

RESEARCH ARTICLE

Genetic variations associated with response to dutasteride in the treatment of male subjects with androgenetic alopecia

Arang Rhie^{1,2}✉, Ho-Young Son^{1,3,4}✉, Soo Jung Kwak^{1,3,4}, Seungbok Lee^{1,3,4}, Dong Young Kim^{5,6,7}✉, Bark-Lynn Lew⁸, Woo-Young Sim⁸, Jeong-Sun Seo^{1,3,4}, Ohsang Kwon^{5,6,7}, Jong-Il Kim^{1,3,4}, Seong Jin Jo^{5,6,7}*

1 Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul, Korea, **2** Genome Informatics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States of America, **3** Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea, **4** Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul, Korea, **5** Department of Dermatology, Seoul National University College of Medicine, Seoul, Korea, **6** Laboratory of Cutaneous Aging and Hair Research, Biomedical Research Institute, Seoul National University Hospital, Seoul, Korea, **7** Institute of Human-Environmental Interface Biology, Medical Research Center, Seoul National University College of Medicine, Seoul, Korea, **8** Department of Dermatology, College of Medicine, Kyung Hee University, Seoul, Republic of Korea

✉ These authors contributed equally to this work.

* sj.jo@snu.ac.kr



OPEN ACCESS

Citation: Rhie A, Son H-Y, Kwak SJ, Lee S, Kim DY, Lew B-L, et al. (2019) Genetic variations associated with response to dutasteride in the treatment of male subjects with androgenetic alopecia. PLoS ONE 14(9): e0222533. <https://doi.org/10.1371/journal.pone.0222533>

Editor: Liang Liu, University of Minnesota, UNITED STATES

Received: April 5, 2019

Accepted: August 31, 2019

Published: September 16, 2019

Copyright: © 2019 Rhie et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: According to the Personal Information Protection Act (Clause 17) in Korea, genetic data from Korean individuals is not allowed to be published on public websites unless every single participant provides their consent. The targeted-sequencing data supporting our findings are available from the corresponding author and the Institutional Review Board of Seoul National University Hospital (contact via irb@snuh.org) for researchers who meet the criteria for access to confidential data.

Abstract

Dutasteride, a dual inhibitor of both type I and II 5 α -reductases, is used to treat male pattern hair loss (MPHL). However, patient response to dutasteride varies in each individual, the cause of which is yet to be identified. To identify genetic variants associated with response to dutasteride treatment for MPHL, a total of 42 men with moderate MPHL who had been treated with dutasteride for 6 months were genotyped and analysed by quantitative linear regression, case-control association tests, and Fisher's exact test. The synonymous single nucleotide polymorphism (SNP) rs72623193 in *DHRS9* was most significantly associated with response to dutasteride, followed by the non-synonymous SNP rs2241057 in *CYP26B1*. Additionally, variants in *ESR1*, *SRD5A1*, *CYP19A1*, and *RXRG* are suggested to be associated with response to dutasteride. Cumulative effect and interaction among these SNPs were presented in both additive and non-additive models.

Introduction

Male pattern hair loss (MPHL) is an androgen-dependent progressive hair loss caused by shortening of anagen duration leading to miniaturization of involved hair follicles [1, 2]. About 80% of Caucasian men have MPHL by 70 years of age [3]. MPHL develops in genetically predisposed individuals. In a twin study, the heritability was estimated to be over 80% [4]. A polymorphism in the androgen receptor (AR) gene is associated with MPHL [5–10]. Recently, genome-wide association studies identified several loci associated with MPHL [9–13]. However, MPHL genetics are still not fully understood.

Funding: This study was supported by the SK Telecom Research Fund (Grant No. 34-2014-0120 to OK). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Androgen is the most important factor in the pathophysiology of MPHL. Although the major circulating androgen is testosterone, the androgen responsible for MPHL is dihydrotestosterone (DHT), which is synthesized from testosterone by 5 α -reductase (5AR) in hair follicles and sebaceous glands. Thus, 5AR inhibition is a primary treatment target for MPHL. Varying amounts of hair regrow in MPHL patients treated with finasteride, a selective type II 5AR [14], and dutasteride, a dual inhibitor of both type I and type II 5ARs [15, 16],

However, 5AR inhibition does not always guarantee hair regrowth in the patients with MPHL, and response to medication varies in each individual. In a previous study, about 30% of patients treated with dutasteride for 6 months revealed 'no global change' as determined by investigator photographic assessment and subject self-assessment [15, 16]. The cause of the individual difference in response to 5AR inhibition for MPHL has not yet been identified.

In this study, we classified MPHL patients treated with dutasteride for 6 months into good- and poor-response groups according to the degree of hair regrowth and performed association analysis by targeting 154 genes related to the androgen pathway and drug metabolism.

Materials and methods

Subjects

This study protocol was approved by the Institutional Review Board of Seoul National University Hospital (H-1007-015-322). The target subjects were men with moderate MPHL (IIIv, IV, and V according to the modified Norwood-Hamilton classification scale) who had participated in the previous phase III study to evaluate the efficacy and safety of dutasteride (0.5 mg) [15]. They were treated with dutasteride or a placebo for 6 months. We contacted the 64 subjects who had been assigned to the dutasteride treatment group and had completed the clinical trial, and took peripheral blood samples from them after obtained a written informed consent. Their epidemiological characteristics and hair counts before and after treatment were obtained from the phase III study of dutasteride. To evaluate treatment response, the hair counts were measured at the same circular area marked with ink tattoo before and after 6 months of treatment (S1 Fig); that is, a substantial increase in hair count was considered a good response to dutasteride.

Library construction and targeted sequencing

DNA was extracted from peripheral blood samples of the subjects. Target genes were carefully selected among genes known to be associated with MPHL in previous genetic studies including genome-wide association study and genes related to the steroid hormone biosynthesis (S1 Table) [5–10, 17–19]. In total, 154 genes were targeted (120 for whole gene and 34 for exon only) with the NimbleGen SeqCap EZ Choice Library. We performed sequencing with HiSeq2000 (Illumina, Inc. San Diego, CA, USA).

Variant calling and SNP filtering

We aligned the DNA with the human genome reference sequence (GRCh37) using Bowtie2 [20]. Single nucleotide polymorphisms (SNPs) were called with Genome Analysis Toolkit (GATK, version 2.5) SNPs were filtered as following criteria. (i) SNP candidates were called from a sequencing region with a minimum coverage of 10 \times , with at least 5 \times the appearance of a minor allele. (ii) Among these candidate SNPs, only variants with an average quality score of 20 (Phred scale) were chosen. (iii) Allelic frequency ($100 \times$ Minor allele counts / Total allele counts) of less than 20% were considered as a sequencing error, 20–80% were considered as a heterozygote, and over 80% were considered as a homozygote [21, 22]. We annotated genetic

variants using the software ANNOVAR [23] and called copy number variation using the software CoNIFER [24].

