Estimating the Heritability of Hair Curliness in Twins of European Ancestry

Sarah E. Medland, Gu Zhu and Nicholas G. Martin

Queensland Institute of Medical Research, Brisbane, Australia

Recent studies in Asian populations have identified variants in the *EDAR* and *FGFR2* genes that arose following the divergence of Asians and Europeans and are associated with thick straight hair. To date no genetic variants have been identified influencing hair texture in Europeans. In the current study we examined the heritability of hair curliness in three unselected samples of predominantly European ancestry (N_{S1} = 2717; N_{S2} = 3904; N_{S3} = 5079). When rated using a three point scale (Straight/Wavy/Curly) males were ~5% more likely to report straight hair than females and there were suggestions in the data that curliness increased slightly with age. Across samples significant additive and dominant genetic influences were detected resulting in a broad sense heritability of 85–95%. Given the magnitude and the specificity of the *EDAR* effect on hair morphology in Asian populations we are hopeful that future association studies will detect similar genetic influences in European populations.

Keywords: hair texture, biometrics, population genetics

Hair morphology (which can be graded into eight categories of increasing curliness (Loussouarn et al., 2007) is one of the more conspicuous features of human variation; categories 1-II reflect the straight hair typical of East Asian populations, while categories VI-VIII reflect extremes of curliness typically found in African populations. In Caucasians, hair texture is highly variable with about 50% falling in categories 1-II ('straight'), 37% in III ('wavy') and 13% in IV ('curly'). There numerous hypotheses regarding the evolutionary significance of different hair textures, including heat insulation, translucence to allow dermal vitamin D synthesis, and sexual selection, but none with any empirical support as yet (Iyengar, 1998).

Biologically, the degree of curliness is correlated with the distribution of hair keratins and cell type within the hair fibre, with the number of mesocortical cells decreasing as the curl intensifies (Thibaut et al., 2007). Recent studies have identified Asian specific alleles of *EDAR* and *FGFR2* that are associated with thick straight hair, suggesting that these variants arose following the divergence of Asians and Europeans (Fujimoto et al., 2008; Fujimoto et al., 2009). It seems likely that parallel mutations may have evolved giving rise to straight hair in Europeans. However, to date no genetic variants have been identified influencing hair texture in Europeans. Here we examine the evidence for genetic effects on hair texture in three twin samples as a precursor to analyses attempting to indentify genes influencing this trait in individuals of European ancestry.

Methods

Participants

The three independent samples analyzed in the current study included one Adolescent and two Adult samples. The ascertainment and phenotypic data collection for these studies is summarized below. All three cohorts were drawn from unselected community samples. Characteristics of the three cohorts are summarized in Table 1.

Sample 1: Adolescent siblings. The data from this sample were collected in the context of an ongoing study (1992–2009) of melanoma risk factors in 12-year-old twins and their siblings. Participants were ascertained from schools in south-east Queensland. As part of a clinical protocol that has been described elsewhere blood was obtained for zygosity confirmation and DNA extraction (Zhu et al., 1999; Zhu et al., 2007), and research nurses recorded whether their hair was 'naturally straight, wavy or curly'. Most twins returned at age 14 and were examined with the same protocol. For 1,404 individuals curliness was recorded at ages 12 and 14 (Table 2; test–retest correlation of .89; SE .01). There was a slight increase in the prevalence of curliness with age. Where hair texture ratings differed at 12 and 14, the 14 year-old rating was used for genetic analyses since this is closer to adulthood. Data were available for a total of 2717 individuals (2082 twins and 635 non-twin siblings) from 1043 families.

Sample 2: Adult twin pairs. Twins from a second cohort born 1903–1964 were asked to donate blood

Received 21 July, 2009; accepted 14 August, 2009.

Address for correspondence: Sarah Medland, Queensland Institute of Medical Research, PO Royal Brisbane Hosptial, Herston Q 4029, Australia. Email: Sarah.medland@qimr.edu.au.

