haplotype in women.

5

Even though the definition of hair loss and the method of sampling differed between the eight study groups, the associations of five genome-wide significant loci were essentially identical in fixed- and random-effects analysis, suggesting that our results are unlikely to be influenced by heterogeneity.

10

We are aware that some other diseases, including progressive supranuclear palsy, corticobasal degeneration, frontotemporal dementia and Pick disease are also strongly associated with the 17q21.31 region[48,49]. However it is difficult to test the association between these diseases and AGA due to their low prevalence[50-52], and the fact that the AGA cases recruited in our study were young.

15

Since 17q21.31 locus has been recognized as a risk factor for Parkinson's disease, it is plausible that an individual's status of Parkinson's disease may have an effect on the association between 17q21.31 locus and early-onset AGA. We undertook an association analysis conditioned on Parkinson's disease in 8,232 individuals for whom both phenotype are available and confirmed that status of Parkinson's disease does not effect the strength of association between AGA and rs12185268, the top variant at 17q21.31 locus in this sample (unconditioned: OR = 0.76, 95% CI = 0.70 - 0.82, $P = 7.9 \times 10^{-13}$; conditioned: OR = 0.76, 95% CI = 0.70 - 0.82, $P = 7.9 \times 10^{-13}$). This result was verified within strata of Parkinson's disease diagnoses: the OR was 0.69 in the Parkinson's disease cases (95% CI = 0.51 - 0.94, $p = 1.8 \times 10^{-2}$) and 0.76 for individuals without Parkinson's disease (95% CI = 0.71 - 0.83, $p = 1.1 \times 10^{-11}$). Further, the relationship between 17q21.31 locus and Parkinson's disease was not affected by conditioning on AGA status although the association on rs12185268 was not significant (unconditioned: OR = 0.99, 95% CI = 0.84 - 1.15, $p = 0.86$; conditioned: OR=1.00, 95% CI = 0.86 - 1.17, $p = 0.97$).

30

The identification of the new associations in this report was driven primarily by augmented power arising from the expanded sample size, which was more than fourfold

that of previous GWA analyses[3,4]. Similar to previous GWA studies[53], this increase in power was associated with a decrease in effect sizes from newly identified loci (top ORs range from 1.27 to 1.33 for six new loci versus OR = 1.60 for the lead SNP on chromosome 20). In this regard, it seems likely that more variants with smaller effect sizes could be uncovered by future larger studies to full describe the allelic architecture of early-onset AGA.

In conclusion, our findings provide fresh insights into the pathogenesis of early-onset AGA. As these newly identified susceptibility loci are also implicated in Parkinson's disease, prostate cancer and fertility, our results highlight the importance of hypothesis-free genetic studies, which allow unexpected genetic relationships between conditions to uncover shared etiologies.

Materials and Methods

15 *Ethics Statement*

All seven studies were approved by institutional ethics review committees at the relevant organizations, and written informed consent was provided by all participating individuals. *Study Participants*

All the participants for this genome-wide meta-analysis were drawn from seven studies: Bonn (582 cases and 347 controls), CoLasu (578 cases and 547 controls), 20 TwinsUK cohort (162 cases and 210 controls), Nijmegen Biomedical Study (73 cases and 132 controls), 23andMe (2,167 cases and 1,753 controls), Iceland (191 cases and 198 controls), Australian population based twin study (138 unrelated cases and 5728 unrelated controls). A detailed description of all these studies and phenotype definitions used in current study is provided in Text S1.

Genotyping and Quality Controls

The genotyping platforms, imputation methods and genome wide association methods used in participant studies are provided in Table S1. Extensive quality control thresholds were applied to include common SNPs (minor allele frequency $\geq 1\%$) with a high call rate ($\geq 95\%$) for genotyped SNPs, and imputed SNPs with high quality metrics (variance ratio ≥ 0.3 for MACH and proper info statistic ≥ 0.4 for IMPUTE)[54,55]. In addition, SNPs demonstrating deviation from Hardy Weinberg Equilibrium ($p > 10^{-6}$) were excluded. The

test statistics for each cohort at each SNP were corrected for their respective genomic inflation factors to avoid inflation of results due to population stratification.

Meta-analysis

We carried out a meta-analysis under both fixed- and random-effects models using the inverse-variance method to combine results from each study using an additive genetic model, while correcting for the genomic inflation factor for each study and the overall meta-analysis. To implement this strategy GWAMA software was used for SNPs on autosomes [56]. Based on the data available, association results using pre-imputation SNPs from Bonn, CoLaus, TwinsUK cohort, Nijmegen Biomedical Study, 23andMe and Australian population based twin study were used for meta-analysis on X chromosome through YAMAS program (<http://yamas.meb.uni-bonn.de/index.html>). Proxy association was applied to gain a higher power. We also tested for evidence of heterogeneity of effects between SNPs and AGA across studies using the Cochran's Q statistics and I^2 measurement. SNPs with low heterogeneity (Q p value > 0.10 and $I^2 < 50\%$) and present in at least three individual studies are reported. In order to test for an inflation of test statistics and the presence of a signal arising from the data for the variants influencing AGA, we constructed quantile-quantile (QQ) plots[57]. Genome-wide significance was set at a p value of 5×10^{-8} [58].

Genome-wide suggestive SNPs ($5 \times 10^{-8} < p < 5 \times 10^{-6}$, $n = 397$) that were imputable using the MetaboChip platform were followed-up in an expanded meta-analysis including the THISEAS study[59] (Text S1, 297 controls and 219 cases) and the original seven cohorts. This additional cohort joined the consortium after completion of the main meta-analysis. No additional loci achieved genome-wide significance after inclusion of this cohort.

Expression Methods

To understand whether the identified SNPs are in close proximity to genes that show differential expression in hair follicles, we performed tissue expression analysis. In brief, total RNA extracted from human hair follicles, skin from temple, scalp, and whole blood were used for array-based gene expression analysis. The differential expression of genes in

these tissues was determined by the average signal of identical probes and detection p values (which is significant, if a gene is reliably expressed). Further details are provided in Text S1.

5 *Genotype Risk Score*

In genotype risk score analysis, the association between a genotype risk score based on the weighted number of susceptibility alleles and AGA status was determined. Six novel susceptibility loci, chromosome 20p11 locus and *AR* locus are included. The weights were established using β coefficients for AGA susceptibility from meta-analysis. The resultant
10 genotypic risk score was divided into quartiles and the risk of AGA for each quartile was tested in an additional set of subjects in 23andMe study, which were not included in the original meta-analysis, using the lowest risk quartile as the reference group. We identified 1582 controls and 1765 cases for this analysis, using the same phenotypic definition used in the meta-analysis. The genotype data is complete without missing values based on
15 phasing and imputation. The trend for risk across the quartiles was tested using the non-parametric trend test[60].

Interaction analysis

To detect the potential modification effect of *AR* locus on any other AGA loci, an
20 interaction analysis was deployed using top SNPs identified from meta-analysis. Each SNP was coded by the number of AGA risk-increasing allele. The test of interaction is based on the coefficient b_3 of the interaction term in the logistic regression model $Y = b_0 + b_1 * A + b_2 * B + b_3 * AB + e$. The Bonferroni correction was applied to address
25 the multiple comparisons problem and the significance level was set to $p \text{ value} < 7 \times 10^{-3}$.

Association with Other Common Diseases

In order to better understand whether the identified risk loci from the meta-analysis were
30 associated with common diseases, we searched the AGA susceptibility loci using the GWAS Integrator embedded in HuGE Navigator website[61]. This online reference tool collects data from all published genetic studies in humans to facilitate the identification of

shared etiologic pathways between diseases.

To examine the relationship between Parkinson's disease and AGA, we performed a logistic regression of Parkinson's disease status against AGA status, evaluating significance of the AGA term by analysis of deviance using a likelihood ratio test. To
5 control for confounding factors, we included age, age², and the top five principal components derived from genotype data as covariates.

References

1. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, et al. (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* 308: 385-389.
- 5 2. Donoso LA, Vrabec T, Kuivaniemi H (2010) The role of complement Factor H in age-related macular degeneration: a review. *Surv Ophthalmol* 55: 227-246.
3. Richards JB, Yuan X, Geller F, Waterworth D, Bataille V, et al. (2008) Male-pattern baldness susceptibility locus at 20p11. *Nat Genet* 40: 1282-1284.
- 10 4. Hillmer AM, Brockschmidt FF, Hanneken S, Eigelshoven S, Steffens M, et al. (2008) Susceptibility variants for male-pattern baldness on chromosome 20p11. *Nat Genet* 40: 1279-1281.
5. Hamilton JB (1951) Patterned loss of hair in man; types and incidence. *Ann N Y Acad Sci* 53: 708-728.
- 15 6. Lo JC, Feigenbaum SL, Yang J, Pressman AR, Selby JV, et al. (2006) Epidemiology and adverse cardiovascular risk profile of diagnosed polycystic ovary syndrome. *J Clin Endocrinol Metab* 91: 1357-1363.
7. Giles GG, Severi G, Sinclair R, English DR, McCredie MR, et al. (2002) Androgenetic alopecia and prostate cancer: findings from an Australian case-control study. *Cancer Epidemiol Biomarkers Prev* 11: 549- 553.
- 20 8. Jemal A, Tiwari RC, Murray T, Ghafoor A, Samuels A, et al. (2004) Cancer statistics, 2004. *CA Cancer J Clin* 54: 8-29.
9. Heinlein CA, Chang C (2004) Androgen receptor in prostate cancer. *Endocr Rev* 25: 276-308.
- 25 10. Jia L, Landan G, Pomerantz M, Jaschek R, Herman P, et al. (2009) Functional enhancers at the gene- poor 8q24 cancer-linked locus. *PLoS Genet* 5: e1000597.
11. Lu Y, Zhang Z, Yu H, Zheng SL, Isaacs WB, et al. (2011) Functional annotation of risk loci identified through genome-wide association studies for prostate cancer. *Prostate* 71: 955-963.
- 30 12. Ellis JA, Sinclair R, Harrap SB (2002) Androgenetic alopecia: pathogenesis and potential for therapy. *Expert Rev Mol Med* 4: 1-11.
13. Suzuki H, Ueda T, Ichikawa T, Ito H (2003) Androgen receptor involvement in the progression of prostate cancer. *Endocr Relat Cancer* 10: 209-216.
- 35 14. Hayes VM, Severi G, Eggleton SA, Padilla EJ, Southey MC, 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: 993-996.

15. Hayes VM, Severi G, Padilla EJ, Morris HA, Tilley WD, et al. (2007) 5alpha-Reductase type 2 gene variant associations with prostate cancer risk, circulating hormone levels and androgenetic alopecia. *Int J Cancer* 120: 776-780.
- 5 16. Cremers RG, Aben KK, Vermeulen SH, den Heijer M, van Oort IM, et al. (2010) Androgenic alopecia is not useful as an indicator of men at high risk of prostate cancer. *Eur J Cancer* 46: 3294-3299.
17. Wright JL, Page ST, Lin DW, Stanford JL (2010) Male pattern baldness and prostate cancer risk in a population-based case-control study. *Cancer Epidemiol* 34: 131-135.
- 10 18. Yassa M, Saliou M, De Rycke Y, Hemery C, Henni M, et al. (2011) Male pattern baldness and the risk of prostate cancer. *Ann Oncol*.
19. Yu X, Gupta A, Wang Y, Suzuki K, Mirosevich J, et al. (2005) Foxa1 and Foxa2 interact with the androgen receptor to regulate prostate and epididymal genes differentially. *Ann N Y Acad Sci* 1061: 77-93.
- 15 20. Mirosevich J, Gao N, Gupta A, Shappell SB, Jove R, et al. (2006) Expression and role of Foxa proteins in prostate cancer. *Prostate* 66: 1013-1028.
21. Qi J, Nakayama K, Cardiff RD, Borowsky AD, Kaul K, et al. (2010) Siah2-dependent concerted activity of HIF and FoxA2 regulates formation of neuroendocrine phenotype and neuroendocrine prostate tumors. *Cancer Cell* 18: 23-38.
- 20 22. Yu X, Wang Y, Degraff DJ, Wills ML, Matusik RJ (2011) Wnt/beta-Catenin activation promotes prostate tumor progression in a mouse model. *Oncogene* 30: 1868-1879.
23. Martin M, Kettmann R, Dequiedt F (2007) Class IIa histone deacetylases: regulating the regulators. *Oncogene* 26: 5450-5467.
- 25 24. Jeong BC, Hong CY, Chattopadhyay S, Park JH, Gong EY, et al. (2004) Androgen receptor corepressor-19 kDa (ARR19), a leucine-rich protein that represses the transcriptional activity of androgen receptor through recruitment of histone deacetylase. *Mol Endocrinol* 18: 13-25.
- 30 25. Suh JH, Shong M, Choi HS, Lee K (2008) CR6-interacting factor 1 represses the transactivation of androgen receptor by direct interaction. *Mol Endocrinol* 22: 33-46.
26. Halkidou K, Cook S, Leung HY, Neal DE, Robson CN (2004) Nuclear accumulation of histone deacetylase 4 (HDAC4) coincides with the loss of androgen sensitivity in hormone refractory cancer of the prostate. *European urology* 45: 382-389; author reply 389.
- 35 27. Petrie K, Guidez F, Howell L, Healy L, Waxman S, et al. (2003) The histone deacetylase 9 gene encodes multiple protein isoforms. *J Biol Chem* 278: 16059-16072.
28. Amann JM, Nip J, Strom DK, Lutterbach B, Harada H, et al. (2001) ETO, a target of t(8;21) in acute leukemia, makes distinct contacts with multiple histone deacetylases

and binds mSin3A through its oligomerization domain. *Mol Cell Biol* 21: 6470-6483.