Association test analysis

By setting the phenotype as the change in hair count after the 6-month treatment with dutasteride (0.5 mg), quantitative regression-based association tests between phenotype and genotype were performed for all SNPs, following which quantitative association tests using linear regression were performed for the data generated from all subjects with unadjusted and age adjusted models. Furthermore, we conducted quantitative association tests with good- and poor-response groups as cases and controls, respectively. The case/control association test was conducted twice using 2 different standards for classification: upper half vs. lower half and top quartile vs. bottom quartile response to dutasteride. To enhance confidence in our selected candidates, Fisher's exact tests were performed on the groups as well. All of the association tests were conducted using PLINK [25]. Multiple test corrections were performed using Bonferroni correction, taking into account linkage disequilibrium (LD) blocks with a threshold of $R^2 > 0.8$ using pairwise tagging in HaploView [26].

Gene-based association analysis

Gene-based association tests were conducted using VEGAS [27] 4 times on the statistically significant SNPs, as determined by the 4 prior association tests, with over 10^5 simulations of each gene. To identify a hidden association, i.e. the cumulative effect among exonic SNPs, one exonic variant was selected from each gene as a tagged SNP. Each SNP with a positive effect was rated as "+1" and with a negative effect as "-1". Allele frequencies of good- and poor-response groups were compared that of an Asian population.

Multifactor dimension reduction modelling

To evaluate the non-additive effect, Multifactor Dimension Reduction (MDR) [28], a statistical epistasis analysis method for genetic association, was conducted. We used the 6 exonic variation candidates in Fig 1 to perform a multifactor dimensionality reduction analysis using MDR 3.0.2 (<https://www.multifactor dimensionality reduction.org>). We classified the variations by their haplotypes and considered a poor responder group (class "1") and a good responder group (class "2") as discrete variables. Before the analysis, the attribute count range was configured as 1:6 to determine the number of attributes that fit the best model by considering the balancing accuracy and cross-validation consistency using all 6 SNPs.

Results

In total, 42 out of 64 MPHL subjects enrolled in the previous phase III clinical trial that received dutasteride (0.5 mg) for 6 months agreed to participate and were included in this pharmacogenetic study (S2 Table). To evaluate non-genetic factors associated with response to dutasteride, we compared the age, height, and weight and the duration of MPHL between the 10 highest- and 10 lowest-ranked patients based on their hair count increase after 6 months of treatment; however, we did not find any significant differences among the subjects with respect to these parameters (S3 Table).

The outline of the genetic analysis is shown in S2 Fig. Subjects were sequenced with a customized targeted capture kit for next-generation sequencing. On average, 4.3 Gb were sequenced per sample, and 92.37% of the target region was covered over 30×, with a mean depth of coverage 296.3× (S4 Table). Among the 9524 SNPs called, 3430 were left in the target

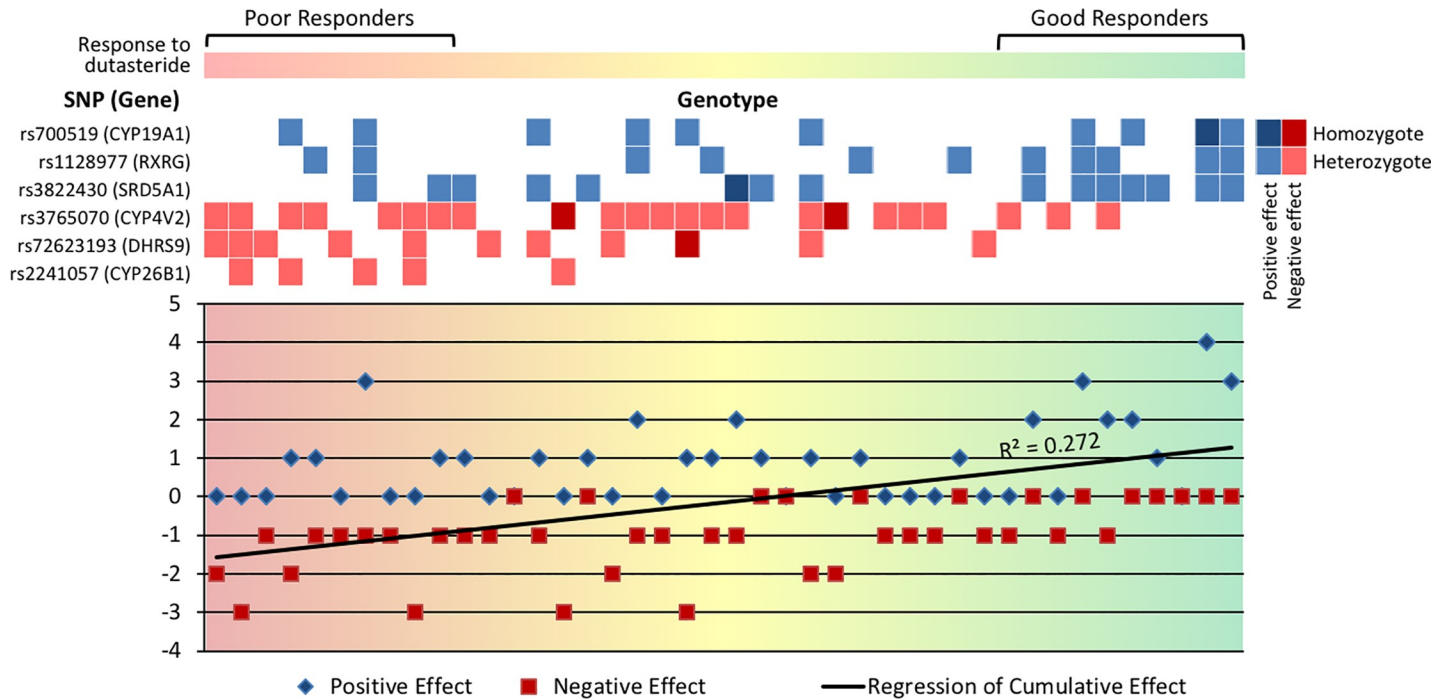


Fig 1. Genotypic landscape of 42 patients and the cumulative effect of each allele count and their positive or negative effect. In the upper figure, boxes represent the genotypes of each sample. Light-coloured boxes represent heterozygote SNPs and dark-coloured boxes represent homozygote SNPs. SNPs with negative effects and positive effects, as determined from the regression test, are in the red boxes and blue boxes, respectively. The lower figure shows cumulative effect by counting the number of SNPs. By considering SNPs with a negative effect as -1 and those with a positive effect as +1 as well as by applying regression tests to these cumulative scores, a positive relationship between the cumulative effect of SNPs and the change in hair count was observed. In both figures, samples are ordered along the X-axis from those showing the poorest to the best response in terms of change in hair count.

