**Twin Research and Human Genetics**

**Supplementary materials**

**The Genetic Overlap Between Hair and Eye Color**

Results of a genome-wide association study for eye color in the Netherlands Twin Register (NTR)

Bochao D. Lin, Gonneke Willemsen, Abdel Abdellaoui, Meike Bartels, Erik A. Ehli, Gareth E. Davies, Dorret I. Boomsma, and Jouke J. Hottenga

Netherlands Twin Register, Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands

**Participants**

Within the NTR, there were 7063 Dutch ancestry participants, clustered in 3407 families (2641 men, age: 44.99±19.15; 4546 women, age: 45.15±16.72) with information on both genotype and eye color. Data on eye color were obtained from surveys. Participants reported their eye color by choosing from one of three answer possibilities: “blue/grey", "green/hazel" and "brown". Written informed consent was obtained from all participants.

**Genotyping data**

Buccal or blood DNA samples were collected for multiple NTR projects. DNA extraction and purification of these samples were performed at various points in time (Boomsma et al., 2008; Willemsen et al., 2010), following several manufacturer specific protocols to obtain the best quality and concentration prior to SNP platform genotyping. Quality control was done within and between platforms and subsets prior to imputation. For each platform, the individual SNP markers were lifted over to build 37 (HG19) of the Human reference genome, using the LiftOver tool. SNPs that were not mapped at all, SNPs that had ambiguous locations, and SNPs that did not have matching (or strand opposite alleles) were removed. Samples were excluded from the data if their expected sex did not match their genotyped sex, if the genotype missing rate was above 10% or if the Plink F inbreeding value was either >0.10 or <−0.10. Quality control details were detailed in Lin et al. (2015).

Phasing of all samples and imputing cross-missing platform SNPs was done with MACH 1 (Li et al., 2010). The phased data were then imputed with MINIMAC (Howie et al., 2012) against the 1000 Genomes Phase 1 Reference panel in batches of around 500 individuals for 561 chromosome chunks obtained by the CHUNKCHROMOSOME program (Liu et al., 2013). After imputation, SNPs were filtered based on the Mendelian error rate (> 2%) in families. If the imputed allele frequency differed more than 0.15 from the 1000G reference allele frequency, SNPs were removed.

**GWA Analysis**

We performed 3 case-control GWAS on binary eye color variables: brown versus non-brown eye color, blue/grey versus non-blue/grey eye color, green/hazel versus non-green/hazel eye color, with logistic regression, having age, sex, 3 Dutch PCs and corrections for genotype platform as covariates. Analyses were performed with the PLINK 1.07 software running a logistic regression on each SNP, taking genotype inaccuracy into account by using dosage data (Purcell et al., 2007). Familial structure was taken into account using "--family" option. Filtering on MAF > 0.01, imputation quality R2> 0.80, and Hardy-Weinberg equilibrium (HWE) p-values>0.0001 were done after the GWA analyses within all eye color informative individuals. This left 5,834,593 SNPs for generating Manhattan and QQ plot.

**Results**

The eye color prevalence was 62.28% for blue/grey eyes, 19.58% for green/light eyes and 18.14% for brown eyes. The top SNPs in a LD block for each eye color are shown in Table 1. The resulting Q-Q and Manhattan plots for all eye colors are shown in the supplemental figures (S1–S3).

For brown eye color, we found that the presence of the T allele at rs74940492, an intron variant for HERC2, significantly decreases the probability of brown eye color (OR=0.09, p=5.4E-8). This SNP is in the same LD block with top SNP rs2240202 for blue eye color (OR=13.55, p=1.0E-47). This locus was also found to be associated with blond hair color, brown hair color and dark hair color in our study. HERC2, which harbors this SNP, has been identified as an eye iris color gene by multiple studies (Kayser et al., 2008; Sturm et al., 2008).

The top SNP rs4904871, an intron genetic variant in SLC24A4, was significantly associated with both blue eye color (OR=0.71, p=2.8E-13) and green eye color (OR=1.52, p=3.8E-20). This SNP has also been associated with hair color in GWA studies (Lin et al., 2015). The rs12896399, which is located in the same LD block of rs4904871 (LD r2=0.95 distance~22kb), was found to be associated with blond versus brown hair color, blue versus green eye color (Sulem et al., 2007) and black versus blond hair color (Han et al., 2008) in other GWA studies.

