CAMBRIDGE
UNIVERSITY PRESS

Article

Comparison of Genome-Wide Association Scans for Quantitative and Observational Measures of Human Hair Curvature

Yvonne Y. W. Ho[1],*, Angela Mina-Vargas[1],*, Gu Zhu[1], Mark Brims[2], Dennis McNevin[3], Grant W. Montgomery[1,4], Nicholas G. Martin[5], Sarah E. Medland[1] and Jodie N. Painter[1]

[1]Psychiatric Genetics, Genetics and Computational Biology Department, QIMR Berghofer Medical Research Institute, Brisbane, Australia, [2]BSC Electronics Pty Ltd, Ardross, Australia, [3]Centre for Forensic Science, School of Mathematical & Physical Sciences, Faculty of Science, University of Technology Sydney, Sydney, Australia, [4]Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia and [5]Genetic Epidemiology, Genetics and Computational Biology Department, QIMR Berghofer Medical Research Institute, Brisbane, Australia

Abstract

Previous genetic studies on hair morphology focused on the overall morphology of the hair using data collected by self-report or researcher observation. Here, we present the first genome-wide association study (GWAS) of a micro-level quantitative measure of hair curvature. We compare these results to GWAS results obtained using a macro-level classification of observable hair curvature performed in the same sample of twins and siblings of European descent. Observational data were collected by trained observers, while quantitative data were acquired using an Optical Fibre Diameter Analyser (OFDA). The GWAS for both the observational and quantitative measures of hair curvature resulted in genome-wide significant signals at chromosome 1q21.3 close to the trichohyalin (*TCHH*) gene, previously shown to harbor variants associated with straight hair morphology in Europeans. All genetic variants reaching genome-wide significance for both GWAS (quantitative measure lead single-nucleotide polymorphism [SNP] rs12130862, $p = 9.5 \times 10^{-09}$; observational measure lead SNP rs11803731, $p = 2.1 \times 10^{-17}$) were in moderate to very high linkage disequilibrium (LD) with each other (minimum $r^2 = .45$), indicating they represent the same genetic locus. Conditional analyses confirmed the presence of only one signal associated with each measure at this locus. Results from the quantitative measures reconfirmed the accuracy of observational measures.

Keywords: Genome-wide association study; *TCHH* gene; human hair shape; SNPs

Human scalp hair has a notoriously variable morphology. The shape and degree of the curl in hair fibers are associated with the number of neocortical cells and the distribution and types of keratins (Thibaut et al., 2007). Previous research has shown that hair follicle morphogenesis is affected by a complex interaction between growth and homeostasis hormone pathways, as well as regulatory pathways (Vogt et al., 2008). The basic structure of hair (cuticle, cortex and medulla) is common to all humans; however, the dimensional shape of the hair fiber and ultrastructure varies considerably among and within populations (Koch et al., 2019; Loussouarn et al., 2007).

The heritability of self-reported hair characteristics has typically been found to be high; hair color was reported to show a 73−99% broad-sense heritability in a Dutch sample (Lin et al., 2015), while hair curliness was found to be 85−95% heritable in an Australian cohort (Medland, Nyholt et al., 2009). Moreover, a previous

analysis using the same sample utilized in this present study showed that hair curvature is a heritable trait with strong genetic influence ($h^2 = .52−.64$; Ho et al., 2016).

Previous genome-wide association studies (GWAS) discovery and meta-analyses have identified 12 loci (located near the trichohyalin [*TCHH*], ectodysplasin A receptor [*EDAR*], *WNT10A*, *FRAS1*, *OFCC1*, *TRAF2*, *PRSS53*, *PADI3*, *LOC105373470*, *S100A11*, *LCE3E* and *LOC391485* genes) associated with hair morphology and structure (Adhikari et al., 2016; Fujimoto et al., 2008; Liu et al., 2018; Medland, Nyholt et al., 2009; Pospiech et al., 2018; Westgate et al., 2017). These studies found that the *TCHH* gene has a major effect on hair morphology in European and Latin American populations, but no effect in East Asian populations (Eriksson et al., 2010; Medland, Nyholt et al., 2009). By contrast, the *EDAR* gene, on chromosome 2, has a strong association and high prediction accuracy for straight hair in East Asians (Tan et al., 2013) and a minor (less than 0.2%) effect on hair structure in Europeans (Liu et al., 2018).

The studies cited above utilized macro-level categorical measures of hair curvature (e.g., straight, wavy or curly), assessed via self-report or researcher classification. While these studies demonstrated that simple classification schemes can be used to identify genetic variants influencing hair curvature, the present study sought to assess whether the use of micro-level quantitative assessments of hair curvature would increase the power to identify

additional genomic loci influencing this trait. A GWAS utilizing digital quantification of continuous variation in human eye color found three new loci in comparison to the use of categorical eye color measures in the same individuals (Liu et al., 2010). This indicates that using quantitative measures may be one way of finding novel genetic variants associated with human morphological traits.