<https://doi.org/10.1371/journal.pone.0222533.g001>

region after removing low quality, homogeneous SNPs, and singletons. From these SNPs, we could identify 86 exonic variants (34 non-synonymous, 49 synonymous, 1 stop-gain, and 2 unknown) and 1 splice-junction variant that affects protein coding. Copy number variation was assessed using CoNIFER [24]. We found 2 deletions in the poor-response group: 1 was found in *CYP4Z2P* of the poorest responder and the other in *CYP2A6* of 2 patients who showed 2nd and 9th poorest responses (S3 Fig). However, because of the small sample size, we could not determine the statistical significance of these deletions.

To identify variants associated with response to dutasteride, we first performed quantitative regression tests on all 42 samples' genotypes and found 162 candidate variants that were significantly associated with an increase in hair count after dutasteride treatment (S5 Table). Interestingly, *ESR1* had 51 variants that were strongly associated with response to dutasteride, all of which are located in the intronic region.

Further, to identify the relevant genetic variants more precisely, we performed a case-control association test by comparing the good- and poor-response groups. We divided subjects into 2 groups based on the increase in hair count after treatment with dutasteride and set the upper half as cases and the lower half as controls. A total of 137 variants were significant with $p < 0.05$, and *ESR1* was again determined to be the gene with the most number of significant variants (a total of 40) (S6 Table).

To identify genotypes associated with extreme phenotypes, we performed additional case-control association tests with the top and bottom quartiles of subjects and found 127 variants that were significant (S7 Table). Additionally, Fisher's exact test on these 2 quartiles showed 53

Table 1. SNPs in the exonic region that are statistically associated with response to dutasteride.

Chr	Position	SNP	Cyto band	Allele frequency			Gene	Base Change	A. a. Change	Func. Impt.	P-value				
				Poor	ASN	Good					AQt	AQt*	H	Q	QF
2	169,952,166	rs72623193	2q31.1	0.25	0.13	0.00	<i>DHRS9</i>	T > C	-	GERP++ 4.17	0.019	0.021	0.014	0.0098	0.047
2	72,361,960	rs2241057	2p13.2	0.20	0.08	0.00	<i>CYP26B1</i>	A > G	L>S	Benign (0.78, 0.00) GERP++ 3.32	0.023	0.021	0.017	0.025	0.106
1	165,389,129	rs1128977	1q23.3	0.10	0.11	0.25	<i>RXRG</i>	G > A	-	GERP++4.3	0.034	0.050	0.292	0.16	0.41
5	6,651,970	rs3822430	5p15.31	0.10	0.15	0.35	<i>SRD5A1</i>	A > G	-		0.035	0.033	0.082	0.025	0.13
5	6,652,009	rs8192186	5p15.31	0.10	0.15	0.35	<i>SRD5A1</i>	G > A	-		0.035	0.033	0.082	0.025	0.13
5	6,656,210	rs3736316	5p15.31	0.10	0.15	0.35	<i>SRD5A1</i>	G > A	-		0.035	0.033	0.082	0.025	0.13
15	51,507,968	rs700519	15q21.2	0.10	0.18	0.25	<i>CYP19A1</i>	G > A	R>C	Impt (0.04, 0.00) GERP++ 4.4	0.040	0.044	0.072	0.212	0.408
4	187,122,355	rs3736456	4q35.2	0.35	0.37	0.15	<i>CYP4V2</i>	T > C	-		0.041	0.050	0.183	0.074	0.273

ASN, 1000G Phase1 East Asian; A. a, Amino acid; AQt, All Quantitative association test; Chr, Chromosome; Func Impt, Functional Important; Good, Good responders; H, Half case-control association test; Poor, Poor responders; Q, Quartile case-control association test; QF, Quartile Fisher’s exact test.

*Age adjusted

Listed SNPs are associated with good/poor response in the group-based tests or are predicted to be functionally important. Forma numbers were predicted with SIFT (0–0.05, where lower scores are more important) and PolyPhen2 (0–1, where higher scores are more important). GERP++ score indicates evolutionally conserved regions (GERP++>2). P-values < 0.05 are indicated in bold. None of these p-values survived multiple test correction due to the small sample size. Regardless, we identified the same candidate locus ranked from 4 different association tests and considered it worthy of reporting.

<https://doi.org/10.1371/journal.pone.0222533.t001>

variants that were more stringently associated with response to dutasteride (S8 Table). In addition to variants in *ESR1*, a few intronic variants in *DHRS9* were highly ranked in these 2 tests.

To identify the functional effects of SNPs, we summarized the exonic variants identified from these 4 association tests (Table 1). Synonymous SNP rs72623193 in the *DHRS9* gene was the most significant variant and the only exonic variant that passed all 4 association tests. A linear regression test revealed a negative effect of this SNP on response to dutasteride ($\beta = -11.70$).

Two missense mutations, rs2241057 in *CYP26B1* and rs700519 in *CYP19A1*, were also associated with response to dutasteride. Rs2241057 showed a negative effect ($\beta = -17.66$) and was found only in subjects of the poor-response group with heterozygote genotypes. On the other hand, rs700519 had a positive effect ($\beta = 10.56$) with a tendency of increased allelic frequency in good responders. However, these mutations were not predicted to cause a dramatic damage in gene function. PolyPhen2 predicted both mutations as ‘benign,’ whereas SIFT predicted the latter mutation as deleterious.

To determine the gene-level association, we performed gene-based association tests using previously determined significant genotypes of the SNPs described above (Table 2). *DHRS9* and *ESR1* passed all 4 gene-based association tests. Interestingly, most of the variants in these genes had shared haplotypes (S4A and S4B Fig) and the same type of effect in linear regressions. SNPs in *DHRS9* had a negative effect ($\beta \leq -10.13$) compared to SNPs in *ESR1* ($\beta \geq 8.378$) (S9 Table). Although they did not pass all of the association tests, *RARG*, *H3F3A*, *SRD5A1*, and *HDAC1* are also suggested to have substantial effects on response to dutasteride. SNPs in *SRD5A1* showed a positive association with drug response ($\beta \geq 9.095$). An LD plot of these SNPs shows that they are significantly associated ($p < 0.05$) in a 25-Kb long region of *SRD5A1* (S4(C) Fig).