Table 1

Characteristics of the Three Samples

	Sample 1		Sample 2		Sample 3	
Birth years	1982–96		1903–81		1922–79	
Age mean (range)	14 (9–24)		47 (23–94)		39 (22–80)	
N Individuals	Females	Males	Females	Males	Females	Males
	1381	1336	2546	1358	2822	2257
% straight	63.4%	70.5%	70.1%	66.9%	49.3%	58.3%
% wavy	27.7%	21.4%	—	—	35.1%	30.2%
% curly	8.9%	8.1%	29.9%	33.1%	15.6%	11.6%

Table 2

Frequencies for 1404 Twins Whose Hair Texture Was Rated at Both 12 and 14 Years

Age 12 \ Age 14	Straight	Wavy	Curly	Total
Straight	889	154	0	1043
Wavy	65	170	41	276
Curly	0	14	71	85
Total	954	338	112	1404

samples in the context of a study on the genetics of alcohol related behaviors (Heath et al., 1997) in the years 1993–96, when they were aged 29–94. At this time they were asked by questionnaire 'Which of the following best describes your natural hair texture at age 20? Straight or curly?' Data were available for 3907 individuals from 2252 families.

Sample 3: Adult twin pairs. A third cohort of twins born 1965–1972 and with mean age 39 at time of survey, took part in telephone interview studies of gambling habits (Slutske et al., 2009) and male pattern baldness (Nyholt et al., 2003) in which participants were asked 'Which of the following best describes your *natural* hair texture at age 21? Straight, wavy or curly?' Data were available for 5017 individuals from 2895 families.

Statistical Methods

The data were analyzed using a multifactorial threshold model in which discrete traits are assumed to reflect an underlying normal distribution of liability (or predisposition). Liability, which represents the sum of all the multifactorial effects, is assumed to reflect the combined effects of a large number of genes and environmental factors each of small effect and is characterized by phenotypic discontinuities that occur when the liability reaches a given threshold (Neale & Cardon, 1992). All data analyses were conducted using maximum likelihood analyses of raw data within Mx (Neale et al., 2006).

Corrections for sex, age (both linear and quadratic effects), and age by sex interactions were included in the threshold models in all data analyses, such that the trait value for individual *j* from family *i* was parameterized as:

$$x_{ij} = \beta_{age} + \beta_{age^2} + \beta_{sex} + \beta_{sex*age} + \mu$$

The relatedness between the participants was explicitly modeled accounting for the sex of relative pairs. The phenotypic variances of males (σ^2_M) and females (σ^2_F), which were constrained to unity, were parameterized as:

$$\sigma^2_M = \sigma^2_{Am} + \sigma^2_{Dm} + \sigma^2_{Em} \text{ and, } \sigma^2_F = \sigma^2_{Af} + \sigma^2_{Df} + \sigma^2_{Ef}$$

where, σ^2_{Am} and σ^2_{Af} represent the additive genetic effects of males and females; σ^2_{Dm} and σ^2_{Df} represent the nonadditive genetic effects and σ^2_{Em} and σ^2_{Ef} represent the nonshared or unique environmental effects of males and females respectively. While the covariance terms were parameterized as:

$$Cov_{Male\ MZs} = \sigma^2_{Am} + \sigma^2_{Dm}$$

$$Cov_{Female\ MZs} = \sigma^2_{Af} + \sigma^2_{Df}$$

$$Cov_{Male\ DZ/Siblings} = .5\sigma^2_{Am} + .25\sigma^2_{Dm}$$

$$Cov_{Female\ DZ/Siblings} = .5\sigma^2_{Af} + .25\sigma^2_{Df}$$

$$Cov_{Opposite\ sex\ DZ/Siblings} = .5\sigma_{Am}\sigma_{Af} + .25\sigma_{Dm}\sigma_{Df}$$

The significance of variance components was tested by comparing the fit (minus twice log-likelihood) of the full model that included the effect to that of a nested model in which the effect had been dropped from the model. The difference in log-likelihoods follows an asymptotic chi-square distribution with the degrees of freedom equal to the difference in estimated parameters between the two models.