29. Haiman CA, Chen GK, Blot WJ, Strom SS, Berndt SI, et al. (2011) Genome-wide association study of prostate cancer in men of African ancestry identifies a susceptibility locus at 17q21. *Nat Genet* 43: 570-573.
30. Schumann G, Coin LJ, Lourdasamy A, Charoen P, Berger KH, et al. (2011) Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. *Proc Natl Acad Sci U S A* 108: 7119-7124.
31. Herold C, Steffens M, Brockschmidt FF, Baur MP, Becker T (2009) INTERSNP: genome-wide interaction analysis guided by a priori information. *Bioinformatics* 25: 3275-3281.
32. Massie CE, Adryan B, Barbosa-Morais NL, Lynch AG, Tran MG, et al. (2007) New androgen receptor genomic targets show an interaction with the ETS1 transcription factor. *EMBO Rep* 8: 871-878.
33. Hu S, Yao G, Guan X, Ni Z, Ma W, et al. (2010) Research resource: Genome-wide mapping of in vivo androgen receptor binding sites in mouse epididymis. *Mol Endocrinol* 24: 2392-2405.
34. Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, et al. (2010) Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet* 74: 97-109.
35. Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, et al. (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet* 41: 1308-1312.
36. Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, et al. (2010) Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat Genet* 42: 781-785.
37. Spencer CC, Plagnol V, Strange A, Gardner M, Paisan-Ruiz C, et al. (2011) Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum Mol Genet* 20: 345-353.
38. Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, et al. (2011) Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* 377: 641-649.
39. Do CB TJ, Dorfman E, Kiefer AK, Drabant EM, et al. (2011) Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genetics*.
40. de Lau LM, Breteler MM (2006) Epidemiology of Parkinson's disease. *Lancet Neurol* 5: 525-535.

41. Miwa H, Kondo T (2003) Hair loss induced by dopamine agonist: case report and review of the literature. *Parkinsonism Relat Disord* 10: 51-52.
42. Cunningham RL, Giuffrida A, Roberts JL (2009) Androgens induce dopaminergic neurotoxicity via caspase-3-dependent activation of protein kinase Cdelta. *Endocrinology* 150: 5539-5548.
43. Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, et al. (2005) A common inversion under selection in Europeans. *Nat Genet* 37: 129-137.
44. Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. *PLoS biology* 4: e72.
45. Nestler JE (2008) Metformin for the treatment of the polycystic ovary syndrome. *N Engl J Med* 358: 47- 54.
46. Carey AH, Waterworth D, Patel K, White D, Little J, et al. (1994) Polycystic ovaries and premature male pattern baldness are associated with one allele of the steroid metabolism gene CYP17. *Hum Mol Genet* 3: 1873-1876.
47. Govind A, Obhrai MS, Clayton RN (1999) Polycystic ovaries are inherited as an autosomal dominant trait: analysis of 29 polycystic ovary syndrome and 10 control families. *J Clin Endocrinol Metab* 84: 38-43.
48. Pastor P, Ezquerro M, Perez JC, Chakraverty S, Norton J, et al. (2004) Novel haplotypes in 17q21 are associated with progressive supranuclear palsy. *Ann Neurol* 56: 249-258.
49. Rademakers R, Cruts M, van Broeckhoven C (2004) The role of tau (MAPT) in frontotemporal dementia and related tauopathies. *Hum Mutat* 24: 277-295.
50. Schrag A, Ben-Shlomo Y, Quinn NP (1999) Prevalence of progressive supranuclear palsy and multiple system atrophy: a cross-sectional study. *Lancet* 354: 1771-1775.
51. Togasaki DM, Tanner CM (2000) Epidemiologic aspects. *Adv Neurol* 82: 53-59.
52. Ratnavalli E, Brayne C, Dawson K, Hodges JR (2002) The prevalence of frontotemporal dementia. *Neurology* 58: 1615-1621.
53. Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, et al. (2010) Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet* 42: 570- 575.
54. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 34: 816-834.
55. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome- wide association studies by imputation of genotypes. *Nat Genet* 39: 906-913.
56. Magi R, Morris AP (2010) GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 11: 288.

57. Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nat Rev Genet* 7: 781-791.
58. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449: 851-861.
- 5 59. (2011) A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* 43: 339-344.
60. Cuzick J (1985) A Wilcoxon-type test for trend. *Stat Med* 4: 87-90.
61. Yu W, Yesupriya A, Wulf A, Hindorff LA, Dowling N, et al. (2011) GWAS Integrator: a bioinformatics tool to explore human genetic associations reported in
10 published genome-wide association studies. *Eur J Hum Genet*.

Description of participant studies and phenotype definition

Bonn: All cases were recruited by dermatologists at the Department of Dermatology, University of Düsseldorf, Germany. Males who were <30 years of age having AGA of
15 grades IV-VII or <40 years having AGA of grades V-VII[1]. Population-based controls were recruited at the University of Essen for the Heinz Nixdorf Recall cohort (Risk Factors, Evaluation of Coronary Calcium and Lifestyle). All the cases and controls were of German descent and their baldness status were assessed by dermatologist and documented by photographs. The current sample size consists of 582 cases and 347 controls.

20 **CoLaus:** Members of this cohort were randomly selected from permanent residents of Lausanne, Switzerland between 2003 and 2006[2,3]. Only individuals with 4 European grandparents were eligible for inclusion in the GWA study. Men were considered to have androgenic alopecia if they experienced Hamilton grade V-VII alopecia between the ages of 35 and 65, while controls were men displaying no hair loss (Hamilton grade I) between
25 ages 45 and 55 or grade II between the ages of 56 and 75. Age of onset was assessed by questionnaire. Participants were recruited and assessed by nurses who had been specifically trained to assess for androgenic alopecia. The current study sample consists of 578 cases and 547 controls.

TwinsUK: The TwinsUK cohort is a population-based sample of Britons, unselected for
30 any disease or trait, which is representative of the UK singleton population[4]. A dermatologist using standardized photographs assessed all subjects. Cases were defined as

Hamilton grade V-VII and controls were those without androgenic alopecia (Hamilton grade I). The current study included 162 cases and 210 controls[5].

Nijmegen Biomedical Study. The details of this study were reported previously[6]. Briefly, this is a population-based survey conducted by the Department of Epidemiology and Biostatistics and the Department of Clinical Chemistry of the Radboud University Nijmegen Medical Center (RUNMC), in which 9,371 individuals participated from a total of 22,500 age and sex stratified, randomly selected inhabitants of Nijmegen. Individuals in the Nijmegen Biomedical Study were invited to participate in a study on gene-environment interactions in multifactorial diseases, such as cancer. All participants are of self-reported European descent. In a second phase of this project, participants completed an additional questionnaire and matched their hair pattern at 20 years, 40 years and present age to a Hamilton grading schema[7]. Affected individuals were those reporting Hamilton grade IV-VII by age 40 and controls were those with Hamilton grade I at age 50 years or older, or Hamilton grade II-III at age 60 years or older. A total of 73 cases and 132 controls were included in the current analyses.

23andMe: Participants in this study are unrelated male individuals of predominantly European ancestry, selected from users of the 23andMe Personal Genome Service[8]. Status of male pattern baldness was measured by self assessment through a web based survey. Individuals with Hamilton grade III or higher, age of onset < 40 were defined as cases, and individuals with Hamilton grade I and age ≥ 30 , or Hamilton grade II and age ≥ 50 were considered as controls. In the current study 2,167 cases and 1,753 controls were included.

Iceland: The hair loss of study subjects and the age of onset were identified from questionnaire data. Individuals with considerable hair loss before the age of 50 years were defined as cases. Individuals older than 50 years who reported no hair loss were defined as controls. The current study includes 191 cases and 198 controls.

Australian population based twin study: Measures of hair loss were obtained in the course of an extensive semi-structured telephone interview with respondent booklet, designed to assess physical, psychological and social manifestations of alcoholism and related disorders, conducted with 6265 twins born 1964-71 from the volunteer based

Australian Twin Registry. All males (45% of the sample) were asked to rate their degree of hair loss, if any, using the Hamilton-Norwood Baldness scale, which was printed in the respondent booklet). This data collection scheme was validated in a study by Ellis et al.[9]. Individuals with hair loss Hamilton-Norwood type III or greater were classified as cases and population based individuals were recruited as controls. The current analyses consist of 138 unrelated cases and 5728 unrelated controls.

THISEAS: The THISEAS study is a case- control study for coronary artery disease. As part of the study, hair status was assessed in male individuals by trained scientists. The pattern of AGA was determined using the Hamilton and Norwood classification. The controls were those with no hair loss or those with Hamilton grade II at the age of 50 years or older. Participants with Hamilton grade III or higher, age of onset < 50 years were defined as cases. 297 controls and 219 cases were included in the analyses of the current study.

Tissue expression analysis

Total RNA from human hair follicles, skin from temple, scalp and whole blood was extracted (RNeasy Micro Kit, Qiagen, Hilden, Germany). Quality and quantity of the RNA was analyzed on a NanoDrop ND-1000 spectrophotometer (PepLab Biotechnologie, Erlangen, Germany). For quality reasons, each RNA sample was additionally checked for degradation via gel electrophoresis in a BioAnalyzer 2100 (Agilent Technologies, Waldbronn, Germany) with RNA 6000 nano lab chips following manufacturer's instructions. Array-based gene expression analysis of individual tissue samples was performed on Illumina's human HT-12v3 Expression BeadChips using standard protocols (Illumina, San Diego, USA). In brief, 50ng of total RNA was reverse transcribed into cRNA and biotin-UTP labelled using the Illumina TotalPrep 96-RNA Amplification Kit (Ambion/Applied Biosystems, Darmstadt, Germany). Labelled cRNA was hybridized on Illumina HT-12v3 Expression BeadChips. The obtained fluorescence expression data were background subtracted and quantile normalized using GenomeStudio software (Illumina, San Diego, USA). Supplementary Table3 lists all genes shown in Figure 1 that neighbour the SNP achieving the most significant p-value at the respective identified risk loci for AGA. The average signal (AVG-signal) of identical probes and the detection p-values (which is significant, if a gene is reliably expressed) are shown for each transcript

tested. If the expression of a respective gene was detected by more than one Illumina probe, the probe with the best expression values based on all samples tested is presented. The same applies, if several probes detected different transcript variants of a respective gene. Non-present fluorescence intensity after background subtraction was marked as “-“

5 and defined as not expressed in the corresponding tissue.

Figure 1. Genome-wide meta-analysis results for AGA in MAAN.

(A) Manhattan plot showing the $-\log_{10}$ p value of SNPs against their chromosomal positions. The genome-wide significant SNPs are green (p value $< 5 \times 10^{-8}$). The points with p value $< 1 \times 10^{-40}$ were truncated; the smallest p value was 2.4×10^{-91} at AR locus.

5 (B-H) Regional association plots for seven loci associated with AGA. In each panel, the lead SNP is denoted in purple with its rs ID and association p value. The color of other SNPs indicates the LD with the lead SNP as red ($0.8 : r^2 : 1$), orange ($0.6 : r^2 < 0.8$), green ($0.4 : r^2 < 0.6$), light blue ($0.2 : r^2 < 0.4$), and dark blue ($r^2 < 0.2$). Estimated recombination rates are in light blue.

10 **Figure 2 S1. Quantile-quantile plot of Meta-Analytic Results.** (Genomic Control was applied to each individual cohort prior to meta-analysis and the overall results).

Figure 3 S2. Quantile-quantile plot of Meta-Analytic Results after removing SNPs of AGA locus on chromosome 20 and SNPs of AR locus on chromosome X. (Genomic Control was applied to each individual cohort prior to meta-analysis and the overall results).

15

Figure 4 S3. Forest Plots of the Lead SNP from Genome-Wide Significant Loci.