By taking 1 exonic variant from each gene as a tagged SNP, each variant with a positive effect was treated as a protective factor and that with a negative effect was treated as a risk factor. We empirically calculated the cumulative effect by adding allele counts of each SNP and

Table 2. Results of the gene-level association tests.

Chr	Gene	Start	Stop	P-values			
				AQt	H	Q	QF
2	<i>DHRS9</i>	169,629,544	169,660,923	0.007	0.018	0.005	0.009
6	<i>ESR1</i>	152,053,323	152,466,101	0.04	0.046	0.048	0.047
12	<i>RARG</i>	51,890,619	51,912,303	0.012	0.318	0.022	0.056
1	<i>H3F3A</i>	224,317,043	224,326,326	0.037	0.594	0.094	0.166
5	<i>SRD5A1</i>	6,686,499	6,722,675	0.04	0.135	0.063	0.219
1	<i>HDAC1</i>	32,530,294	32,571,811	0.029	0.106	0.096	0.357

AQt, All Quantitative association test; H, Half case-control association test; Q, Quartile case-control association test; QF, Quartile Fisher's exact test.

This table shows genes significantly associated with response to dutasteride ($p < 0.05$). This test was performed using VEGAS, with start-stop positions in hg18.

<https://doi.org/10.1371/journal.pone.0222533.t002>

subtracting the risk factors from the protective factors. The regression of the cumulative effect was positively correlated with response to dutasteride (Fig 1). Allele frequencies of each SNP were different between poor and good responders. Asian frequency obtained from the 1000 Genomes Project was in-between allele frequencies of the poor-/good-response groups in all SNPs except for *CYP4V2* (S5 Fig).

To examine hidden associations, also called as gene-gene interactions or epistasis, we performed a non-parametric, genetic model-free data mining strategy MDR by using the 6 loci described above. The results of the cross-validation statistics used for training and testing the MDR model and the predicted class of each genotype combination were shown in S10 and S11 Tables. For a detailed model description of values in entropy-based measures for information gain (IG) and interaction (I), see S12 and S13 Tables. For equations used to calculate IG and I [28]. This model had cross-validation consistency of 10 with a balanced accuracy of 87.19% (Fig 2 and S10–S13 Tables). Polymorphisms in *CYP19A1* and *DHRS9* synergistically interact (i.e. non-additive or exclusive-OR interactions) and had the most entropy (3.77%), followed by *CYP26B1* and *CYP4V2* (3.19%). *DHRS9* and *CYP26B1* are highly additive (redundancy or lack of information gain, -7.35%). This result indicates existence of nonlinear interactions between SNPs that underlie the cumulative effect.

Discussion

Dutasteride is a dual inhibitor of type I and II 5ARs and has proven to be an effective and tolerable treatment modality for MPHL. It improves hair growth in a dose-dependent manner by reduction of serum and scalp DHT [16]. The response to dutasteride varies in each individual, but genetic factors related to individual differences are unclear. In this study, we found several candidate regions associated with response to dutasteride in MPHL treatment through genotypic association tests.

In our study, the most significant exonic variant was rs72623193 located in *DHRS9*, that was ranked the highest in the gene-based tests. *DHRS9* (dehydrogenase/reductase SDR family member 9) has been identified as a 3α -hydroxysteroid dehydrogenase (3α -HSD) [29]. It is a microsomal enzyme whereas other 3α -HSDs in the aldoketoreductase gene superfamily (AKR1C1–AKR1C4) are cytosolic enzymes. This enzyme is expressed in the human epidermis, hair follicles, and sebaceous glands [30, 31]. Since *DHRS9* is involved in the synthesis of DHT from 3α -androstenediol [29], upregulation would facilitate the backdoor pathway to DHT in the scalp tissue even when 5AR is inhibited by dutasteride (S6(A) Fig). A previous study showed that there was negative correlation between the change in hair count and the percent change of scalp DHT after treatment with the 5AR inhibitor [16].