Results

Where curliness was measured using a 3-point scale males had a significantly higher prevalence of straight hair (Table 1, Table 3). However, in sample 2 where a two category measure was used more females than males reported straight hair. In the adolescent sample,

https://doi.org/10.1375/twin.12.5.514 Published online by Cambridge University Press

Table 3

Thresholds and Covariate Effects

	Sample 1	Sample 2	Sample 3
Threshold 1	.362	.548	–.003
Threshold 2	1.300	–	.995
βsex (M = 1, F = 0)	.145	–.106	.169
βage (Z score)*	–.042	–.043	–.004
βage^2	.005	–.005	.010
βsex*age	–.145	.074	–.112
Dropping βsex*age (χ^2_1)	17.34	147.33	4.35
Dropping βage, βage^2 & βsex*age (χ^2_3)	26.92	149.13	6.36
Dropping βsex (χ^2_1)	18.37	147.46	18.68

Note: Z scores calculated separately for the adolescents (Mean 13.92, sd 1.349) and the adults (samples 2 and 3 combined; Mean 42.5, sd 9.77). Note: the covariate effects reported here should be interpreted with reference to the liability threshold model. For example the first threshold (dividing straight hair from wavy hair) for a male aged 16 (zscore = 1.48) from sample 1 would be .362 + .145 + (1.48*–.042) + (2.16*.005) + (1.48*–.145) = .243 which equates to a prevalence of 59.6%.

there were also significant effects of age, age squared and age by sex (Table 3). In the adult samples, where participants were asked to report their hair texture at age 20 or 21, there were significant age effects were also seen (Table 3), although the magnitude of these effects differed between samples, and the interpretation of these effects is complicated by potential recall effects. However, as shown in Figure 1, there was a general trend for the prevalence of wavy and curly hair to increase with age in males but no clear effect of age in females.

As the adolescent sample was composed of twins and their non-twin siblings we compared the magnitude of the DZ twin and non-twin sibling correlations, allowing for differences between male, female, and opposite sex pairs. As no significant differences were observed (χ^2_3 = 3.2, p = .35) these correlations were equated in all further analyses. As shown in Table 4 significant co-twin correlations were observed across samples, with the relative magnitude of MZ and DZ correlations suggesting the presence of dominance.

For all three samples the ADE model provided a better fit (lower AIC) than the ACE model (Table 5). Across samples the majority of variance could be explained by genetic influences. In males these genetic effects were more dominant than additive, while the converse was true for females. However, given the high correlation between estimates of additive and dominance effects (r = –.97) the reliability of these estimates is low compared to the estimates of the broad sense heritability (Martin & Eaves, 1977). The concurrent analysis of the three samples revealed significant heterogeneity (χ^2_{12} = 100.8) primarily driven by the larger estimate of unique environmental effects in sample 3.

The estimates of AD and E could be equated for males and females in samples 1 and 2. Additive and dominant genetic effects accounted for the vast majority of the variance, yielding broad sense heritabilities of 98 and 96% in samples 1 and 2. Significant sex limitation was observed in sample 3 where, the majority of the variance in males was due to dominant genetic effects whereas in females additive genetic effects explained the majority of the variance. However, curliness was significantly heritable for both sexes. The broad sense heritability in sample 3 was 79% for the males and 87% for the females. Although lower than the other two samples, the magnitude of the heritability was similar to that of the repeatability of the measurement observed in sample 1 (.89).

Discussion

The current study of three independent twin samples has shown that hair curliness is strongly heritable in unselected samples of European ancestry (with a broad sense heritability ranging from .79 to .98 across samples). Across studies the majority of participants reported straight hair, with only a small percentage reporting truly curly hair. A comparison of the prevalence across studies suggests that when offered two response choices, straight or curly, participants who would otherwise have selected the 'wavy' category are divided fairly equally between the wavy and curly response choices. These findings are consistent with the findings of Loussouarn et al. (2007).

Hair texture was significantly heritable across the three samples, and all three samples showed significant non-additive genetic effects. While, the two adult studies did ask the participants to answer with refer-

Table 4

MZ and DZ Co-Twin Correlations (and 95% Confidence Intervals)

	Sample 1*		Sample 2		Sample 3	
	N twin pairs	*r* (95%)	*N* twin pairs	*r* (95%)	*N* twin pairs	*r* (95%)
MZ Male	170	.973 (.937, .990)	302	.925 (.850, .967)	351	.781 (.709, .837)
MZ Female	186	.996 (.987, .996)	671	.979 (.960, .986)	509	.870 (.834, .899)
DZ Male	184	.345 (.220, .394)	233	.309 (.059, .529)	250	.287 (.173, .398)
DZ Female	176	.380 (.266, .492)	482	.375 (.212, .522)	368	.378 (.254, .488)
DZ OS	318	.355 (.267, .435)	564	.440 (.295, .441)	413	.242 (.140, .339)