Table 1 Demographic properties of the study subjects in participant studies

Study	Country of origin	Descent	Study type	Sample size
Bonn	Germany	German	Case-control (population based)	582 Cases 347 Controls
Colums	Switzerland	European	Cohort (community sample)	578 Cases 347 Controls
Iceland	Iceland	European	Cohort (population based)	191 Cases 198 Controls
Nijmegen	Netherlands	European	Cohort (population based)	73 Cases 132 Controls
TwinsUK	UK	UK	Cohort (population based)	162 Cases 218 Controls
23andMe	America	European	Cross-sectional	2,167 Cases 1,753 Controls
Australian	Australia	European	Cohort (population based)	138 Cases 5728 Controls
THRSEAS	Greece	European	Case-control	219 Cases 287 Controls

Table 2 Summary result for the lead SNP from the genome-wide significant loci

Genetic Variant	Chr.	Position ^a	EA ^b	EAF	OR (95% CI)	p value	Gene ^c	Number of GW Significant SNPs	P
rs12365727	1	10855669	A/G	0.788	1.33 (1.25-1.45)	9.07×10^{-11}	<i>ZAR102P</i>	8	<0.01
rs9287618	2	239339379	A/C	0.562	1.31 (1.21-1.41)	1.01×10^{-13}	<i>BDNF4</i>	9	0.10
rs2071963	7	18844399	G/T	0.530	1.29 (1.20-1.38)	1.08×10^{-13}	<i>BDNF9</i>	30	<0.01
rs6945541	7	68249896	C/T	0.539	1.27 (1.18-1.38)	1.71×10^{-8}	<i>ALTS2</i>	6	<0.01
rs12373124	17	41279099	T/C	0.438	1.33 (1.21-1.45)	5.07×10^{-10}	17q21.31	118	0.24
rs10902861	18	41054146	C/T	0.775	1.28 (1.18-1.39)	2.62×10^{-8}	<i>SEZNP1</i>	16	0.18
rs6447844	20	21985575	T/C	0.460	1.60 (1.49-1.72)	1.71×10^{-10}	<i>PMY1, FOXM1</i>	277	<0.01
rs2497918	X	66479743	T/C	0.830	2.20 (2.04-2.37)	2.40×10^{-16}	<i>AR</i>	181	0.32

Abbreviations: Chr., chromosome; EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval.

^a-Chromosome position located in reference sequence of Genome Build 36.3. ^b-AGA risk increasing allele on forward strand. ^c-Gene harboring the SNP or nearest to the SNP. See Figure S3 for forest plots of these SNPs. See Table S2 for an expanded version of this table that shows all SNPs with p values less than 5×10^{-8} .

Table 3 Genotype score associated with the risk of androgenetic alopecia

Genotype score ^a	OR (95% CI)	p value	p value for Trend across all 4 quartiles
Quantile 1	1.00 (Reference)		
Quantile 2	2.20 (1.87 - 2.59)	6.1×10^{-23}	
Quantile 3	3.50 (2.96 - 4.13)	1.8×10^{-46}	
Quantile 4	5.78 (4.86 - 6.87)	1.4×10^{-88}	2.3×10^{-121}

Abbreviations: OR, odds ratio; CI, confidence interval.

^a-The top SNPs from each genome-wide significant loci as shown in Table 2 are used in risk score analysis.

Table S1 Genotyping, imputation and statistical analysis used in the GWA studies

Study	Genotyping		Imputation method	Genome-wide association analysis	
	platforms/chips	calling algorithm		software ^a	λ_{gc}
Bonn	Illumina 317K, 317Kduo, 550K, 610K	Illumina BeadStudio/Genome Studio	IMPUTE[1]	SNPTEST[1]	1.062
ColLum	Affymetrix GeneChip Human Mapping 500K	BRLMM algorithm	IMPUTE	PLINK[2]	1.024
Iceland	Illumina 317K, 317Kduo, CNV370K	Illumina BeadStudio	IMPUTE	SNPTEST	1.022
Nijmegen	Illumina CNV370K	Illumina BeadStudio	IMPUTE	SNPTEST	1.001
TwinsUK	Illumina 610K	Illuminus	IMPUTE	GWAF[3]	1.016
23andMe	Illumina 550+	Illumina GenomeStudio	MACH	R	1.033
Australian	Illumina 317K, CNV370K, 610K	Illumina BeadStudio/Genome Studio	MACH	PLINK	1.016
THISEAS	MetaboChip	GenoSNP	IMPUTE	PLINK	N/A

^a-Association analysis on X chromosome was undertaken by using Clayton's methods[4] implemented in INTERSNP[5]

^b-The relatedness between siblings were corrected by generalized estimating equations[6] implemented in GWAF

1. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39: 906-913.
2. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559-575.
3. Chen MH, Yang Q (2010) GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics* 26: 580-581.
4. Clayton D (2008) Testing for association on the X chromosome. *Biostatistics* 9: 589-600.
5. Harold C, Steffens M, Brockschmidt FF, Baur MP, Becker T (2009) INTERSNP: genome-wide interaction analysis guided by a priori information. *Bioinformatics* 25: 3275-3281.
6. Zeng SL, Liang KY (1996) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42: 121-130.

Table S2 Genome-Wide Significant SNPs in fixed-effect model ($p < 5 \times 10^{-8}$)

Chr.	Position	SNP	Effect Allele	Non Effect Allele	Odds Ratio (95% CI)	p-value	Q p-value	r ²
1	105,554,53	rs2555566	C	A	1.51 (1.32, 1.42)	1.06E-09	0.41	0.02
1	105,556,08	rs2566737	G	A	0.75 (0.68, 0.82)	9.87E-13	0.43	0
1	105,556,64	rs6355356	G	A	0.77 (0.71, 0.84)	1.26E-09	0.44	0
1	105,603,23	rs11576073	T	C	0.77 (0.71, 0.84)	1.06E-09	0.42	0
1	105,605,03	rs1157607	T	C	0.77 (0.71, 0.84)	1.07E-09	0.44	0
1	105,627,93	rs7342133	G	A	1.29 (1.18, 1.41)	1.26E-09	0.44	0
1	105,644,48	rs9436133	G	C	1.29 (1.18, 1.41)	1.43E-08	0.5	0
1	105,674,16	rs7342133	T	A	1.29 (1.18, 1.42)	1.30E-08	0.51	0
2	239,977,64	rs4075566	T	A	1.24 (1.12, 1.36)	1.67E-09	0.46	0
2	239,989,78	rs4075566	C	A	0.77 (0.71, 0.82)	1.68E-12	0.39	0.1
2	239,992,52	rs9711321	T	C	1.24 (1.18, 1.30)	4.09E-11	0.4	0
2	239,992,52	rs11685481	T	C	1.24 (1.18, 1.30)	6.09E-11	0.4	0
2	239,997,48	rs9737443	T	C	1.24 (1.14, 1.35)	1.30E-08	0.63	0
2	239,999,34	rs9737443	G	A	0.75 (0.71, 0.80)	1.30E-08	0.66	0
2	239,999,59	rs9737443	G	A	1.27 (1.17, 1.38)	1.31E-08	0.66	0
2	239,999,59	rs9737443	T	C	1.24 (1.17, 1.32)	1.31E-08	0.67	0
2	239,999,59	rs12613823	T	C	0.75 (0.71, 0.80)	1.34E-08	0.68	0
2	105,513,63	rs3632256	G	A	1.24 (1.16, 1.32)	2.66E-08	0.1	0.46
2	105,545,76	rs17549965	T	C	1.24 (1.16, 1.36)	3.66E-09	0.14	0.39
2	105,582,58	rs3632257	T	C	1.24 (1.19, 1.32)	6.79E-11	0.15	0.37
2	105,609,18	rs2548313	T	G	0.75 (0.74, 0.80)	6.79E-12	0.16	0.33
2	105,634,33	rs13238389	C	A	0.75 (0.72, 0.80)	1.46E-11	0.22	0.38
2	105,639,33	rs9579485	C	A	0.75 (0.74, 0.80)	1.52E-09	0.36	0.33
2	105,641,49	rs9579485	G	A	1.24 (1.2, 1.30)	1.47E-12	0.34	0.35
2	105,644,88	rs2573943	T	G	0.77 (0.71, 0.80)	1.68E-12	0.43	0
2	105,646,76	rs2573944	G	A	0.75 (0.72, 0.80)	1.68E-12	0.34	0.34
2	105,676,88	rs13238386	T	C	1.27 (1.18, 1.36)	4.06E-11	0.32	0.35
2	105,702,16	rs4441364	G	A	0.8 (0.75, 0.86)	2.19E-09	0.33	0.34
2	105,769,24	rs754873	G	A	1.27 (1.19, 1.36)	1.46E-11	0.32	0.4
2	105,777,64	rs13245266	G	A	0.75 (0.71, 0.80)	1.38E-11	0.32	0.35
2	105,804,22	rs10279371	T	C	0.75 (0.71, 0.80)	1.30E-11	0.39	0.35
2	105,825,16	rs2246897	G	A	1.27 (1.18, 1.36)	1.36E-11	0.31	0.34
2	105,848,39	rs4441367	G	A	0.75 (0.74, 0.80)	1.24E-11	0.28	0.3
2	105,885,48	rs754873	T	C	1.27 (1.18, 1.36)	1.36E-12	0.39	0.39
2	105,907,58	rs4441369	G	A	1.24 (1.15, 1.33)	3.46E-08	0.4	0.69
2	105,933,14	rs9737444	G	C	1.24 (1.15, 1.34)	1.09E-09	0.39	0.66
2	105,948,32	rs9737443	G	A	0.8 (0.75, 0.87)	1.32E-08	0.49	0
2	105,959,13	rs4441369	G	A	0.8 (0.74, 0.87)	1.24E-08	0.49	0
2	105,959,49	rs4441369	T	G	0.81 (0.75, 0.87)	1.30E-08	0.33	0.38
2	105,964,16	rs4441364	G	C	1.25 (1.16, 1.34)	9.56E-09	0.46	0
2	105,969,64	rs7754341	C	A	0.81 (0.75, 0.87)	2.67E-08	0.29	0.38
2	105,987,76	rs2246733	G	C	1.24 (1.15, 1.34)	1.46E-08	0.39	0.38
2	105,999,73	rs10277149	G	A	0.81 (0.75, 0.87)	1.33E-08	0.34	0.36
2	105,928,84	rs10277149	G	C	0.81 (0.75, 0.87)	4.36E-08	0.39	0.64
2	105,931,82	rs10277149	T	C	1.24 (1.15, 1.33)	1.46E-08	0.39	0.64
2	105,932,36	rs13238383	C	A	0.81 (0.75, 0.87)	1.61E-08	0.31	0.37
2	105,939,89	rs10486734	G	A	1.23 (1.14, 1.32)	4.96E-08	0.37	0.38
2	602,497,13	rs9497344	T	C	1.27 (1.18, 1.36)	1.14E-09	0.56	0
2	602,497,59	rs4719865	G	A	0.75 (0.71, 0.80)	1.31E-09	0.64	0
2	602,498,96	rs4719865	T	C	0.75 (0.71, 0.80)	1.31E-09	0.67	0
2	602,498,96	rs4719865	G	A	0.75 (0.71, 0.80)	1.41E-09	0.66	0
2	602,498,96	rs13197228	G	A	0.75 (0.71, 0.80)	9.92E-09	0.69	0
27	412,339,48	rs17742384	T	C	0.75 (0.68, 0.82)	4.07E-08	0.44	0
27	412,344,32	rs13238383	G	C	1.25 (1.17, 1.30)	1.38E-08	0.46	0
27	412,344,84	rs13238383	T	C	0.75 (0.72, 0.80)	2.85E-08	0.46	0
27	412,357,12	rs13238383	T	C	1.25 (1.17, 1.30)	2.66E-08	0.46	0

Table S3 Average Expression Signal for All Genes Shown in Figure 1 across Three Tissue Types and Blood

GENE	Chr.	Blumina PROBE_ID	Hair			Skin			Scalp (temples)			Blood	
			AVG Signal	Detection p value		AVG Signal	Detection p value		AVG Signal	Detection p value		AVG Signal	Detection p value
<i>ANGPTL7</i>	1	BLMN_1813361	362.8	< 10E-03		16.2	9.22E-03		236.9	< 10E-03		-	not significant
<i>C1orf127</i>	1	BLMN_1744889	20.3	9.22E-03		22.5	2.64E-03		10.6	4.87E-02		29.8	not significant
<i>C45Z1</i>	1	BLMN_1655191	285.8	1.32E-03		287.3	< 10E-03		410.6	< 10E-03		122.7	2.64E-03
<i>EXOSC10</i>	1	BLMN_1670796	345.0	1.32E-03		299.6	< 10E-03		346.4	< 10E-03		182.1	< 10E-03
<i>M4SP2</i>	1	BLMN_1709408	-	not significant		-	not significant		-	not significant		-	not significant
<i>MTRR</i>	1	BLMN_1769031	90.7	2.64E-03		277.1	< 10E-03		286.4	< 10E-03		115.0	3.95E-03
<i>PEN34</i>	1	BLMN_1763634	60.9	2.64E-03		33.7	< 10E-03		44.6	< 10E-03		28.9	not significant
<i>SRR</i>	1	BLMN_1661337	180.0	2.64E-03		84.4	< 10E-03		79.9	< 10E-03		220.4	< 10E-03
<i>TARDP</i>	1	BLMN_1677532	4.4	not significant		44.8	< 10E-03		54.9	< 10E-03		70.8	1.32E-02
<i>UBIAD1</i>	1	BLMN_1631872	36.9	6.59E-03		63.21	< 10E-03		390.6	< 10E-03		24.8	not significant
<i>ASB1</i>	2	BLMN_1683096	26.0	1.32E-02		45.8	< 10E-03		60.7	< 10E-03		3.3	not significant
<i>ESPNL</i>	2	BLMN_2115862	-	not significant		-	not significant		-	not significant		-	not significant
<i>HDAC4</i>	2	BLMN_1764396	84.1	2.64E-03		63.5	< 10E-03		102.2	< 10E-03		79.2	1.05E-02
<i>HE56</i>	2	BLMN_1694268	82.4	2.64E-03		19.0	6.59E-03		44.3	< 10E-03		74.2	1.39E-02
<i>KLHL30</i>	2	BLMN_1780663	-	not significant		3.7	not significant		7.8	not significant		-	not significant
<i>MGC16023</i>	2	BLMN_1757361	0.7	not significant		-	not significant		1.1	not significant		-	not significant
<i>PER2</i>	2	BLMN_1738095	93.8	2.64E-03		129.1	< 10E-03		112.4	< 10E-03		50.3	3.43E-02
<i>SCXY</i>	2	BLMN_1722742	1.0	not significant		7.1	not significant		-	not significant		35.5	not significant
<i>TRAF3IP1</i>	2	BLMN_2067032	22.5	2.24E-02		6.7	not significant		12.7	1.19E-02		4.3	not significant
<i>UBE2F</i>	2	BLMN_2164242	287.8	1.32E-03		325.6	< 10E-03		256.3	< 10E-03		384.5	< 10E-03
<i>ACTS2</i>	7	BLMN_1749081	57.3	2.64E-03		331.4	< 10E-03		270.0	< 10E-03		9.5	not significant
<i>FERD3L</i>	7	BLMN_1721475	-	not significant		-	not significant		-	not significant		5.4	not significant

Table S4 Summary result for the top SNPs from the genome-wide significant loci using random effect models.