With the aim of identifying additional loci associated with hair morphology using a quantitative measure, we follow-up on an earlier study by our group that evaluated the heritability and genetic correlations of a quantitative micro-level measure of hair curvature in a sample of twins and siblings of European descent (Ho et al., 2016). In the current study, we evaluate whether a GWAS of the quantitative micro-level hair curvature measure replicates the GWAS results based on observational macro-level measures of hair morphology in the same sample.

Methods

Sample

Details of the sample used in this study are reported in Ho et al. (2016). Briefly, measures of hair curvature were obtained for 3057 twins and their siblings, within the context of the Brisbane Longitudinal Twin Study (BLTS; Wright & Martin, 2004). The BLTS is an ongoing study initiated in 1992 in Brisbane, Australia. Its initial focus was research on melanoma; however, the study has collected information on a wide range of traits and, over the years, many parallel studies have been derived from it. Participants in the BLTS provide information on many traits such as height, skin color and wellbeing. Participants also provided a blood sample that is used for biochemical and DNA analyses. For hair studies, clippings of 2−4 cm were taken from the scalp of participants.

Written informed consent was obtained from participants and their parents or guardians. Ethics approval for the study was obtained from the QIMR Berghofer Medical Research Institute Human Research Ethics Committee.

Measures

Quantitative measure of hair curvature. The quantitative micro-level measures of hair curvature were obtained using an Optical Fibre Diameter Analyzer (OFDA 2000, BSC Electronics, Myaree, WA, Australia). The OFDA is an industry-standard technology used for measuring wool fiber diameter and quality (Wortmann & Schwan-Jonczyk, 2006). For this study, we used hair segments of 200 µm in length. The OFDA measured the angle of curvature of the hair in degrees; this measure was multiplied by 5 to represent the angle of a 1 mm segment of the hair strand. The data provided by the OFDA was the number of snippets (hairs) per bin (with bins ranging from 0 to 360°). In a previous publication (Ho et al., 2016), we show that the distribution of this data is approximately log-normal, therefore for this study, the data were log-normalized. Measurements beyond three standard deviations above and below the mean were removed from analyses.

Observational measure of hair curvature. The observational macro-level measure of hair curvature was conducted by trained research nurses who scored participants' hair according to a 3-point scale: straight, wavy or curly. The data were collected at up to three time points for each individual included in the study, at ages 12, 14 and 16 years. For the present study, we used the data collected at age 12.

Structural equation modeling (SEM). To determine possible common genetic effects underlying the correlation between the quantitative and observational measures of hair curvature, we used a bivariate Cholesky twin model (Neale & Cardon, 1992). Previous univariate twin models for this cohort have found that the genetic variance of hair curvature is better explained by a model that includes additive, dominant and environmental components (ADE; Ho et al., 2016; Medland, Zhu et al., 2009). In the current study, we started by fitting a bivariate Cholesky ADE model without sex limitation. Simplified AE and E models were tested for goodness of fit to determine whether parameters could be dropped from the full model. The change in model fit was tested using likelihood ratio chi-square tests. The distribution of twice the difference in log-likelihood ($-2LL$) approximates $\Delta\chi^2$, with the degrees of freedom (df) for this test equal to the difference in df between the models. SEM analyses were performed using the R package OpenMx version 2.6.7 (Boker et al., 2011; Neale et al., 2016).

Genotyping and GWAS. Genotyping data were available for 2975 individuals (twins and siblings) with hair measurement data. After quality control, data for 2225 individuals were included in the GWAS analyses. Participants were previously genotyped using the Illumina 610K genotyping array and imputed to the Haplotype Reference Consortium reference (McCarthy et al., 2016, Medland, Nyholt et al., 2009). Genome-wide association analyses were conducted using RARERMETALWORKER (Feng et al., 2014). This program assesses the evidence for association at each variant using the imputed dosage format, parameterized as $x_{ij} = \beta_{dose} + \beta_{sex} + \beta_{age} + \mu$ where j denotes the trait value for an individual within family i, while accounting for familial relatedness and zygosity. Univariate GWAS analyses were performed for each measure. Sex and age were included as covariates in both GWAS. Following each GWAS analysis, we applied standard quality control procedures using thresholds of quality score $R^2 > .5$ and minor allele frequency (MAF) > 0.05, including only those single-nucleotide polymorphisms (SNPs) reaching quality control thresholds in subsequent analyses ($N_{SNP} = 7,533,529$).