Note: *DZ twin and Twin-Sibling correlations were equated.

https://doi.org/10.1375/twin.12.5.514 Published online by Cambridge University Press

Table 5

Summary of the Variance Components Analyses for the 3 Samples

		Sample 1	Sample 2	Sample 3
ADE sex-limitation model	–2LL	3940.46	4253.49	9129.81
ACE sex-limitation model	–2LL	3952.45	4256.78	9138.20
Male A (95% CI)		.391 (.236,.924)	.588 (.122,.958)	.183 (.145,.784)
Male D (95% CI)		.581 (.307,.892)	.337 (.000,.810)	.601 (.000,.797)
Male E (95% CI)		.027 (.010,.061)	.075 (.033,.150)	.216 (.162,.286)
Female A (95% CI)		.520 (.275,.984)	.632 (.185,.987)	.686 (.094,.897)
Female D (95% CI)		.476 (.285,.869)	.348 (.000,.795)	.183 (.000,.780)
Female E (95% CI)		.004 (.001,.013)	.021 (.010,.040)	.130 (.101,.167)
Equating Male and Female VC	χ^2_3	5.953	5.813	9.48
	p-value	.114	.121	.024
A (95% CI)		.446 (.156,.729)	.605 (.211,.969)	—
D (95% CI)		.542 (.259,.833)	.362 (.001,.757)	—
E (95% CI)		.011 (.005,.023)	.033 (.019,.054)	—
AE model	χ^2_1	14.688	3.71	14.437
	p-value	.0001	.0541	.0001

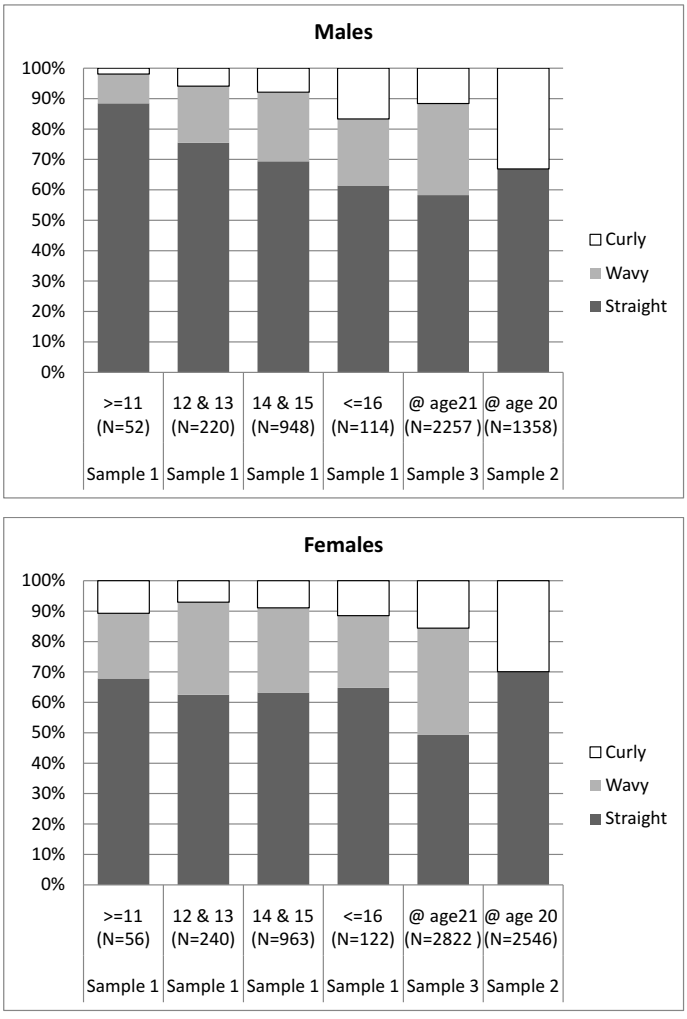

Figure 1

Prevalence of straight, wavy and curly hair by sex and age.