Genetic Variant	Chr.	Position ^a	EA ^b	EAF	OR (95% CI)	p value	Gene ^c	Number of GW Significant SNPs	Q p value	r ²
rs12365727	1	10935669	A/G	0.789	1.33 (1.22-1.45)	9.07×10^{-13}	ERBBP	8	0.43	< 0.01
rs11683401	2	239362355	T/C	0.447	1.26 (1.18-1.35)	6.02×10^{-11}	MDAC4	9	0.60	< 0.01
rs2073963	7	18844099	G/T	0.530	1.29 (1.20-1.38)	1.27×10^{-13}	MDAC9	4	0.42	< 0.01
rs945541	7	68249896	C/T	0.539	1.27 (1.18-1.38)	1.71×10^{-6}	ACYS2	6	0.97	< 0.01
rs17650991	17	41400344	A/C	0.799	1.29 (1.19-1.41)	6.85×10^{-8}	ITG2131	88	0.45	< 0.01
rs6047644	20	21985575	T/C	0.460	1.60 (1.40-1.72)	1.71×10^{-20}	PRX1, PRX42	193	0.58	< 0.01
rs1511061	X	66232849	A/G	0.822	2.40 (2.21-2.62)	6.82×10^{-66}	AR	177	0.71	< 0.01

Abbreviations: Chr., chromosome; EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval.

^aChromosome position located in reference sequence of Genome Build 36.3. ^bAGA risk increasing allele on forward strand. ^cGene harboring the SNP or nearest to the SNP.

Table S5 Interaction results between top SNPs of *AR* locus and other AGA loci

Chr.	SNP	Interaction OR (95% CI)	Interaction p value	I^2
1	rs12565727	0.99 (0.94-1.05)	0.83	< 0.01
2	rs9287638	1.03 (0.99-1.07)	0.10	< 0.01
7	rs2073963	1.01 (0.98-1.05)	0.42	< 0.01
7	rs6945541*	1.04 (1.00-1.09)	0.07	0.57
17	rs12373124	1.00 (0.92-1.09)	0.97	< 0.01
18	rs10502861	1.01 (0.97-1.04)	0.75	0.38
20	rs6047844	1.02 (0.99-1.05)	0.31	< 0.01

Abbreviations: Chr., chromosome; OR, odds ratio; CI, confidence interval.

* I^2 is greater than 50%, meta-analysis was also performed under random effect model with OR (95% CI) = 1.07 (0.98-1.16), P value = 0.14.

Table S6 Association test results for AGA₁ at Parkinson's disease loci

SNP	chr	position	NEA/ EA	Gt. rate	EAF	HWE p value	p value	Odds Ratio	95% CI	Gene
rs34637584	12	39625469	A/G	1.000	0.998	0.012	0.56	0.769	0.314,1.886	<i>LEK2</i>
rs356219	4	90856624	A/G	0.974	0.370	0.75	0.86	0.993	0.919,1.073	<i>GPRD3-SNCA</i>
rs009415	1	153472258	C/T	0.991	0.995	0.59	0.34	0.766	0.481,1.331	<i>GBA</i>
rs12185268	17	41279463	A/G	0.998	0.237	0.32	9.3e-08	0.782	0.714,0.856	<i>TMPS</i>
rs10313789	3	144242767	G/T	1.000	0.796	0.78	0.39	1.041	0.950,1.140	<i>MCC1</i>
rs6812193	4	77418889	C/T	1.000	0.368	0.93	0.27	0.958	0.887,1.034	<i>FAAH7D</i>
rs823156	1	204031263	A/G	0.999	0.180	0.32	0.44	0.963	0.876,1.059	<i>SLC41A1</i>
rs2599054	2	135257016	A/G	1.000	0.462	0.0032	0.27	0.958	0.887,1.034	<i>TMEM163-ACMED</i>
rs6599389	4	929113	A/G	1.000	0.925	0.56	0.32	0.932	0.810,1.072	<i>TMEM175</i>
rs11868035	17	17655826	A/G	0.999	0.686	0.051	0.23	1.049	0.970,1.134	<i>RAT1-SREBF1</i>
rs10886515	10	121333579	C/T	0.999	0.712	0.60	0.50	0.972	0.897,1.054	<i>TIAL1</i>
rs6475667	9	22598036	A/G	1.000	0.135	0.34	0.72	0.980	0.881,1.091	<i>DAB2IP1</i>
rs660895	6	32685358	A/G	1.000	0.191	0.052	0.74	1.016	0.934,1.117	<i>HLA-DRE3-HLA-DQA1</i>
rs2823357	21	15836776	A/G	0.999	0.620	0.19	0.44	0.970	0.899,1.047	<i>NRIP1-USP25</i>
rs11065598	12	120418046	C/T	0.997	0.095	0.73	0.61	0.968	0.853,1.097	<i>FBXL10</i>
rs9379968	6	27352369	A/G	1.000	0.528	0.94	0.56	0.978	0.904,1.053	<i>PRES16-FXSG83</i>
rs9917256	2	168851281	A/G	0.999	0.875	0.70	0.40	1.049	0.999,1.172	<i>STX39-LASS6</i>
rs10999435	10	72045807	A/G	1.000	0.927	0.23	0.57	0.940	0.893,1.105	<i>PRF1-ADAMTS14</i>
rs4130047	18	38932233	C/T	1.000	0.682	0.89	0.60	0.979	0.905,1.060	<i>RIT2</i>
rs35883	3	29573795	A/G	1.000	0.546	0.95	0.52	0.976	0.907,1.051	<i>RBMS3</i>
rs943914	14	54509818	A/C	0.997	0.708	0.47	0.059	1.080	0.997,1.171	<i>WDRD1</i>
rs2778020	7	48427057	A/G	0.999	0.905	0.94	0.41	0.940	0.857,1.076	<i>ABCA13</i>
rs530586	6	31968316	C/T	0.708	0.309	0.018	0.16	1.072	0.974,1.179	<i>EBMT2</i>
rs17763599	19	2320415	G/T	1.000	0.279	0.38	0.31	1.043	0.962,1.132	<i>SPPL28-TMPRSS9</i>
rs2414739	15	59781426	A/G	1.000	0.269	0.86	0.59	0.977	0.899,1.062	<i>VPS13C</i>
rs624032	10	28998382	C/T	0.998	0.355	0.85	0.13	1.058	0.980,1.143	<i>WAC-BAMBI</i>
rs182549	2	136303224	C/T	1.000	0.578	2.4e-29	0.89	1.006	0.926,1.092	<i>ARCMB</i>
rs6464536	7	142399594	C/T	0.999	0.085	0.92	0.84	1.034	0.888,1.157	<i>KEL-OR9A2</i>
rs2785606	5	39243174	G/T	0.357	0.999	0.93	0.49	0.554	0.099,3.088	<i>FYB</i>
rs17326237	8	18025369	C/T	0.998	0.118	0.57	0.64	1.028	0.916,1.154	<i>ASAH1-NAT7</i>

Abbreviation: chr, chromosome; NEA, non-effect allele; EA, effect allele; Gt. rate, genotyping rate; EAF, effect allele frequency.

Note: significant SNP is in bold

Table S7 Association test results for Parkinson's disease, at AGA loci

SNP	chr	position	NEA/ EA	GL rate	EAF	HWE p value	p value	Odds Ratio	95% CI	Gene
rs918696	X	66242280	C/T	0.999	0.146	0.15	0.54	1.020	0.956,1.085	<i>EDAR-AR</i>
rs6113491	20	22005415	A/C	1.000	0.489	0.037	0.82	1.007	0.952,1.064	<i>PAX1-FOXA2</i>
rs13245206	7	13637784	A/G	1.000	0.606	0.087	0.19	1.039	0.981,1.100	<i>HDAC9</i>
rs287638	2	238359379	A/C	1.000	0.676	0.12	0.31	0.980	0.924,1.040	<i>ASB1-HDAC4</i>
rs2003046	1	10953414	A/C	1.000	0.772	0.13	0.38	1.030	0.963,1.102	<i>Ctcf127-TARDBP</i>
rs956087	7	68233508	C/T	0.999	0.538	0.066	0.33	0.973	0.921,1.029	<i>AUTS2</i>
rs7349332	2	219464627	C/T	1.000	0.136	0.23	0.79	1.011	0.932,1.097	<i>WNT10A</i>
rs2668682	17	41648797	A/G	0.992	0.784	0.14	2.8e-12	1.283	1.195,1.378	<i>KIAA1267-LRRRC37A</i>
rs6668810	12	28317687	C/T	1.000	0.265	0.0010	0.62	1.016	0.954,1.082	<i>SIPN-ITPR2</i>
rs11079281	17	52586786	C/T	1.000	0.539	0.45	0.72	1.010	0.953,1.069	<i>AKAP1-MEF2</i>
rs885664	18	41068134	A/C	0.999	0.726	0.66	0.65	1.015	0.953,1.081	<i>SETBP1-SLC14A2</i>
rs10859017	12	116801217	C/T	1.000	0.280	0.89	0.58	1.019	0.954,1.088	<i>KSR2</i>
rs10935800	3	153137960	A/G	1.000	0.713	0.86	0.50	1.018	0.957,1.082	<i>SUCNR1-MBNL1</i>
rs707918	5	158202803	A/C	1.000	0.403	0.029	0.59	1.016	0.960,1.075	<i>ENF1</i>
rs2708482	12	49369532	C/T	1.000	0.389	0.77	0.32	1.000	0.972,1.029	<i>DUF8</i>
rs12194169	6	9601238	G/T	1.000	0.051	0.034	0.66	1.029	0.906,1.168	<i>OFCC1</i>
rs1907354	10	77882399	A/G	0.998	0.679	0.26	0.31	0.960	0.913,1.029	<i>C10orf11</i>
rs13130633	4	8345096	C/T	0.679	0.226	0.27	0.051	1.073	1.000,1.151	<i>HTRAB</i>
rs801307	1	25331538	A/G	1.000	0.336	0.25	0.37	0.973	0.917,1.033	<i>RINX3-STF2</i>
rs10812586	9	27435104	A/G	0.679	0.301	0.67	0.78	0.991	0.929,1.057	<i>MOBKL20</i>
rs4940285	18	45291571	A/G	0.999	0.281	0.19	0.65	1.015	0.954,1.080	<i>COML2-EMAD2</i>
rs12026324	1	119240233	C/T	1.000	0.709	0.26	0.79	0.992	0.932,1.054	<i>TBX15</i>
rs8053509	16	8760772	G/T	0.999	0.865	0.12	0.98	0.999	0.921,1.084	<i>ARAT</i>
rs2861754	2	69931211	A/G	1.000	0.694	0.74	0.37	1.027	0.968,1.090	<i>ETAA1-C1D</i>
rs762174	22	26195344	A/C	0.998	0.390	0.15	0.49	0.980	0.925,1.038	<i>CBYRA4-MXI</i>
rs733710	8	130612833	A/C	1.000	0.787	0.96	0.14	1.050	0.985,1.120	<i>GSDMC</i>

Abbreviation: chr, chromosome; NEA, non-effect allele; EA, effect allele; GL, rate, genotyping rate; EAF, effect allele frequency.

What is claimed:

1 A method of treating a mammal exhibiting one or more symptoms of hair follicle
disregulation/dysfunction comprising administering one or more histone deacetylase 4
5 (HDAC4) inhibitors.

2 A method of stimulating hair growth in a mammal comprising administering one
or more HDAC4 inhibitors.

3 A method of preventing hair loss in a mammal comprising administering one or
more HDAC4 inhibitors.

10 4 The method of claims 1-3, wherein the mammal has a condition selected from the
following: alopecia, hair loss, or hair thinning.