Conditional association analyses on the most significant SNP for each GWAS was performed with COJO, a utility included in the GCTA software package (Yang et al., 2011). The conditional analysis determines whether the significant signals from multiple SNPs in the same genomic region were due to high linkage disequilibrium (LD) between SNPs (i.e., SNPs in high LD representing one association signal) or due to the presence of multiple causal variants. GWAS p values were used to perform gene-based association analyses with fastBAT (Bakshi et al., 2016). This analysis takes into account the aggregated effect of a set of SNPs defined by the degree of LD between them. The 1000 Genomes Build37 (hg19; 1000 Genomes Project Consortium et al., 2015) was used as a reference genome for both the conditional analyses and gene association analyses. Bioinformatic analyses to investigate possible functional effects of the most likely candidate SNPs were conducted using FUMA (Watanabe et al., 2017), SIFT (Schneider et al., 2012) and PolyPhen-2 (Adzhubei et al., 2010).

Results

The distributions of quantitative measurements and the observational measures of straight, wavy and curly hair are presented in Figure 1. Means were not significantly different ($p > .05$), and boxplots whiskers overlap between categories ($M_{Straight} = 0.68$ $SD = 0.07$, $M_{Wavy} = 0.71$ $SD = 0.07$, $M_{Curly} = 0.80$, $SD = 0.10$).

Table 1. Results of bivariate modeling of the genetic variance of hair curvature

Variable	Model	Proportion of total variance[a]			Model fit[b]					
		A	D	E	−2LL	df	Δdf	Δχ²	AIC	p value
Quantitative	ADE	0.510 [0.26,0.62]	0.090 [1.48E-24,0.33]	0.416 [0.35,0.47]	15,537	3803			7930.6	
Observational		0.867 [0.53,0.99]	0.124 [3.5E-20,0.45]	0.008 [0,0.02]						
Quantitative	**AE**	**0.574 [0.51,0.63]**		**0.425 [0.36,0.49]**	**15,539**	**3806**	**3**	**2.73**	**7927**	**.43**
Observational		**0.991 [0.97,0.1]**		**0.008 [0.002,0.02]**						
Quantitative	E			1	16,106	3809	3	566.24	8487.5	<.001
Observational				1						

[a]Variance was decomposed into additive (A), no-additive (D) and unique environmental (E) factors for the quantitative micro-level and observational macro-level measures of hair curvature. The 95% confidence intervals are given within square brackets, with the best fitting model in bold.
[b]–2LL minus twice the log-likelihood of the model, *df*, degrees of freedom; Δ*df*, difference of degree of freedom; Δχ², change in chi-squared; AIC, Akaike information criterion.
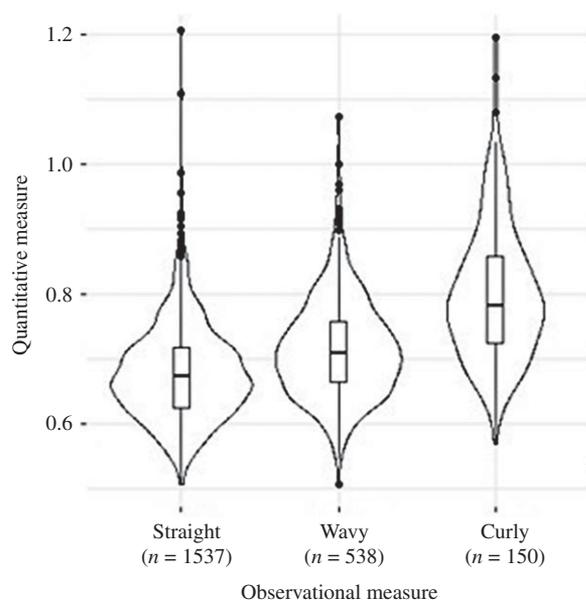

Fig. 1. Distribution of the hair curvature data. *Note:* Quantitative micro-level measures are shown on the *y*-axis and observational macro-level measures on the *x*-axis. The plot includes data for 2225 individuals for whom there were genotype data and observational and quantitative hair curvature measures available.

Fig. 2. Bivariate Cholesky model (AE model) for quantitative and observational measures of hair curvature. *Note:* This figure shows the estimated values for additive genetic (A) and unique environmental (E) components of quantitative micro-level and observational macro-level measures of hair curvature, with the 95% confidence intervals in square brackets.

Results for the bivariate ADE decomposition are presented in Table 1. The goodness of fit of the model did not decrease significantly if the component D was omitted from the model ($p = .43$). In contrast, the fit changed significantly when A and D were dropped ($p < .001$). The AE model was adopted as the best fit. Path coefficients from the AE model are presented in Figure 2. The estimated genetic variance suggests a very strong genetic component for hair curvature, 90% for the observational measures, and 58% for the quantitative measurements (Table 1). The cross-trait genetic correlations from the AE model were $rg_A = .41$ (95% CI [0.35, 0.46]).