The T allele of rs67279079 at TYR has been found to significantly decrease the probability of blue eye color (OR=0.70, p=3.1E-11), and simultaneously to increase the probability of green/hazel eye color (OR=1.49, p=3.6E-10). The TYR gene codes tyrosinase located in melanocyte, which is responsible for the first step in melanin production. This gene is associated oculocutaneous albinism and skin tanning ability (J. Liu et al., 2010; Morice-Picard et al., 2014; Zhang et al., 2013). Three known pigment genetic loci were thus confirmed, but no new genetic variants for eye color were identified.

**Conclusion**

In this study, we have replicated genetic variants for eye color: HERC2 for brown eye color and blue/grey eye color; TYR and SLC24A4 for blue/grey and green/hazel eye color.

Table 1: SNP Associations for Three Eye Colors

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Locus | Chr  location | Top SNPs | MAF | Brown eye color | | | Blue eye color | | | Green eye color | | |
| OR | SE | p | OR | SE | p | OR | SE | p |
| HERC2 | 15q13 | rs2240202 | 0.036 | 0.09 | 0.123 | **5.4E-83** | 14.53 | 0.189 | **2.5E-45** | 0.89 | 0.128 | 0.343 |
| HERC2 | 15q13 | rs74940492 | 0.037 | 0.10 | 0.123 | **2.2E-81** | 13.55 | 0.179 | **1.0E-47** | 0.89 | 0.128 | 0.3655 |
| SLC24A4 | 14q32.12 | rs4904871 | 0.478 | 1.09 | 0.056 | 0.1218 | 0.71 | 0.046 | **1.3E-13** | 1.52 | 0.053 | **1.5E-15** |
| TYR | 11q14.3 | rs67279079 | 0.236 | 1.17 | 0.066 | 0.0209 | 0.70 | 0.054 | **3.1E-11** | 1.49 | 0.064 | **3.6E-10** |

Figure 1: Manhattan and QQ plot for brown eye color (MAF > 0.01, λ =1.03114).