5 The method of claims 1-4, wherein the condition is alopecia, and wherein the
alopecia is androgenic alopecia.

6 The method of claims 1-5, wherein the mammal is a human.

15 7 The method of claims 1-6, further comprising administering an additional molecule
for treatment of the symptoms of hair follicle disregulation.

8 The method of claim 7, wherein the additional treatment comprises one or more of
an HDAC9 inhibitor, finasteride, dutasteride, minoxidil, dutasteride, spironolactone,
latanoprost, or oral contraceptives.

20 9 Any of the methods above, wherein the treatment comprises a pharmaceutically
acceptable salt or composition thereof, of any inhibitor or molecule of any of the
preceding claims.

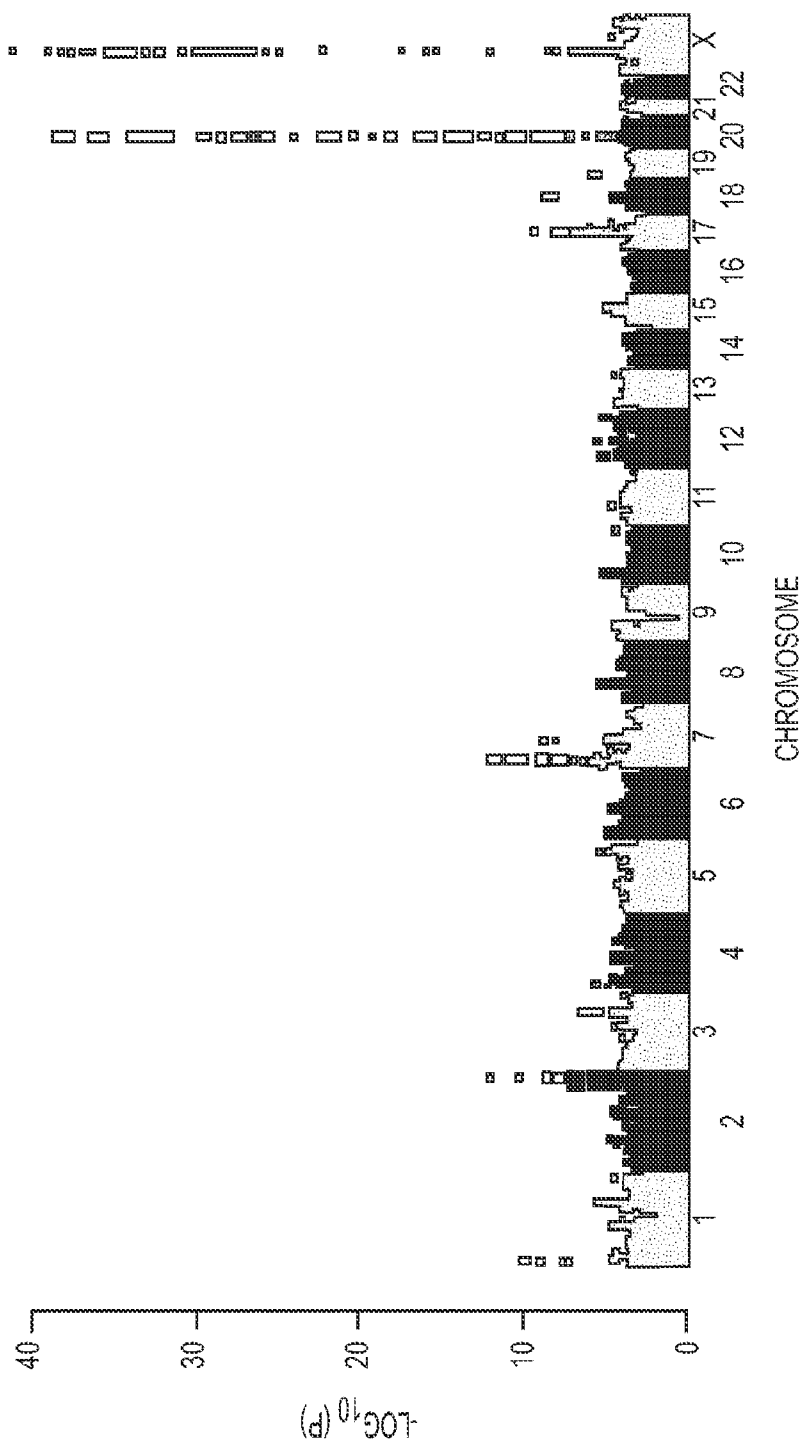