Genome-wide Association Analysis

The GWAS of quantitative hair curvature measurements resulted in one genome-wide significant (defined as $p < 5.0 \times 10^{-8}$) association signal on chromosome 1q21.3 (Figure 3A; Supplementary Table S1). This signal is represented by five significantly associated SNPs (lead SNP rs12130862, $\beta = -0.16$, $p = 9.5 \times 10^{-09}$) that are in
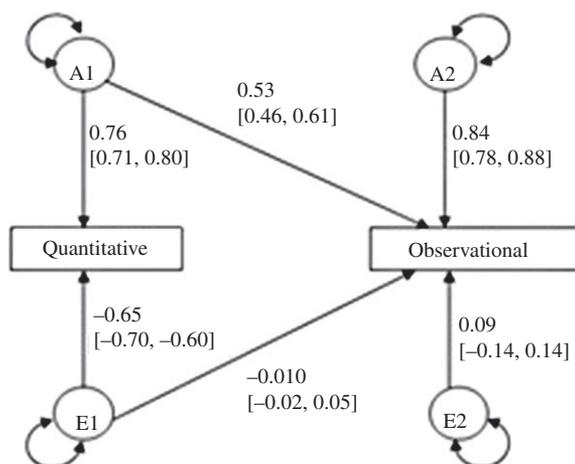
very high LD with each other (minimum r^2 between SNPs = .93; Supplementary Table S2).

The GWAS of observed measures of hair curvature also resulted in one genome-wide significant association signal, also on chromosome 1q21.3 (Figure 3B; Supplementary Table S1), with 12 significantly associated SNPs (lead SNP rs11803731, $\beta = -0.24$, $p = 2.1 \times 10^{-17}$) in moderate to very high LD with each other (minimum $r^2 = .45$; Supplementary Table S2). This result replicates previous findings (Medland, Nyholt et al., 2009) of the influence of rs11803731 on observable hair curvature. The two lead SNPs from the current analyses, rs12130862 and rs11803731, are in very high LD with each other ($r^2 = .95$). Association analyses conditioning on the top SNPs for each of the hair curvature measures revealed no additional significant SNPs for either measure at this locus, indicating these SNPs represent one association signal.

We plotted the effect sizes for the top 1000 SNPs (Figure 4) for both GWAS. The effect sizes correlate significantly ($r^2 = .815, p < .01$).