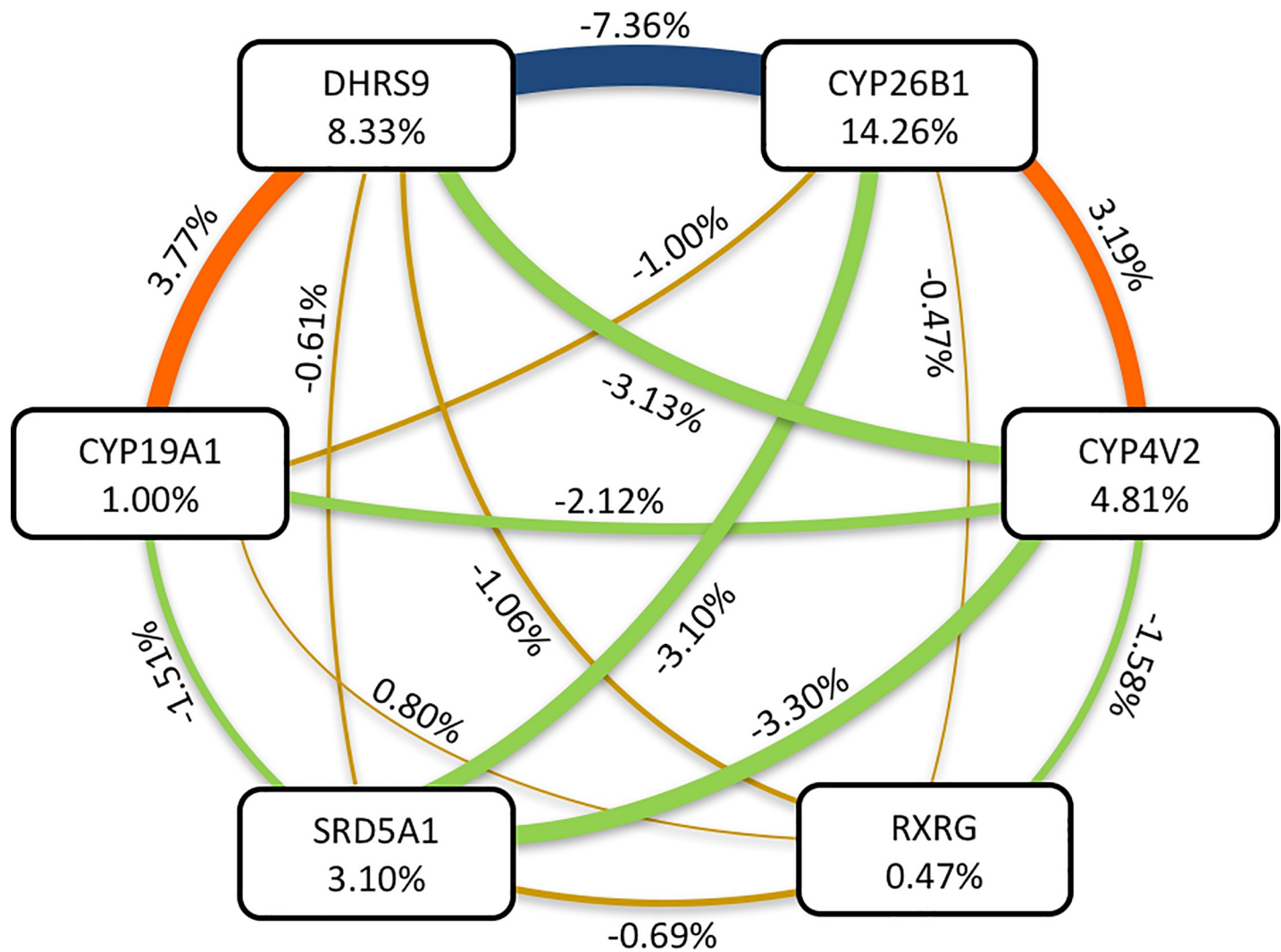


Fig 2. Entropy-based interaction graph for exonic variations drawn with Multifactor Dimension Reduction (MDR) analysis of good-/poor responders treated as cases/controls. Positive entropy indicates information gain or synergistic effect (i.e. non-additive), whereas negative entropy indicates redundancy or lack of information gain. Model statistics as well as single and pairwise attribute values for information gain and interactions are listed in S10–S13 Tables.

<https://doi.org/10.1371/journal.pone.0222533.g002>

Additionally, DHRS9 possesses retinol dehydrogenase (RoldH) activity [32, 33], which converts retinol to retinal in retinoid acid (RA) synthesis (S6(B) Fig). Because DHRS9 can oxidize retinol that is bound to the cellular retinol-binding protein while many other RoldHs cannot [34, 35], it is regarded as a physiological enzyme [35]. Its expression changes during hair cycling in mice and increases in hair follicles of C57BL/6J mice that frequently develop dorsal skin alopecia. DHRS9 also increased in the hair follicles of patients with central centrifugal cicatricial alopecia [30]. Previous studies demonstrated that RA synthesis and the RA signalling system are related to hair growth and cycling [35, 36]. Furthermore, RA was reported to regulate bone morphogenetic protein and Wnt signalling pathways, which are involved in hair growth [37].

Rs2241057, the second most significant exonic SNP in this study, is located on *CYP26B1*, which is also involved in regulation of RA level. RA is metabolized by cytochrome P450 26 family members (*CYP26A1*, *B1*, and *C1*) (S6(B) Fig) [38]. *CYP26A1* and *CYP26B1* are expressed and are RA-inducible in both reconstructed and *in vivo* human epidermis [39, 40], but *CYP26B1* was suggested to be more effective under physiological conditions [41],

CYP26B1 is able to inactivate both all-*trans*-RA and 9-*cis*-RA [42]. Additionally, we found that rs1128977 on *RXRG* is associated with response to dutasteride. This gene encodes retinoid X receptor- γ (*RXRG*), a nuclear receptor that is activated by 9-*cis*-RA. Taken together, these findings strongly suggest that RA metabolism and the RA signalling pathway are associated with response to dutasteride in MPHL, although mechanism of action needs to be elucidated in future studies.

We also found that response to dutasteride may be associated with variants of genes involved in steroid hormone metabolism, such as *SRD5A1*, *ESR1*, and *CYP19A1* (S6(A) Fig). The product of *SRD5A1* is type I 5AR, which is a target of dutasteride. Dutasteride inhibits both type I and II 5ARs, while finasteride selectively inhibits type II 5AR. Therefore, a variation in *SRD5A1* might affect the subject's response to dutasteride during treatment for MPHL. However, *SRD5A2* encoding type II 5AR was not associated with response to dutasteride.

In addition to androgens, it has long been known that oestrogen also effects hair follicle growth by binding locally expressed oestrogen receptors (ERs) [43]. *ESR1* encodes ER α , which is 1 of 2 distinct isoforms of ERs. *CYP19A1* converts androstenedione to estrone (E1) and testosterone to 17 β -estradiol (E2). It has been reported that ER α is maximally expressed during telogen, and E2 inhibits hair growth [44, 45]. Moreover, E2 modifies androgen metabolism in pilosebaceous units [43].

In this study, no individual variant passed the filtering criteria after multiple tests correction was applied. This might be due to the small number of participants enrolled in this study. However, we cannot exclude the effects of interaction between multiple variants, considering that MPHL is very complex, and the hair-count change in each patient after treatment was distributed continuously and was not clearly divided into different groups. We hypothesize that the individual difference in response to dutasteride comes from the cumulative effects of multiple variants rather than that of a single responsible variant and depicts a genotypic landscape with good correlation between the allele frequency scores of the tagged exonic SNPs and response to dutasteride. Results from MDR analysis are also consistent with the existence of an interaction between SNPs. Therefore, we propose that these SNPs, when used simultaneously, could be suitable markers for clinicians to predict response to dutasteride in patients with MPHL.

In conclusion, we identified candidate variants associated with response to dutasteride for MPHL. According to the association tests, potential candidate variants are located on *DHRS9*, *CYP26B1*, *ESR1*, *SRD5A1*, *CYP19A1*, and *RXRG*. These genes encode proteins involved in the backdoor pathway of DHT synthesis, RA metabolism and the RA signalling pathway, and steroid hormone metabolism. To the best of our knowledge, this study is the first to show genetic variation associated with drug response in patients with MPHL. Moreover, this study could be developed further to investigate the cumulative effect of multiple variants. Our findings would contribute to the investigation on the pathophysiology of MPHL and can potentially help clinicians improve treatment of MPHL patients by using 5AR inhibitors.

Supporting information

S1 Table. List of the target genes.

(XLSX)

S2 Table. Epidemiologic data of the individual subjects.