FIG. 1A

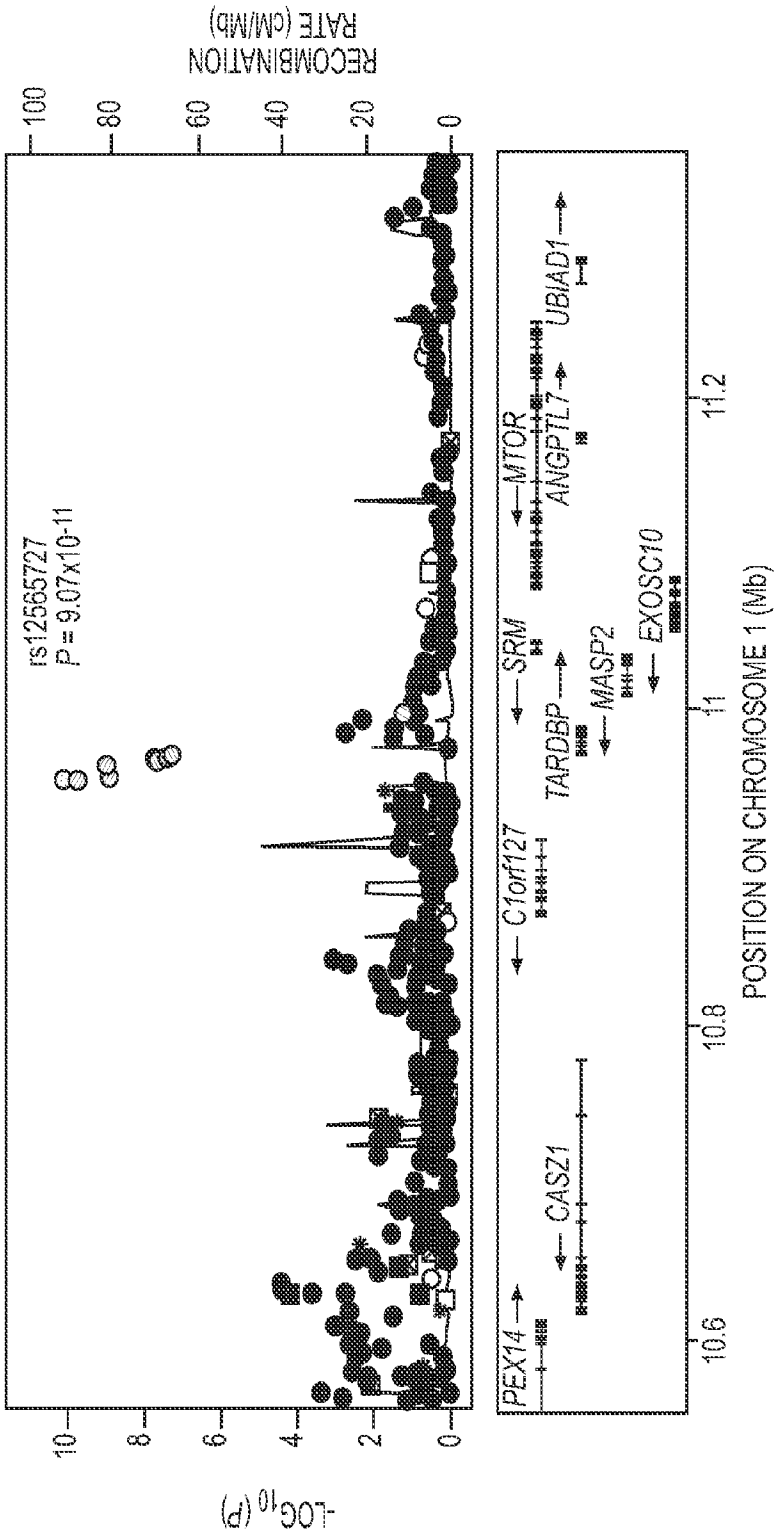


FIG. 1B

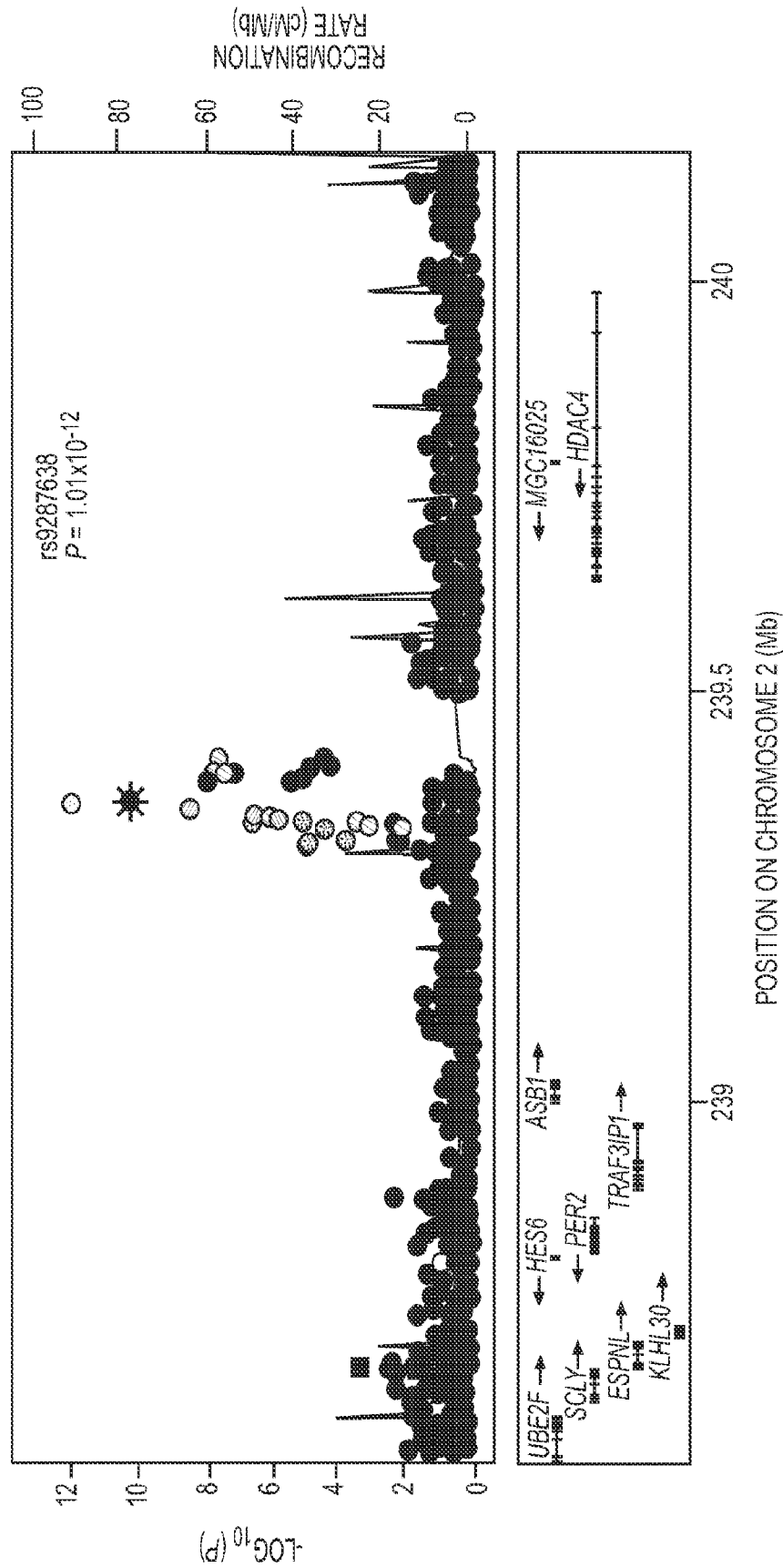


FIG. 1C

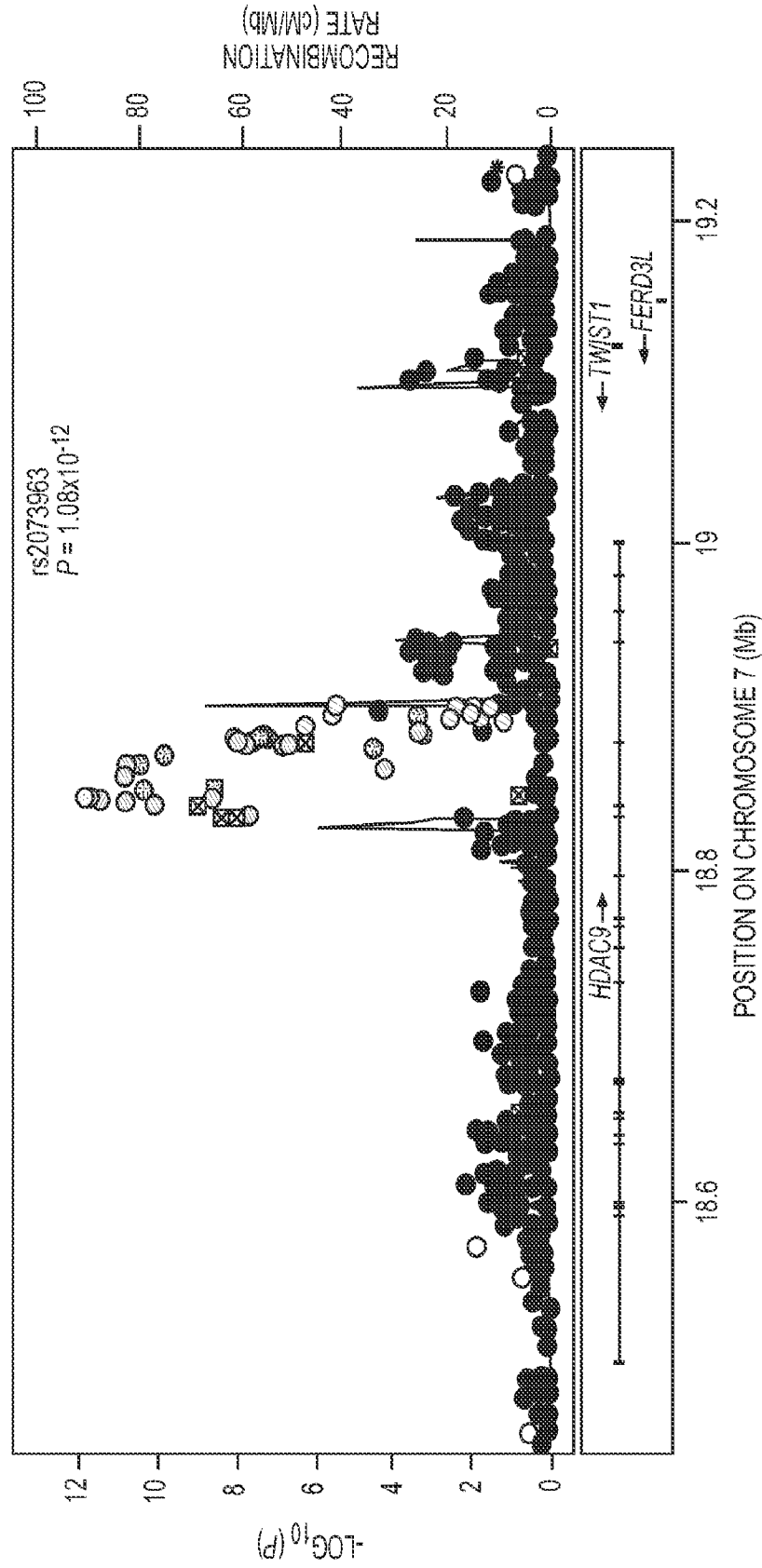


FIG. 1D

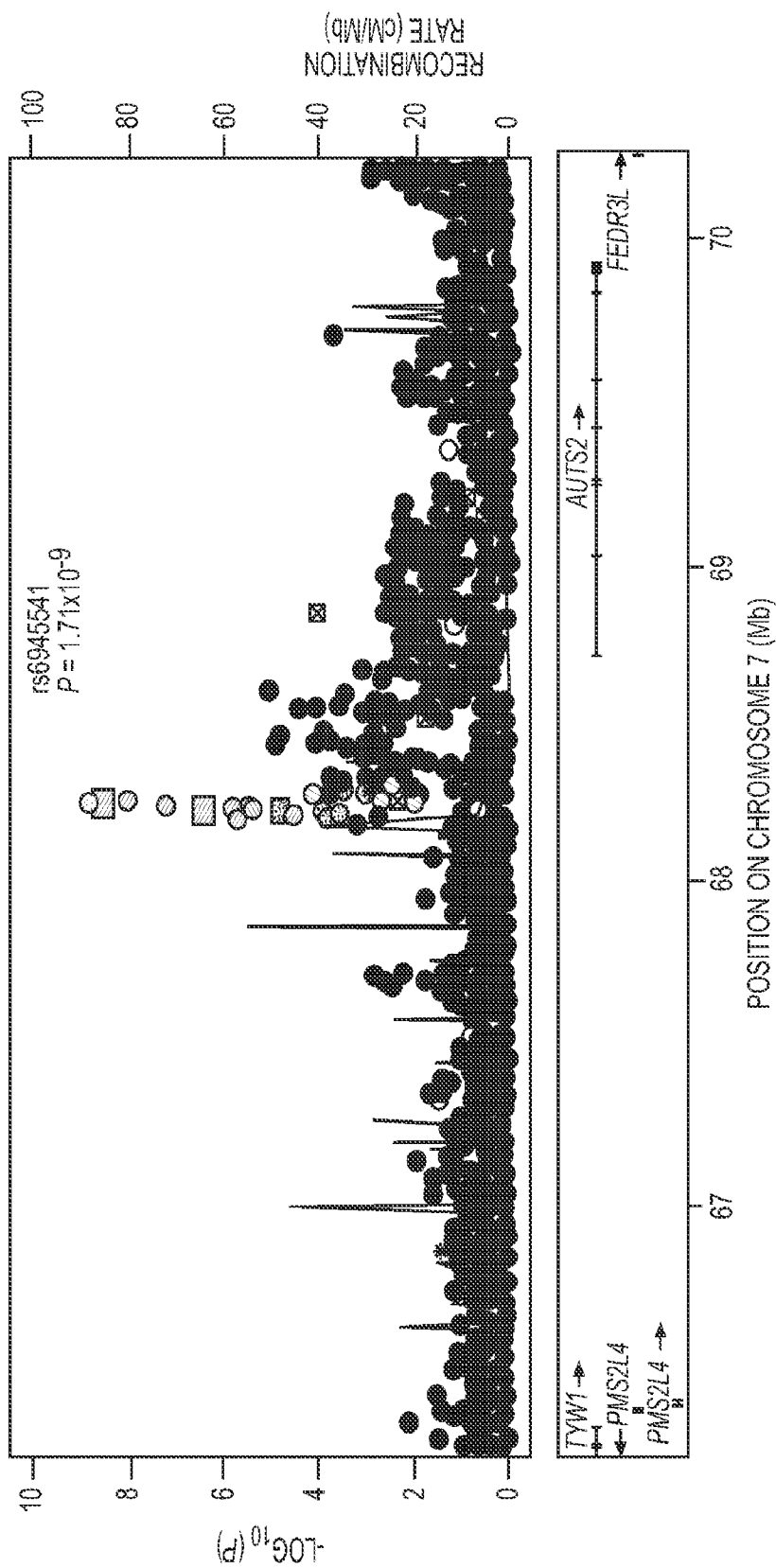


FIG. 1E

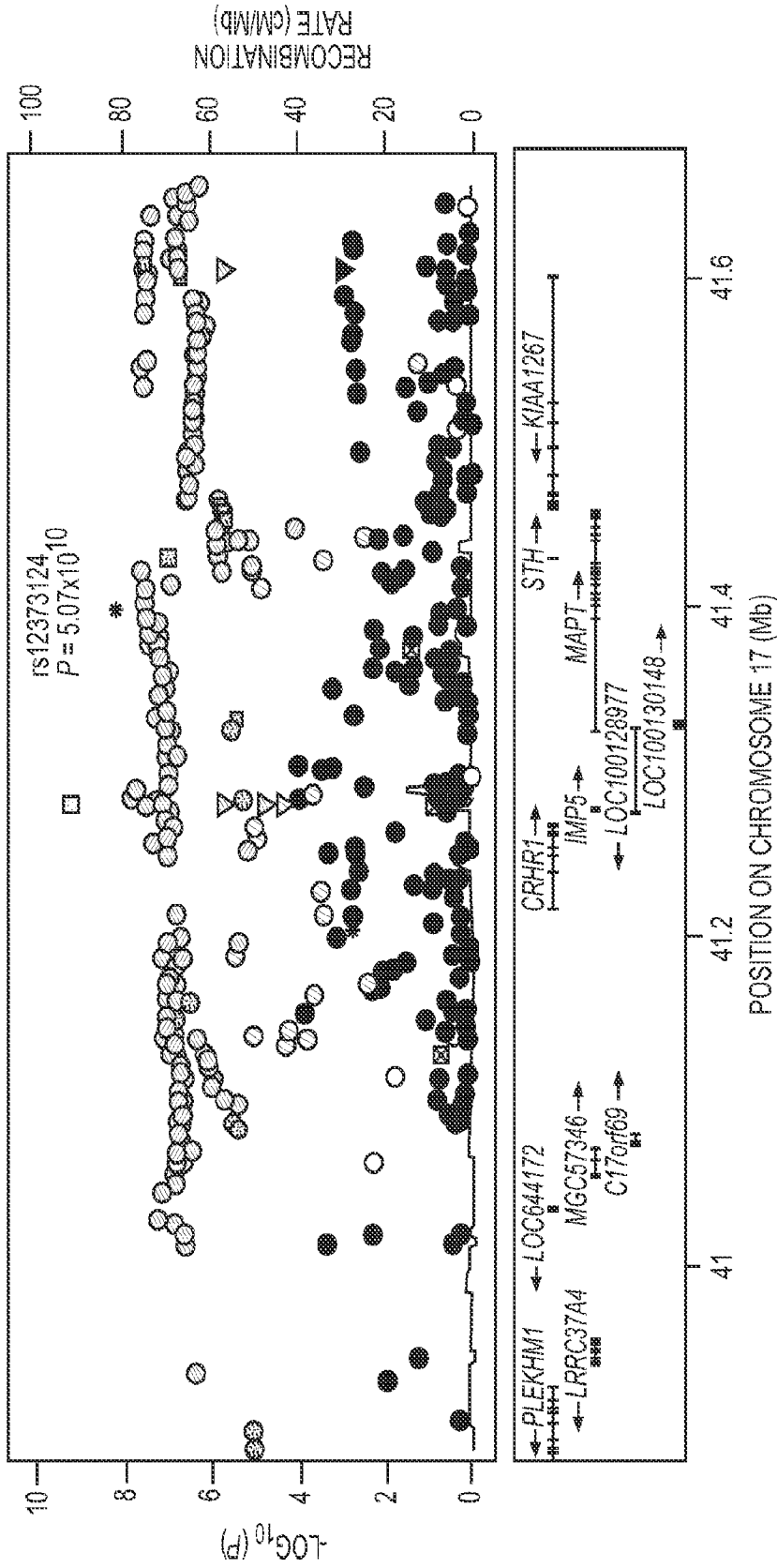


FIG. 1F

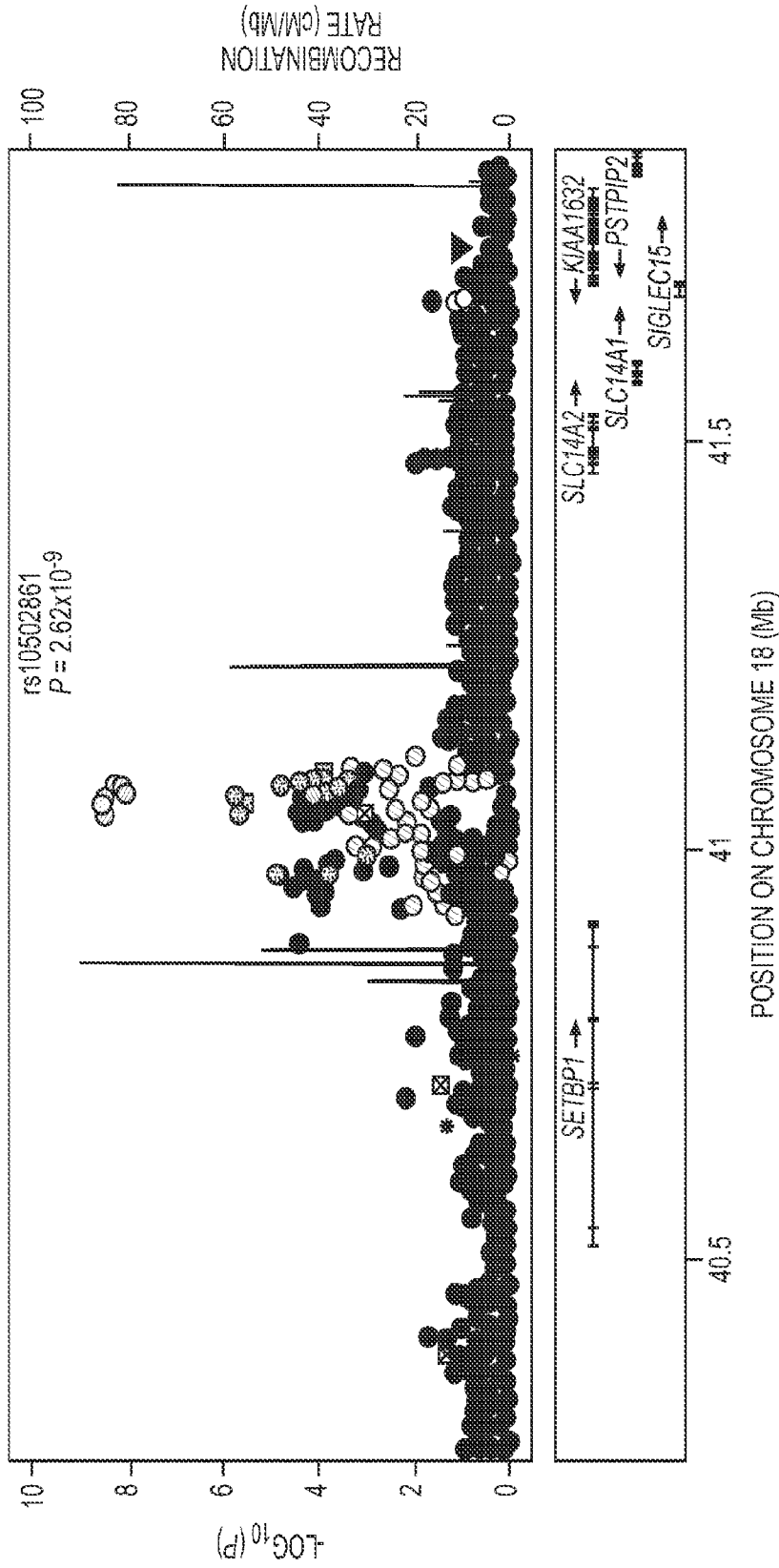


FIG. 1G

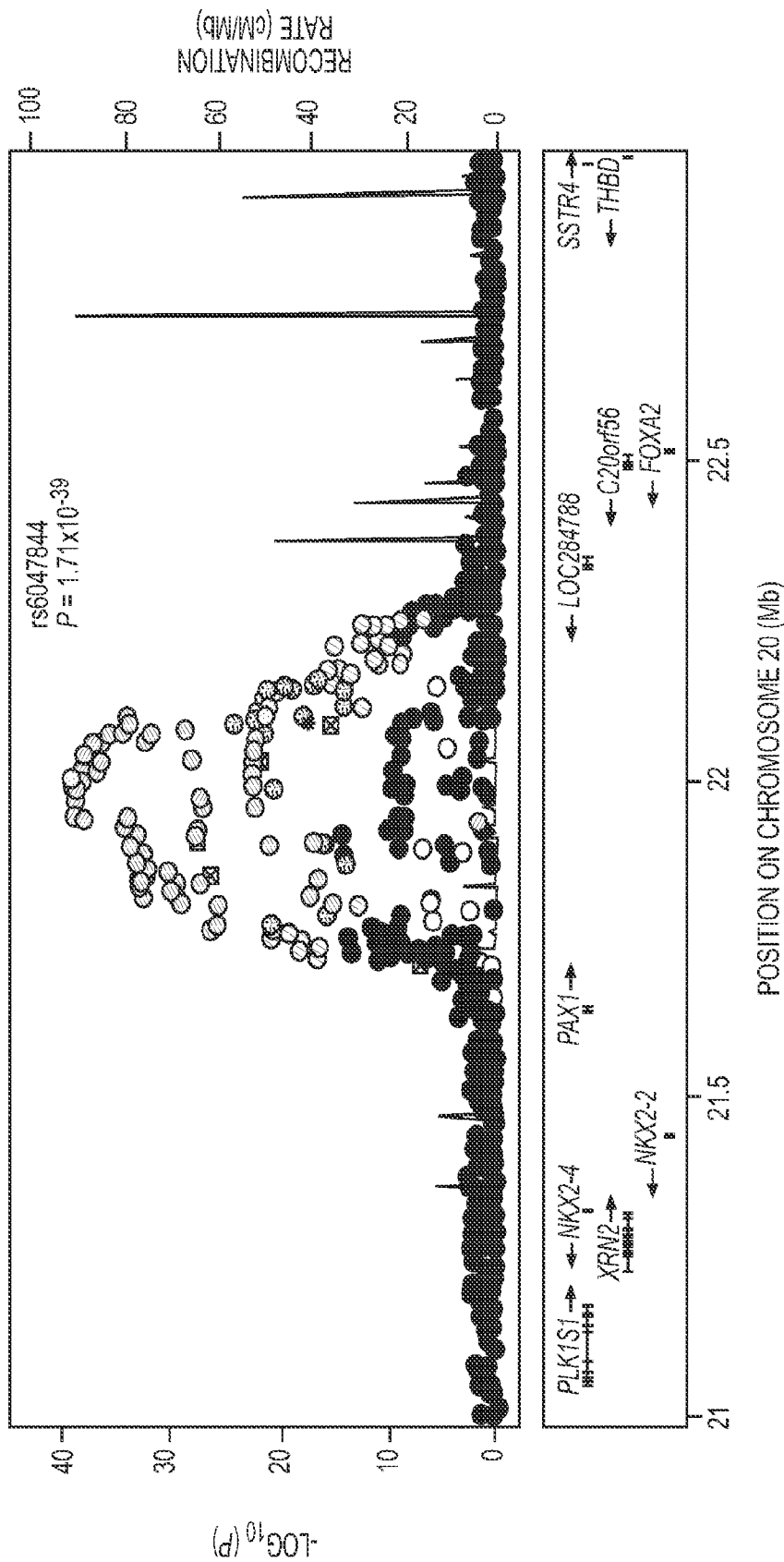


FIG. 1H