Gene-based analyses of both hair curvature GWAS supported the individual SNP findings of significant association signal (with a Bonferroni-corrected gene-based $p < 1 \times 10^{-6}$ to account for tests on 24,036 genes) on chromosome 1q21.3. For the observational measure, there were six significantly associated genes: *TCHH*, *S100A11*, *TCHHL1*, *NBPF18p*, *RPTN* and *S100A10*. All six genes
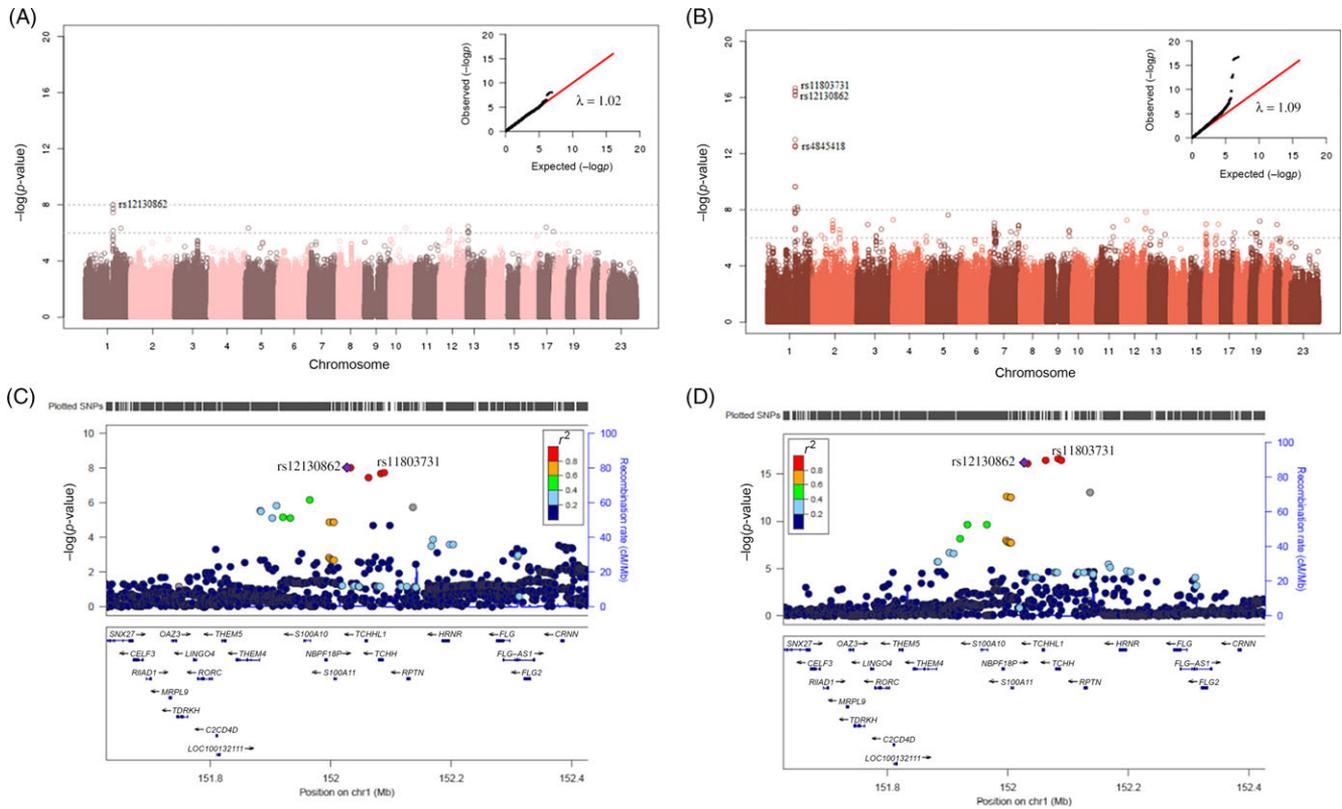
Fig. 3. Genome-wide associations for quantitative and observational measures of hair curvature and regional plots for locus 1q21.3. *Note:* The top row of panels shows Manhattan and Quantile-Quantile (QQ) plots for the (A) quantitative micro-level measure of hair curvature and (B) observational macro-level measures of hair curvature. The Manhattan plots show the distribution of the −log 10 (*p* values) observed for the SNP associations by chromosome. QQ plots show the distribution of −log10 *p* values of the observed association and the genomic inflation factor (λ). The most significant associations are annotated. The bottom row of panels shows the regional plots for the (C) quantitative measure (C) and (D) observational measure of hair curvature. These figures show the position of rs12130862 and rs11803731 and the nearest genes to these two SNPs. The colors represent the degree of linkage disequilibrium (r^2) between each SNP and the sentinel quantitative measure SNP (rs12130862).
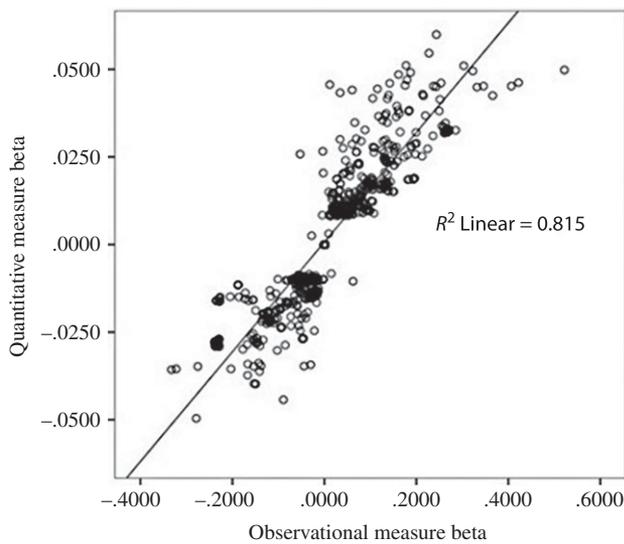

Fig. 4. Bivariate plot of allelic effect sizes (betas) for the quantitative micro-level measure on the *y*-axis and observational macro-level measure of hair curvature on the *x*-axis. *Note:* Correlation ($r^2 = .815$, $p < .01$).

are located within a region of approximately 400 Mb on chromosome 1q23.1. For the quantitative measure, there were no significant gene associations: the highest association was for *TCHH* ($p = 1.0 \times 10^{-5}$; Table 2, Figure 5).

The SNP rs11803731 is an exonic SNP that would result in the replacement of the amino acid lysine by methionine in the TCHH protein (p.L790M). This change is predicted to be 'damaging' by SIFT and 'probably damaging' by PolyPhen, and it has previously been suggested that rs11803731 could be the functional SNP at this locus (Medland, Nyholt et al., 2009). Functional analyses indicate that rs11803731 and SNPs in high LD with rs11803731 are also eQTLs that influence the expression of other nearby genes, including the hornerin (*HRNR*) and thioesterase superfamily member 4 (*THEM4*) genes in human skin (Figure 3C and 3D, Supplementary Tables S2 and S3).

Discussion

The curvature of human hair has been shown to be a highly heritable trait. While previous association studies have identified some genetic variants influencing hair curvature using observational macro-level measures, this is the first study to use quantitative micro-level measures of hair curvature.

We first reconfirmed the moderated cross-trait genetic correlation ($rg_A = .41$) between the quantitative and observational measures of hair curvature. The estimates of heritability for each measure (quantitative = 57%, observational = 95%) agree with previous estimations of the heritability of hair shape (Ho et al., 2016; Medland, Zhu et al., 2009).