(XLSX)

S3 Table. Epidemiological characteristics of the enrolled subjects with MPHL who had been treated with dutasteride (0.5 mg) for 6 months.

(XLSX)

S4 Table. Sequencing summary of 42 subjects.

(XLSX)

S5 Table. List of SNPs from quantitative linear regression test under p-value < 0.05.

(XLSX)

S6 Table. List of SNPs from upper half (n = 21) and lower half (n = 21) case-control association test under p-value < 0.05.

(XLSX)

S7 Table. List of SNPs from extreme phenotypes (n = 10 on each) case-control association test under p-value < 0.05.

(XLSX)

S8 Table. List of SNPs from extreme phenotypes (n = 10 on each) case-control Fisher's exact test under p-value < 0.05.

(XLSX)

S9 Table. Genes and the number of SNPs commonly and significantly associated (p < 0.05) with response to dutasteride as determined by 4 association tests. Effect, $|\beta|$, and R² show the minimum to maximum range for each SNP used based on quantitative regression test results.

(XLSX)

S10 Table. Results of the cross-validation statistics used for training and testing the MDR model with the 6 genotyped exonic SNPs.

(XLSX)

S11 Table. Predicted class of each genotype combination. Genotype combination is in order of SNPs in *DHRS9*, *CYP26B1*, *RXRG*, *SRD5A1*, *CYP19A1*, and *CYP4V2* used in the MDR model, and their number of alternate alleles noted as 0, 1, and 2.

(XLSX)

S12 Table. Single attribute values in entropy-based measures for information gain and interaction depicted in Fig 2.

(XLSX)

S13 Table. Pairwise attribute values in entropy-based measures for information gain and interaction depicted in Fig 2.

(XLSX)

S1 Fig. Representative phototrichogram for the measurement of hair counts. Total hair counts and growth rate were measured on 1 cm² circular area of clipped hair, selected from the anterior edge of the balding area on the vertex. A small cosmetic ink tattoo was marked in the center of the selected area to identify the same area of measurement at every measurement (yellow arrow). The phototrichogram was converted into dot map, and then converted to hair counts using a computer imaging system.

(PNG)

S2 Fig. Schematic representation of SNP calling, variant filtering steps, and statistical analysis workflow. Tools used for each step are marked in blue; number of variants left after each filtering step are marked in red. After filtration and quality control, 4 different statistical analyses were applied at the SNP and gene levels. First, linear regression tests were applied to all samples; then, we divided our samples into 2 groups for the case/control association test. This

test has been applied twice by taking half of the samples as cases/controls and quartiles of the poor-/good-response groups. Last, Fisher's exact tests have been applied to the poor-/good-response groups.

(TIF)

S3 Fig. Copy number variations found in the poor-response group. Red-marked samples represent poor responders and blue-marked samples represent good responders. (A) Deletion in *CYP4Z2P* in the poorest responder patient (A037; hair count change of -22.5). This is a novel deletion that has not previously been reported [1]. (B) Deletion in *CYP2A6* in a patient with the 2nd poorest response (A007; hair count change of -10.5) and another patient with the 9th poorest response (A043; hair count change of -3.0). This deletion was reported to be commonly found in 20% of the Asian population (Lee C, Seo JS et al. Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. *Nature Genetics*, 2010; 42: 400–405)

(PDF)

S4 Fig. Linkage disequilibrium (LD) blocks of variants in genes associated with drug response in MPHL treatment. Green-coloured rs numbers are SNPs in coding sequences, and black-coloured rs numbers are SNPs in intronic regions. (A) Variants in the LD blocks of *ESR1*. (B) Variants in the LD block of *DHRS9*. (C) Variants in the LD block of *SRD5A1*.

(PDF)

S5 Fig. Comparison of the allele frequency of each group as determined by the present study with the Asian allele frequency obtained from the 1000 Genomes Project. Asian frequency was almost in-between allele frequencies of the poor-/good-responders. ASN, Asian population; good, good-responders; poor, poor-responders.

(PDF)

S6 Fig. Overview of metabolic pathways of steroids and retinoids. (A) Metabolic pathways of steroids. DHT is normally produced from testosterone in target tissues (conventional pathway, black box), but is also synthesized from androstanediol (back-door pathway, blue box) by a reaction in which *DHRS9* is involved. *CYP19* converts androstenedione into estrone and testosterone into estradiol (oestrogen biosynthesis, pink box). (B) Overview of retinoid metabolism in target cells. Retinol is transported into the cell where binds *CRBP*, which is oxidized to retinal by *DHRS9*. Retinal is further oxidized to retinoic acid. *CYP26* is involved in retinoic acid metabolism. *ALDH1*, retinal dehydrogenase 1; *CRABP2*, cellular retinoic acid binding protein II; *CRBP*, cellular retinol-binding protein; *CYP17*, cytochrome P450 17 family members; *CYP19*, cytochrome P450 19 family members; *CYP26*, cytochrome P450 26 family members; *DHRS9*, dehydrogenase reductase member 9; DHT, dihydrotestosterone; 3α -HSD, 3α -hydroxysteroid dehydrogenase; 17β -HSD, 17β -hydroxysteroid dehydrogenase.

(PDF)

Acknowledgments

The biospecimens and data used in this study were provided by the Biobank of Seoul National University Hospital, a member of Korea Biobank Network.

Author Contributions

Conceptualization: Dong Young Kim, Ohsang Kwon, Jong-Il Kim, Seong Jin Jo.

Data curation: Arang Rhie, Ho-Young Son, Bark-Lynn Lew, Woo-Young Sim.

Formal analysis: Arang Rhie, Ho-Young Son.

Funding acquisition: Ohsang Kwon.

Investigation: Arang Rhie, Ho-Young Son, Soo Jung Kwak, Seungbok Lee, Seong Jin Jo.

Methodology: Jeong-Sun Seo.

Software: Arang Rhie, Ho-Young Son.

Supervision: Jeong-Sun Seo, Ohsang Kwon, Jong-Il Kim.

Validation: Soo Jung Kwak, Dong Young Kim.

Writing – original draft: Arang Rhie, Seong Jin Jo.

Writing – review & editing: Ho-Young Son, Soo Jung Kwak, Dong Young Kim, Ohsang Kwon, Seong Jin Jo.