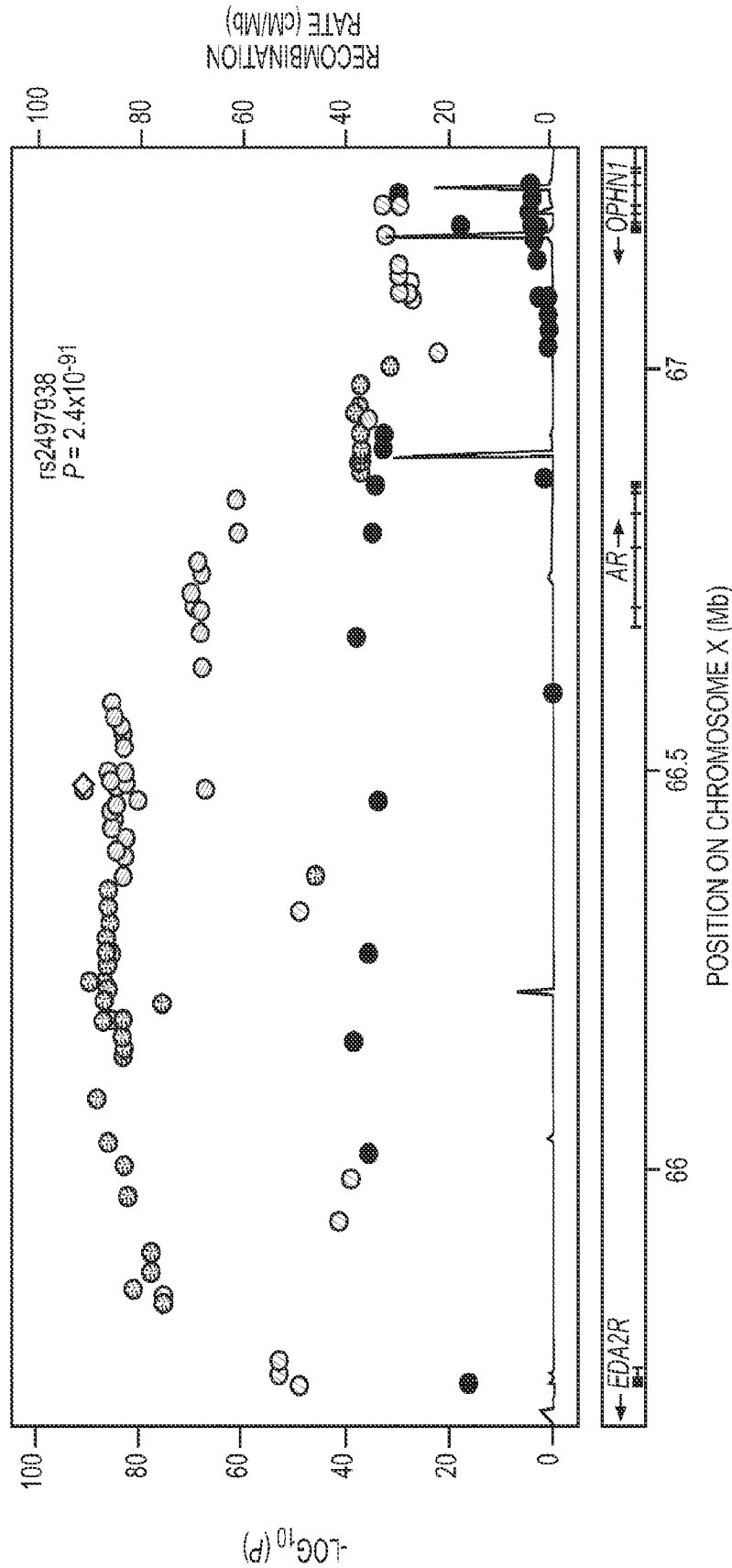


FIG. 11

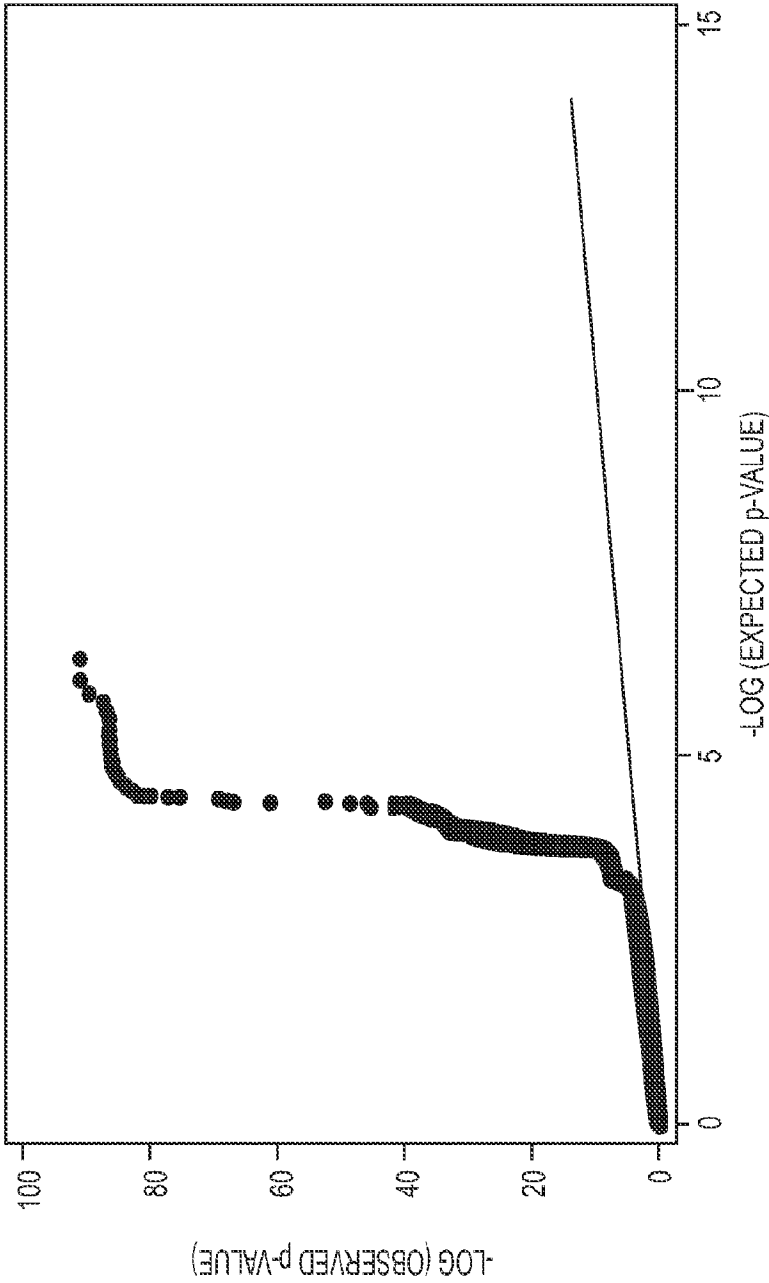


FIG. 2

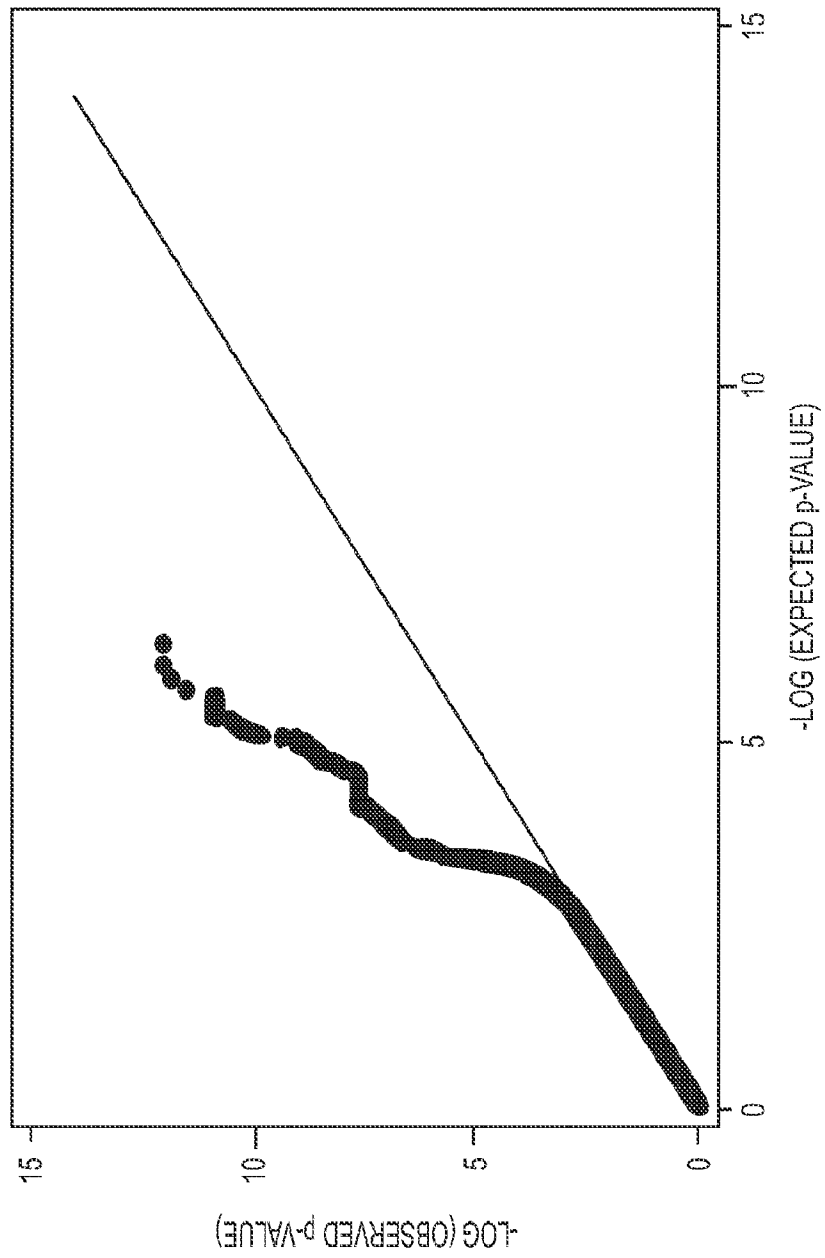


FIG. 3

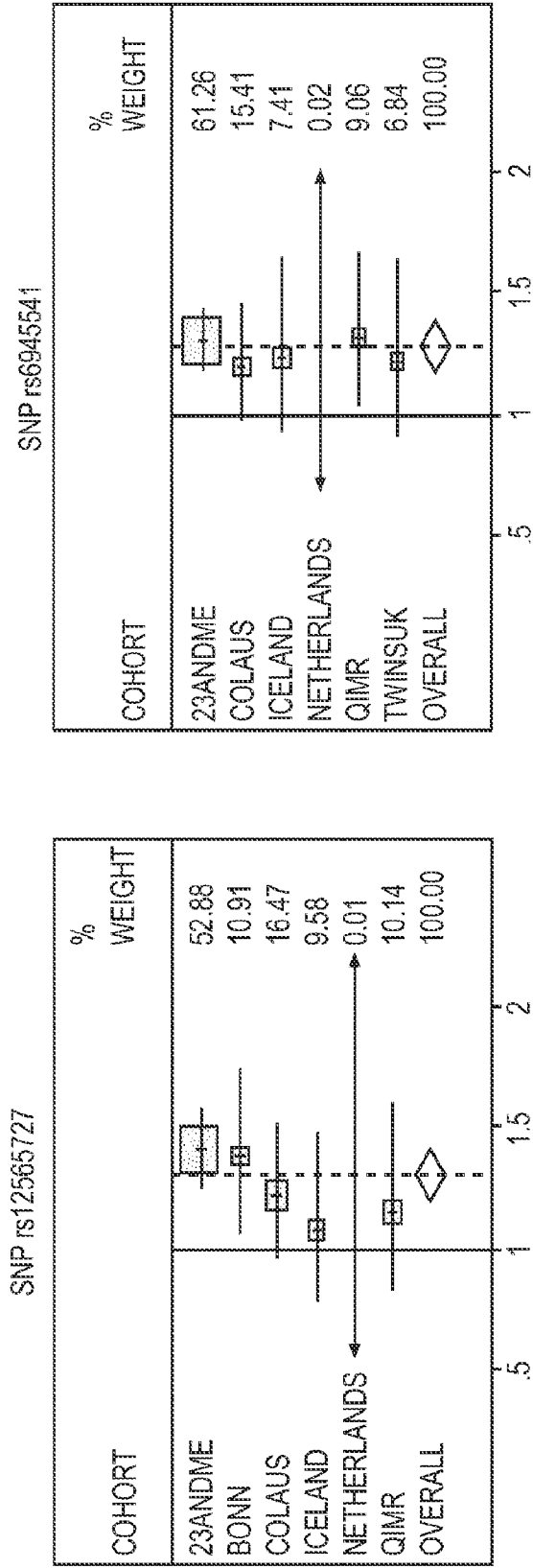


FIG. 4

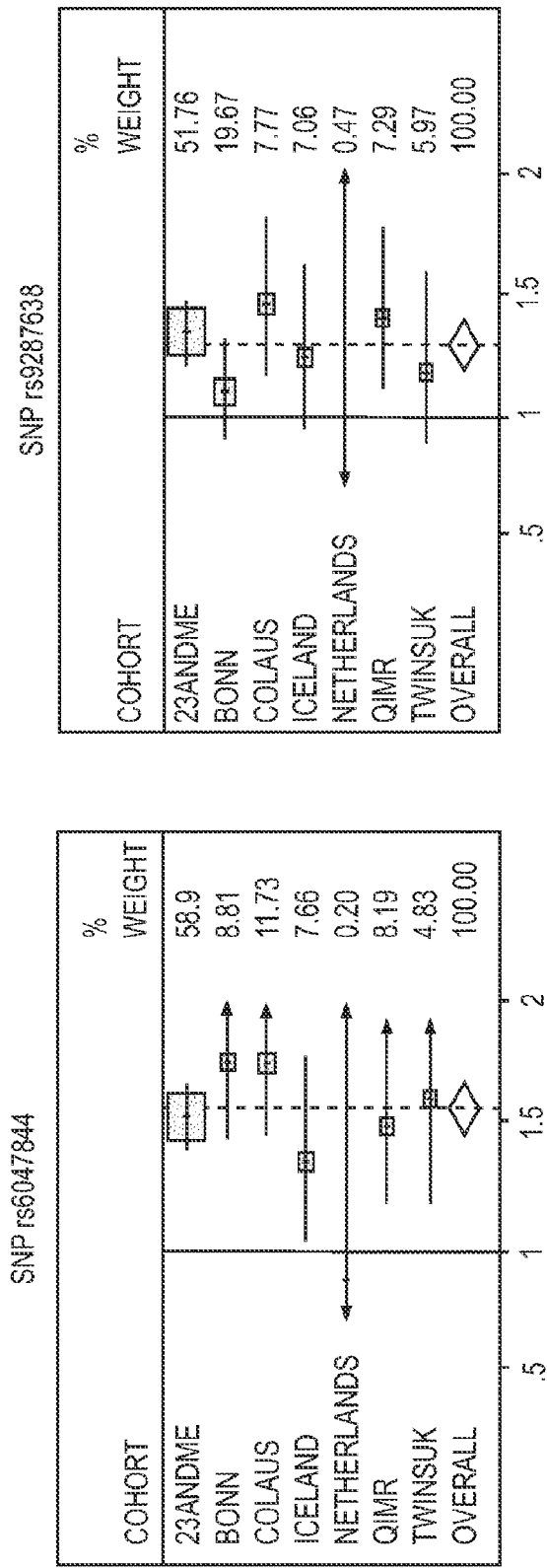


FIG. 4
CONTINUED-1

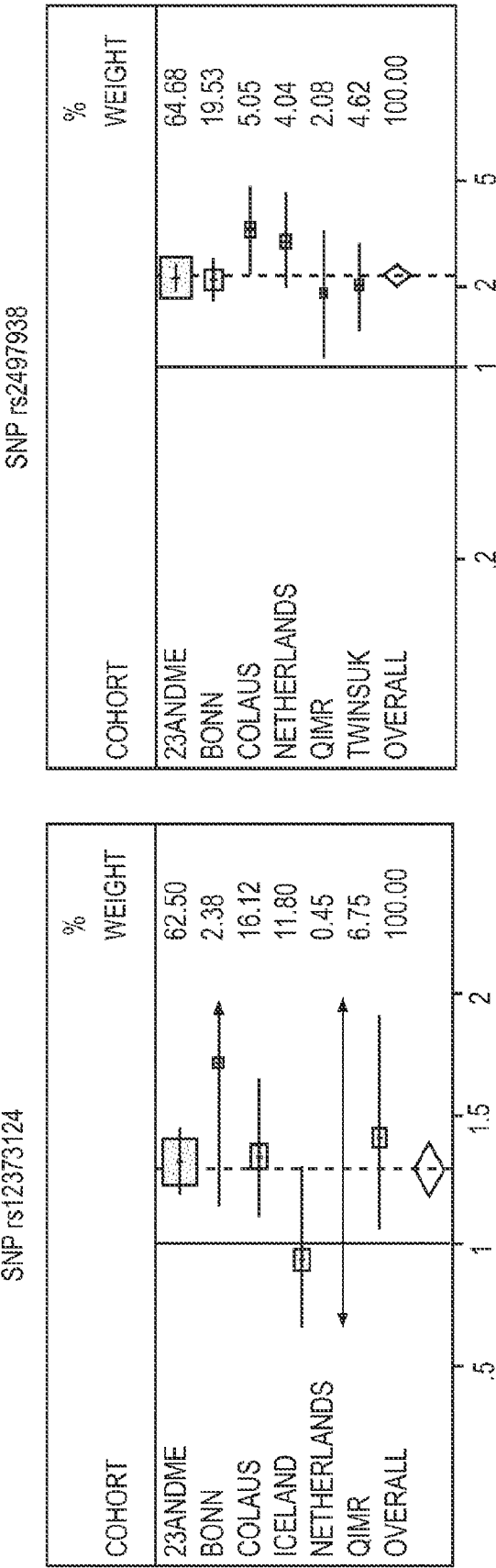


FIG. 4
CONTINUED-2

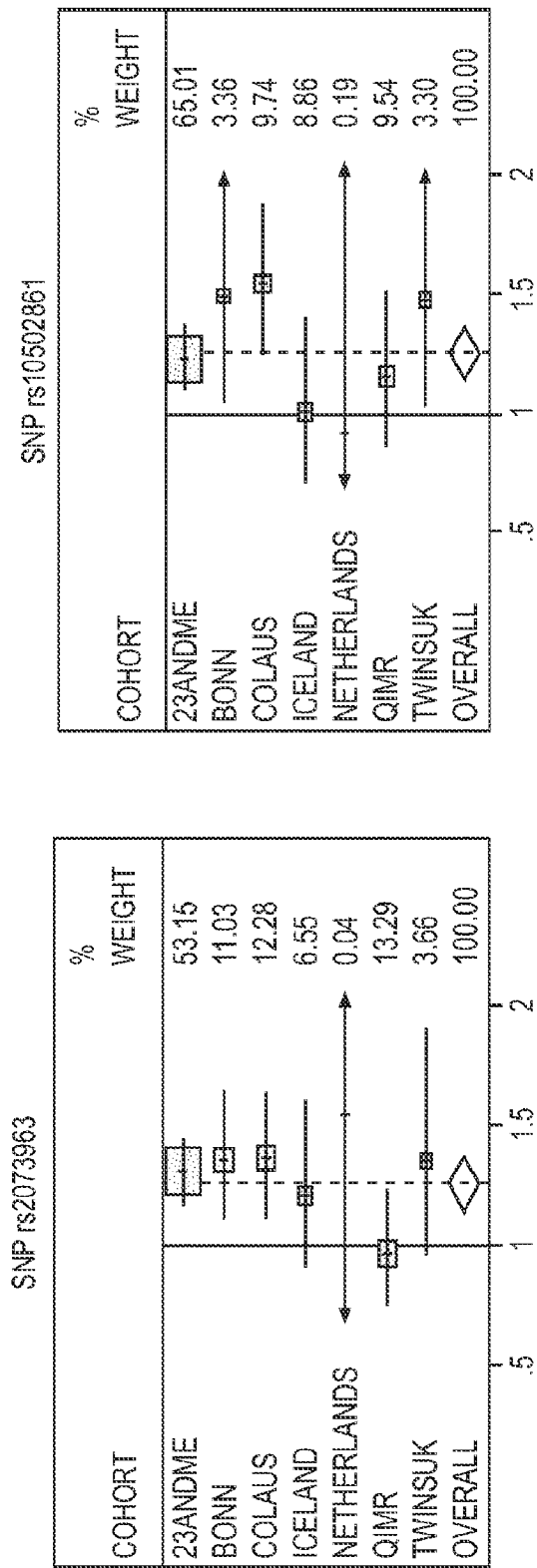


FIG. 4
CONTINUED-3