The comparison of the quantitative and observational measures of hair curvature suggests that the quantitative assessment of

Table 2. Gene association for quantitative and observational measures of hair curvature

Gene	Location chr:bp	No. SNPs	p value	Top SNP	Top SNP p value
Quantitative measure					
TCHH	1:152,078,792	43	2.41E-05	rs55668963	9.88E-09
MIR6871	20:39,797,662	45	3.10E-05	rs118116797	1.37E-04
MRFAP1	4:6,641,817	161	3.32E-05	rs10937747	7.94E-05
PLCG1	20:39,766,160	58	5.07E-05	rs6129760	6.53E-05
CEP83	12:94,702,055	238	5.16E-05	rs1915335	5.32E-06
LOC101927117	20:39,726,633	56	5.81E-05	rs6129760	6.53E-05
LOC100132891	8:72,755,357	342	9.45E-05	rs9694605	5.59E-06
TCHHL1	1:152,056,619	43	1.08E-04	rs12130862	9.45E-09
FMO9P	1:166,573,152	118	1.10E-04	rs6670674	1.46E-05
MSC	8:72,753,776	120	1.12E-04	rs9694605	5.59E-06
Observational measure					
TCHH	1:152,078,792	43	3.24E-10	rs11803731	2.14E-17
S100A11	1:152,004,981	59	2.00E-09	rs12130862	6.32E-17
TCHHL1	1:152,056,619	43	3.38E-09	rs11803731	2.14E-17
NBPF18P	1:151,990,753	57	6.27E-09	rs12130862	6.32E-17
RPTN	1:152,126,070	65	6.79E-08	rs11803731	2.14E-17
S100A10	1:151,955,385	63	1.50E-07	rs34974291	2.67E-13
TCERG1L	10:132,890,654	713	1.98E-06	rs10765044	3.78E-05
PPDPF	20:62,152,132	209	1.07E-05	rs78117224	2.47E-06
SRMS	20:62,171,276	214	1.23E-05	rs78117224	2.47E-06
PTK6	20:62,159,775	221	1.31E-05	rs78117224	2.47E-06

Note: fastBat gene association analysis of quantitative micro-level and observational macro-level measures of hair curvature. Gene, location in chromosome (Chr) and starting position of the gene in base pairs (bp), number of SNPs associated with each gene, gene association *p* value and top SNP *p* value. SNP, single-nucleotide polymorphism.
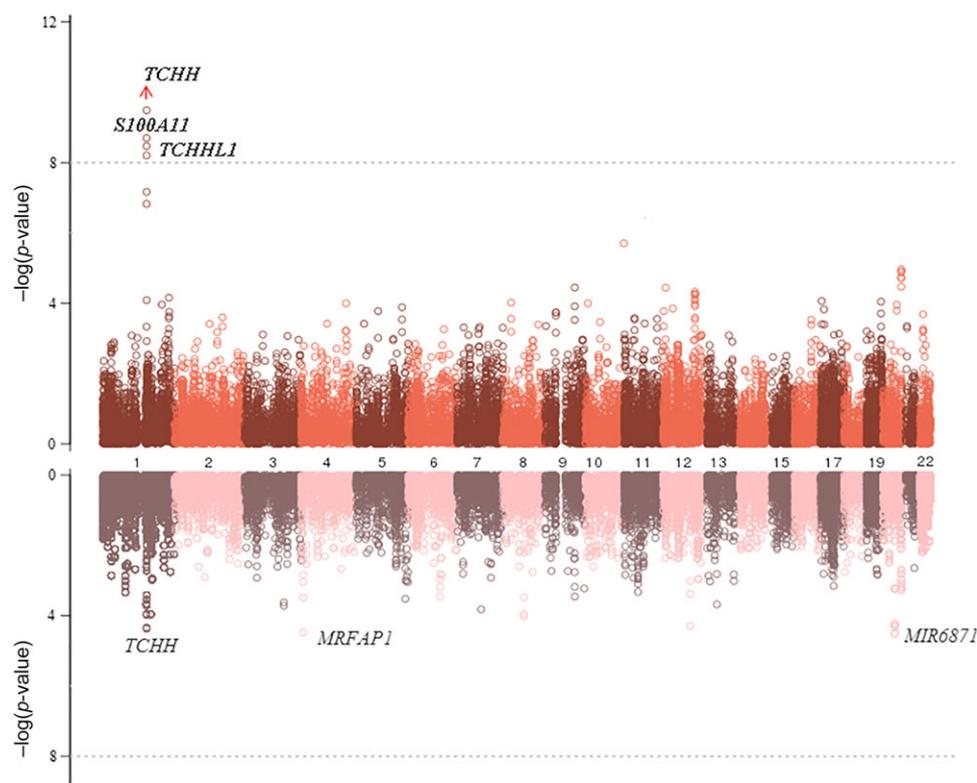

Fig. 5. Gene association analysis results for hair curvature. The top plot shows the gene association results for the observational macro-level measure and the bottom plot the gene association results for the quantitative micro-level measure of hair curvature. Plots show the location and association values for the top three associated genes for each data set.

curvature provided by the OFDA analysis is a good predictor of hair curvature categorized by observation alone. High correlation between the quantitative and observational effect sizes (Figure 4) shows these are predictive of each other, indicating that quantitative measures are comparable to observational measures in the detection of genetic variants predicting hair curvature.