References

1. Sinclair R. Male pattern androgenetic alopecia. *BMJ*. 1998; 317(7162):865–9. <https://doi.org/10.1136/bmj.317.7162.865> PMID: 9748188; PubMed Central PMCID: PMC1113949.
2. Whiting DA. Male pattern hair loss: current understanding. *Int J Dermatol*. 1998; 37(8):561–6. <https://doi.org/10.1046/j.1365-4362.1998.00542.x> PMID: 9731996
3. Norwood OT. Male pattern baldness: classification and incidence. *South Med J*. 1975; 68(11):1359–65. <https://doi.org/10.1097/00007611-197511000-00009> PMID: 1188424.
4. Nyholt DR, Gillespie NA, Heath AC, Martin NG. Genetic basis of male pattern baldness. *J Invest Dermatol*. 2003; 121(6):1561–4. <https://doi.org/10.1111/j.1523-1747.2003.12615.x> PMID: 14675213
5. Ellis JA, Stebbing M, Harrap SB. Polymorphism of the androgen receptor gene is associated with male pattern baldness. *J Invest Dermatol*. 2001; 116(3):452–5. <https://doi.org/10.1046/j.1523-1747.2001.01261.x> WOS:000167222200014. PMID: 11231320
6. Hayes VM, Severi G, Eggleton SA, Padilla EJD, Southey MC, Sutherland RL, et al. The E211 G > A androgen receptor polymorphism is associated with a decreased risk of metastatic prostate cancer and androgenetic alopecia. *Cancer Epidemiol Biomarkers*. 2005; 14(4):993–6. <https://doi.org/10.1158/1055-9965.Epi-04-0778> WOS:000228351300038. PMID: 15824176
7. Hillmer AM, Hanneken S, Ritzmann S, Becker T, Freudenberg J, Brockschmidt FF, et al. Genetic variation in the human androgen receptor gene is the major determinant of common early-onset androgenetic alopecia. *Am J Hum Genet*. 2005; 77(1):140–8. <https://doi.org/10.1086/431425> PMID: 15902657
8. Levy-Nissenbaum E, Bar-Natan M, Frydman M, Pras E. Confirmation of the association between male pattern baldness and the androgen receptor gene. *Eur J Dermatol*. 2005; 15(5):339–40. WOS:000232392300005. PMID: 16172040
9. Richards JB, Yuan X, Geller F, Waterworth D, Bataille V, Glass D, et al. Male-pattern baldness susceptibility locus at 20p11. *Nat Genet*. 2008; 40(11):1282–4. <https://doi.org/10.1038/ng.255> WOS:000260501500017. PMID: 18849991
10. Hillmer AM, Flaquer A, Hanneken S, Eigelshoven S, Kortum AK, Brockschmidt FF, et al. Genome-wide scan and fine-mapping linkage study of androgenetic alopecia reveals a locus on chromosome 3q26. *Am J Hum Genet*. 2008; 82(3):737–43. <https://doi.org/10.1016/j.ajhg.2007.11.014> PMID: 18304493; PubMed Central PMCID: PMC2427264.
11. Li R, Brockschmidt FF, Kiefer AK, Stefansson H, Nyholt DR, Song K, et al. Six novel susceptibility Loci for early-onset androgenetic alopecia and their unexpected association with common diseases. *PLoS Genet*. 2012; 8(5):e1002746. Epub 2012/06/14. <https://doi.org/10.1371/journal.pgen.1002746> PMID: 22693459; PubMed Central PMCID: PMC3364959.
12. Heilmann S, Kiefer AK, Fricker N, Drichel D, Hillmer AM, Herold C, et al. Androgenetic alopecia: identification of four genetic risk loci and evidence for the contribution of WNT signaling to its etiology. *J Invest Dermatol*. 2013; 133(6):1489–96. Epub 2013/01/30. <https://doi.org/10.1038/jid.2013.43> PMID: 23358095.
13. Brockschmidt FF, Heilmann S, Ellis JA, Eigelshoven S, Hanneken S, Herold C, et al. Susceptibility variants on chromosome 7p21.1 suggest HDAC9 as a new candidate gene for male-pattern baldness. *Br J Dermatol*. 2011; 165(6):1293–302. Epub 2011/10/29. <https://doi.org/10.1111/j.1365-2133.2011.10708.x> PMID: 22032556.

14. Kaufman KD, Olsen EA, Whiting D, Savin R, DeVillez R, Bergfeld W, et al. Finasteride in the treatment of men with androgenetic alopecia. *J Am Acad Dermatol*. 1998; 39(4):578–89.
15. Eun HC, Kwon OS, Yeon JH, Shin HS, Kim BY, Ro BI, et al. Efficacy, safety, and tolerability of dutasteride 0.5 mg once daily in male patients with male pattern hair loss: A randomized, double-blind, placebo-controlled, phase III study. *J Am Acad Dermatol*. 2010; 63(2):252–8. <https://doi.org/10.1016/j.jaad.2009.09.018> WOS:000280252100008. PMID: 20605255
16. Olsen EA, Hordinsky M, Whiting D, Stough D, Hobbs S, Ellis ML, et al. The importance of dual 5 α -reductase inhibition in the treatment of male pattern hair loss: results of a randomized placebo-controlled study of dutasteride versus finasteride. *J Am Acad Dermatol*. 2006; 55(6):1014–23. <https://doi.org/10.1016/j.jaad.2006.05.007> PMID: 17110217
17. Prodi DA, Pirastu N, Maninchedda G, Sassu A, Picciau A, Palmas MA, et al. EDAR2R is associated with androgenetic alopecia. *J Invest Dermatol*. 2008; 128(9):2268–70. <https://doi.org/10.1038/jid.2008.60> PMID: 18385763
18. Hillmer AM, Freudenberg J, Myles S, Herms S, Tang K, Hughes DA, et al. Recent positive selection of a human androgen receptor/ectodysplasin A2 receptor haplotype and its relationship to male pattern baldness. *Hum Genet*. 2009; 126(2):255–64. <https://doi.org/10.1007/s00439-009-0668-z> PMID: 19373488
19. Hayes VM, Severi G, Padilla EJ, Morris HA, Tilley WD, Southey MC, et al. 5 α -Reductase type 2 gene variant associations with prostate cancer risk, circulating hormone levels and androgenetic alopecia. *Int J Cancer*. 2007; 120(4):776–80. <https://doi.org/10.1002/ijc.22408> PMID: 17136762
20. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9(4):357–9. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286; PubMed Central PMCID: PMC3322381.
21. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010; 20(9):1297–303. <https://doi.org/10.1101/gr.107524.110> PMID: 20644199; PubMed Central PMCID: PMC2928508.
22. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011; 43(5):491–8. <https://doi.org/10.1038/ng.806> WOS:000289972600023. PMID: 21478889
23. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010; 38(16):e164–e. <https://doi.org/10.1093/nar/gkq603> PMID: 20601685
24. Krumm N, Sudmant PH, Ko A, O’Roak BJ, Malig M, Coe BP, et al. Copy number variation detection and genotyping from exome sequence data. *Genome Res*. 2012; 22(8):1525–32. <https://doi.org/10.1101/gr.138115.112> PMID: 22585873; PubMed Central PMCID: PMC3409265.
25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: 17701901; PubMed Central PMCID: PMC1950838.
26. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2):263–5. <https://doi.org/10.1093/bioinformatics/bth457> PMID: 15297300.
27. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet*. 2010; 87(1):139–45. <https://doi.org/10.1016/j.ajhg.2010.06.009> PMID: 20598278; PubMed Central PMCID: PMC2896770.
28. Moore JH, Gilbert JC, Tsai CT, Chiang FT, Holden T, Barney N, et al. A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J Theor Biol*. 2006; 241(2):252–61. <https://doi.org/10.1016/j.jtbi.2005.11.036> PMID: 16457852.
29. Chetyrkin SV, Belyaeva OV, Gough WH, Kedishvili NY. Characterization of a Novel Type of Human Microsomal 3 α -Hydroxysteroid Dehydrogenase. *J Biol Chem*. 2001; 276(25):22278–86. <https://doi.org/10.1074/jbc.M102076200> PMID: 11294878
30. Everts HB, Silva KA, Montgomery S, Suo L, Menser M, Valet AS, et al. Retinoid metabolism is altered in human and mouse cicatricial alopecia. *J Invest Dermatol*. 2013; 133(2):325–33. <https://doi.org/10.1038/jid.2012.393> PMID: 23096705; PubMed Central PMCID: PMC3546159.
31. Markova NG, Pinkas-Sarafova A, Karaman-Jurukovska N, Jurukovski V, Simon M. Expression pattern and biochemical characteristics of a major epidermal retinol dehydrogenase. *Mol Genet Metab*. 2003; 78(2):119–35. PMID: 12618084.
32. Rexer BN, Ong DE. A novel short-chain alcohol dehydrogenase from rats with retinol dehydrogenase activity, cyclically expressed in uterine epithelium. *Biol Reprod*. 2002; 67(5):1555–64. <https://doi.org/10.1095/biolreprod.102.007021> PMID: 12390888.