https://doi.org/10.1375/twin.12.5.514 Published online by Cambridge University Press

ence to their hair at ages 20 and 21 it is possible that this heterogeneity may reflect the low number of younger participants in sample 3. In addition, the pattern of stronger dominance effects in males, although not significantly different in samples 1 and 2, suggest that researchers working in this area should consider the possibility of sex by genotype interactions. When measured on a 3-point scale, females were more likely to have curly hair. Males appear to be more likely to develop curly hair as they age, this may potentially be due to differences in gene expression or hormonal regulation during the transition to adulthood, however the data used in the present study were predominantly cross sectional limiting our ability to examine this hypothesis.

In conclusion, the present study has shown hair curliness is significantly heritable with a broad sense heritability of ~85–95%. Given the magnitude and the specificity of the *EDAR* effect in hair morphology in Asian populations we are hopeful that future association studies will detect similar genetic influences in European populations.

Acknowledgments

We thank the twins and their families for their participation; Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney (sample collection); David Smyth, Harry Beeby, Daniel Park (IT support). Funding: Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498); Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921); FP-5 GenomEUtwin Project (QLG2-CT-2002-01254); U.S. National Institutes of Health (NIH grants AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206). S.E.M is supported by the NHMRC Fellowship Scheme (Sidney Sax 443036). The authors declare that they have no competing financial interests.

References

Fujimoto, A., Kimura, R., Ohashi, J., Omi, K., Yuliwulandari, R., Batubara, L., Mustofa, M. S., Samakkarn, U., Settheetham-Ishida, W., Ishida, T., Morishita, Y., Furusawa, T., Nakazawa, M., Ohtsuka, R., & Tokunaga, K. (2008). A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. *Human Molecular Genetics, 17*, 835–843.

Fujimoto, A., Nishida, N., Kimura, R., Miyagawa, T., Yuliwulandari, R., Batubara, L., Fujimoto, A., Kimura, R., Ohashi, J., Omi, K., Yuliwulandari, R., Batubara, L., Mustofa, M. S., Samakkarn, U., Settheetham-Ishida, W., Ishida, T., Morishita, Y., Tsunoda, T., Tokunaga, K., & Ohashi, J. (2009). FGFR2 is associated with hair thickness in Asian populations. *Journal of Human Genetics*, *54*, 461–465.

Heath, A. C., Bucholz, K. K., Madden, P. A., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., Statham, D. J., Dunne, M. P., Whitfield, J. B., & Martin, N. G. (1997). Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychological Medicine, 27*, 1381–1396.

Iyengar, B. (1998). The Hair Follicle: A specialised UV Recepter in the Human Skin. *Biological Signals and Receptors, 7*, 188–194.

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

Martin, N., & Eaves, L. (1977). The genetical analysis of covariance structure. *Heredity 38*, 79–95.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2006). *Mx: Statistical Modeling* (6th Edition ed.). VCU Box 900126, Richmond, VA 23298 *http://www.vcu.edu/mx/:* Department of Psychiatry.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Nyholt, D. R., Gillespie, N. A., Heath, A. C., & Martin, N. G. (2003). Genetic basis of male pattern baldness. *Journal of Investigative Dermatology, 121*, 1561–1564.

Slutske, W. S., Meier, M. H., Zhu, G., Statham, D. J., Blaszczynski, A., & Martin, N. G. (2009). The Australian Twin Study of Gambling (OZ-GAM): rationale, sample description, predictors of participation, and a first look at sources of individual differences in gambling involvement. *Twin Research and Human Genetics, 12*, 63–78.

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

Zhu, G., Duffy, D. L., Eldridge, A., Grace, M., Mayne, C., O'Gorman, L., Aitken, J. F., Neale, M. C., Hayward, N. K., Green, A. C., & Martin, N. G. (1999). A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: A maximum-likelihood combined linkage and association analysis in twins and their sibs. *American Journal of Human Genetics, 65*, 483–492.

Zhu, G., Montgomery, G. W., James, M. R., Trent, J. M., Hayward, N. K., Martin, N. G., & Duffy, D. L. (2007). A genome-wide scan for naevus count: linkage to CDKN2A and to other chromosome regions. *European Journal of Human Genetics, 15*, 94–102.