Genome-wide association analyses for the two measures resulted in one genome-wide significant signal on chromosome 1q21.3. Association between variants at 1q21.3 and hair shape was previously reported by Medland, Nyholt et al. (2009) and reconfirmed by a meta-analysis that evaluated a quantitative measure of hair curvature in 28,964 individuals that included cohorts of non-European ancestry (Liu et al., 2018). In the current analysis, a set of five SNPs that includes rs11803731 and rs12130862, in high LD with each other and shown to represent the same signal in conditional GWAS analyses, surpassed the genome-wide significant threshold for both GWAS. These five SNPs are considered as the most likely candidate causal SNPs driving the association signal at the 1q21.3 locus, and are located in the region of the *TCHH* and *TCHHL1* genes. In the gene-based analyses, both genes were significantly associated with observational measures of hair curvature. *TCHH* and *TCHH1* encode the *TCCH* and *TCCH*-like 1 proteins, respectively. *TCCH* binds to keratins and is expressed in the inner root sheath of the hair follicle and in the hair strand. SNP rs11803731, a nonsynonymous SNP that replaces a lysine with a methionine in the TCHH protein, was previously associated with the straightness of hair (Medland, Nyholt et al., 2009). Our bioinformatic analysis also indicates rs11803731, rs12130862 and other SNPs in LD with these influence the expression of other nearby genes, including *HRNR* expressed in the epidermis and with a suggested role in keratinocyte differentiation (Henry et al., 2011). Further studies are, however, required to determine whether these additional genes have roles in influencing hair morphology, and *TCHH* remains the best candidate gene in this chromosomal region. In sheep, genetic variation in the TCHH gene, as well as differing mRNA expression levels over the lifetime, are known to influence wool straightness in various sheep breeds (Gong et al., 2018, Jie et al., 2019). Several studies of human hair have also indicated a role for TCHH in the biology and evolution of straight hair (Koch et al., 2019) and in the straightening of the hair follicle (Steinert et al., 2003, Wu et al., 2011).

The extent of support for the *TCHH* signal was stronger for the observational measure than for the quantitative measure of hair curvature. We hypothesize that this is due to either the difference between the macro (observational) and micro (quantitative) scales of measurement, or to the sampling of only a relatively small number of hairs taken from a single area on the head for the quantitative measure. Nevertheless, the quantitative measure GWAS did detect the signal previously observed at the *TCHH* locus, matching the results of the observational measures in the same individuals and serving as a positive control to reconfirm the accuracy of observational measures.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2020.78.

References

Adhikari, K., Fontanil, T., Cal, S., Mendoza-Revilla, J., Fuentes-Guajardo, M., Chacon-Duque, J. C., . . . Ruiz-Linares, A. (2016). A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. *Nature Communications*, 7.

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, 7, 248–249.

Bakshi, A., Zhu, Z., Vinkhuyzen, A. A., Hill, W. D., McRae, A. F., Visscher, P. M., & Yang, J. (2016). Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. *Science Reports*, 6, 32894.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., . . . Fox, J. (2011). OpenMx: an open source extended structural equation modeling framework. *Psychometrika*, 76, 306–317.

Eriksson, N., Macpherson, J. M., Tung, J. Y., Hon, L. S., Naughton, B., Saxonov, S., . . . Mountain, J. (2010). Web-based, participant-driven studies yield novel genetic associations for common traits. *PLOS Genetics*, 6, e1000993.

Feng, S., Liu, D., Zhan, X., Wing, M. K., & Abecasis, G. R. (2014). RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics*, 30, 2828–2829.

Fujimoto, A., Kimura, R., Ohashi, J., Omi, K., Yuliwulandari, R., Batubara, L., . . . Tokunaga, K. (2008). A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. *Human Molecular Genetics*, 17, 835–843.

Gong, H., Zhou, H., Tao, J, Li, S., Dyer, J. M., Luo, Y., & Hickford, J. G. H. (2018). Variation in the ovine trichohyalin gene and its association with wool curvature. *Small Ruminant Research*, 159, 1–4.

Henry J., Hsu, C-Y., Haftek, M, Nachat, R., de Koning, H. D., Gardinal-Galera, I., . . . Simon, M. (2011). Hornerin is a component of the epidermal cornified cell envelopes. *FASEB Journal*, 25, 1567–1576.