33. Soref CM, Di Y-P, Hayden L, Zhao YH, Satre MA, Wu R. Characterization of a novel airway epithelial cell-specific short chain alcohol dehydrogenase/reductase gene whose expression is up-regulated by retinoids and is involved in the metabolism of retinol. *J Biol Chem*. 2001; 276(26):24194–202. <https://doi.org/10.1074/jbc.M100332200> PMID: 11304534
34. Napoli JL. Interactions of retinoid binding proteins and enzymes in retinoid metabolism. *Biochim Biophys Acta*. 1999; 1440(2):139–62.
35. Everts HB, Sundberg JP, King LE Jr., Ong DE. Immunolocalization of enzymes, binding proteins, and receptors sufficient for retinoic acid synthesis and signaling during the hair cycle. *J Invest Dermatol*. 2007; 127(7):1593–604. <https://doi.org/10.1038/sj.jid.5700753> PMID: 17363914.
36. Everts HB, King LE Jr., Sundberg JP, Ong DE. Hair cycle-specific immunolocalization of retinoic acid synthesizing enzymes Aldh1a2 and Aldh1a3 indicate complex regulation. *J Invest Dermatol*. 2004; 123(2):258–63. <https://doi.org/10.1111/j.0022-202X.2004.23223.x> PMID: 15245423.
37. Balmer JE, Blomhoff R. Gene expression regulation by retinoic acid. *J Lipid Res*. 2002; 43(11):1773–808. <https://doi.org/10.1194/jlr.r100015-jlr200> PMID: 12401878.
38. Roos TC, Jugert FK, Merk HF, Bickers DR. Retinoid metabolism in the skin. *Pharmacol Rev*. 1998; 50(2):315–33. PMID: 9647871.
39. Lorie EP, Chamcheu JC, Vahlquist A, Törmä H. Both all-trans retinoic acid and cytochrome P450 (CYP26) inhibitors affect the expression of vitamin A metabolizing enzymes and retinoid biomarkers in organotypic epidermis. *Arch Dermatol Res*. 2009; 301(7):475–85. <https://doi.org/10.1007/s00403-009-0937-7> PMID: 19294396
40. Pavez Lorie E, Cools M, Borgers M, Wouters L, Shroot B, Hagforsen E, et al. Topical treatment with CYP26 inhibitor talarozole (R115866) dose dependently alters the expression of retinoid-regulated genes in normal human epidermis. *Br J Dermatol*. 2009; 160(1):26–36. <https://doi.org/10.1111/j.1365-2133.2008.08895.x> PMID: 19016711.
41. Pavez Lorie E, Li H, Vahlquist A, Torma H. The involvement of cytochrome p450 (CYP) 26 in the retinoic acid metabolism of human epidermal keratinocytes. *Biochim Biophys Acta*. 2009; 1791(3):220–8. Epub 2009/01/28. <https://doi.org/10.1016/j.bbali.2008.12.004> PMID: 19171200.
42. White JA, Ramshaw H, Taimi M, Stangle W, Zhang AQ, Everingham S, et al. Identification of the human cytochrome P450, P450RAI-2, which is predominantly expressed in the adult cerebellum and is responsible for all-trans-retinoic acid metabolism. *P Natl Acad Sci USA*. 2000; 97(12):6403–8. <https://doi.org/10.1073/pnas.120161397> WOS:000087526300033. PMID: 10823918
43. Ohnemus U, Uenalan M, Inzunza J, Gustafsson JÅ, Paus R. The hair follicle as an estrogen target and source. *Endocr Rev*. 2006; 27(6):677–706. <https://doi.org/10.1210/er.2006-0020> PMID: 16877675
44. Oh H-S, Smart RC. An estrogen receptor pathway regulates the telogen-anagen hair follicle transition and influences epidermal cell proliferation. *P Natl Acad Sci USA*. 1996; 93(22):12525–30.
45. Ohnemus U, Unalan M, Handjiski B, Paus R. Topical estrogen accelerates hair regrowth in mice after chemotherapy-induced alopecia by favoring the dystrophic catagen response pathway to damage. *J Invest Dermatol*. 2004; 122(1):7–13. <https://doi.org/10.1046/j.0022-202X.2003.22120.x> PMID: 14962083.