Ho, Y. Y., Brims, M., McNevin, D., Spector, T. D., Martin, N. G., & Medland, S. E. (2016). Variation and heritability in hair diameter and curvature in an Australian twin sample. *Twin Research and Human Genetics*, 19, 351–358.

Jie, Z., Yufang, L, Lu, M., & Ming, X. (2019). Genetic structure and expression pattern analysis of TCHH in Tan sheep. *China Animal Husbandry & Veterinary Medicine*, 46, 1585–1593.

Koch, S. L., Shriver, M. D., & Jablonski, N. G. (2019). Variation in human hair ultrastructure among three biogeographic populations. *Journal of Structural Biology*, 205, 60–66.

Lin, B. D., Mbarek, H., Willemsen, G., Dolan, C. V., Fedko, I. O., Abdellaoui, A., . . . Hottenga, J. J. (2015). Heritability and genome-wide association studies for hair color in a Dutch twin family based sample. *Genes*, 6, 559–576.

Liu, F., Chen, Y., Zhu, G., Hysi, P. G., Wu, S., Adhikari, K., . . . Kayser, M. (2018). Meta-analysis of genome-wide association studies identifies 8 novel loci involved in shape variation of human head hair. *Human Molecular Genetics*, 27, 559–575.

Liu, F., Wollstein, A., Hysi, P. G., Ankra-Badu, G. A., Spector, T. D., Park, D., . . . Kayser, M. (2010). Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genetics*, 6, e1000934

Loussouarn, G., Garcel, A. L., Lozano, I., Collaudin, C., Porter, C., Panhard, S., . . . de La Mettrie, R. (2007). Worldwide diversity of hair curliness: a new method of assessment. *International Journal of Dermatology*, 46, 2–6.

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., . . . Haplotype Reference Consortium. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, *48*, 1279–1283.

Medland, S. E., Nyholt, D. R., Painter, J. N., McEvoy, B. P., McRae, A. F., Zhu, G., . . . Martin, N. G. (2009). Common variants in the trichohyalin gene are associated with straight hair in Europeans. *American Journal of Human Genetics*, *85*, 750–755.

Medland, S. E., Zhu, G., & Martin, N. G. (2009). Estimating the heritability of hair curliness in twins of European ancestry. *Twin Research and Human Genetics*, *12*, 514–518.

Neale, M., & Cardon, L. (1992). *Methodology for genetic studies of twins and families*. Kluwer Academic Publishers.

Neale, M. C., Hunter, M. D., Pritkin, J., Zahery, M., Brick, T. R., Kirkpatrick, R. M., . . . Boker, S. M. (2016). OpenMx 2.0: extended structural equation and statistical modeling. *Psychometrika*, *81*, 535–549.

1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *26*, 68–74.

Pospiech, E., Chen, Y., Kukla-Bartoszek, M., Breslin, K., Aliferi, A., Andersen, J. D., . . . EUROFORGEN-NoE Consortium. (2018). Towards broadening Forensic DNA Phenotyping beyond pigmentation: improving the prediction of head hair shape from DNA. *Forensic Science International: Genetics*, *37*, 241–251.

Schneider, G., Hu, J., Sim, N.-L., Kumar, P., Henikoff, S., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research*, *40*, W452–W457.

Steinert, P. M., Parry, D. A., & Marekov, L. N. (2003). Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root shealth. *Journal of Biological Chemistry*, *278*, 41409–41419.

Tan, J., Yang, Y., Tang, K., Sabeti, P., Jin, L., & Wang, S. (2013). The adaptive variant EDARV370A is associated with straight hair in East Asians. *Human Genetics*, *132*, 1187–1191.

Thibaut, S., Barbarat, P., Leroy, F., & Bernard, B. A. (2007). Human hair keratin network and curvature. *International Journal of Dermatology*, *46*, 7–10.

Vogt A., McElwee K., & Blume-Peytavi U. (2008). Biology of the hair follicle. In U. Blume-Peytavi, D. A. Whiting & R. M. Trüeb (Eds.), *Hair growth and disorders* (pp. 1–20). Springer.

Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*, 1826.

Westgate, G. E., Ginger, R. S., & Green, M. R. (2017). The biology and genetics of curly hair. *Experimental Dermatology*, *26*, 483–490.

Wortmann, F., & Schwan-Jonczyk, A. (2006). Investigating hair properties relevant for hair 'handle'. Part I: hair diameter, bending and frictional properties. *International Journal of Cosmetic Science*, *28*, 61–68.

Wright, M., & Martin, N. G. (2004). Brisbane Adolescent Twin Study: outline of study methods and research projects. *Australian Journal of Psychology*, *56*, 65–78.

Wu, Z., Latendorf, T., Meyer-Hoffert, U., & Schröder, J.-M. (2011). Identification of Trichohyalin-like 1, an S100 fused-type protein selectively expressed in hair follicles. *Journal of Investigative Dermatology*, *131*, 1761–1763.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*, 76–82.