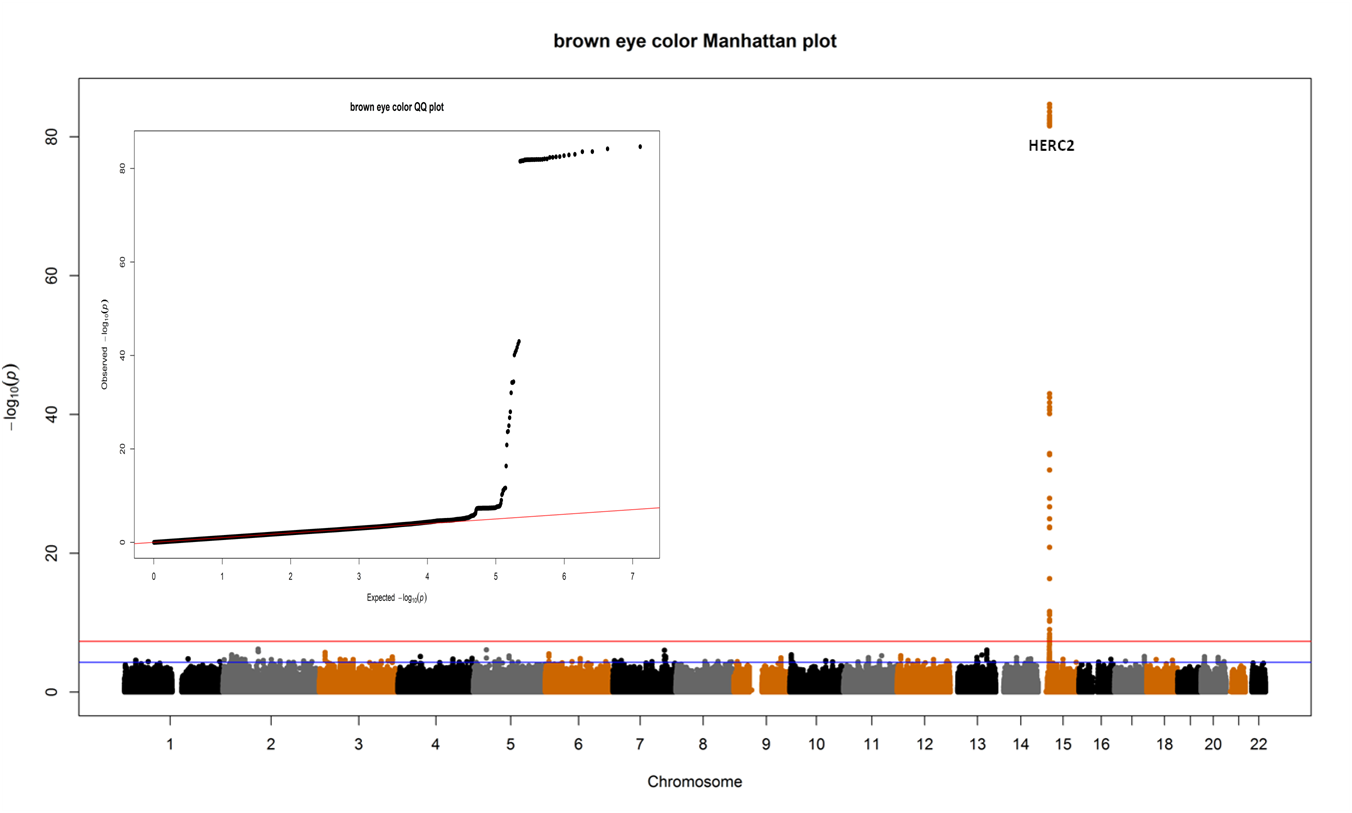


Figure 2: Manhattan and QQ plot for blue/grey eye color (MAF > 0.01, λ =1.0259).

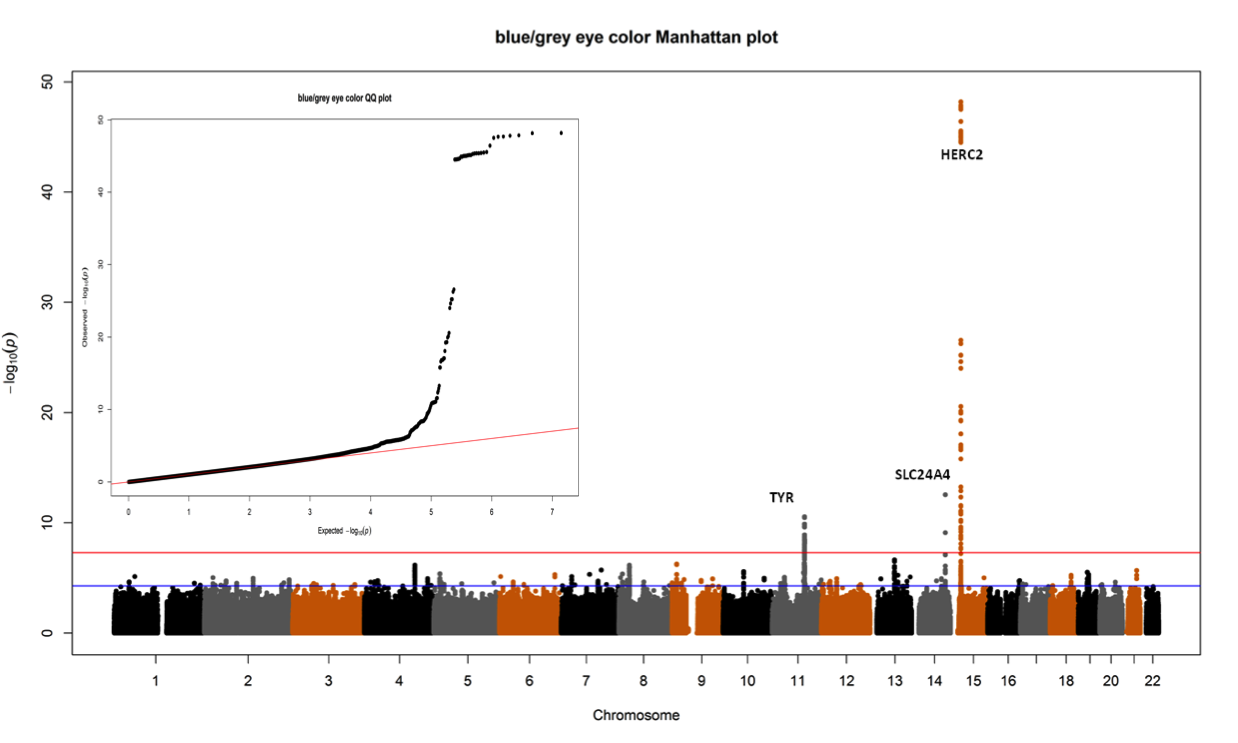
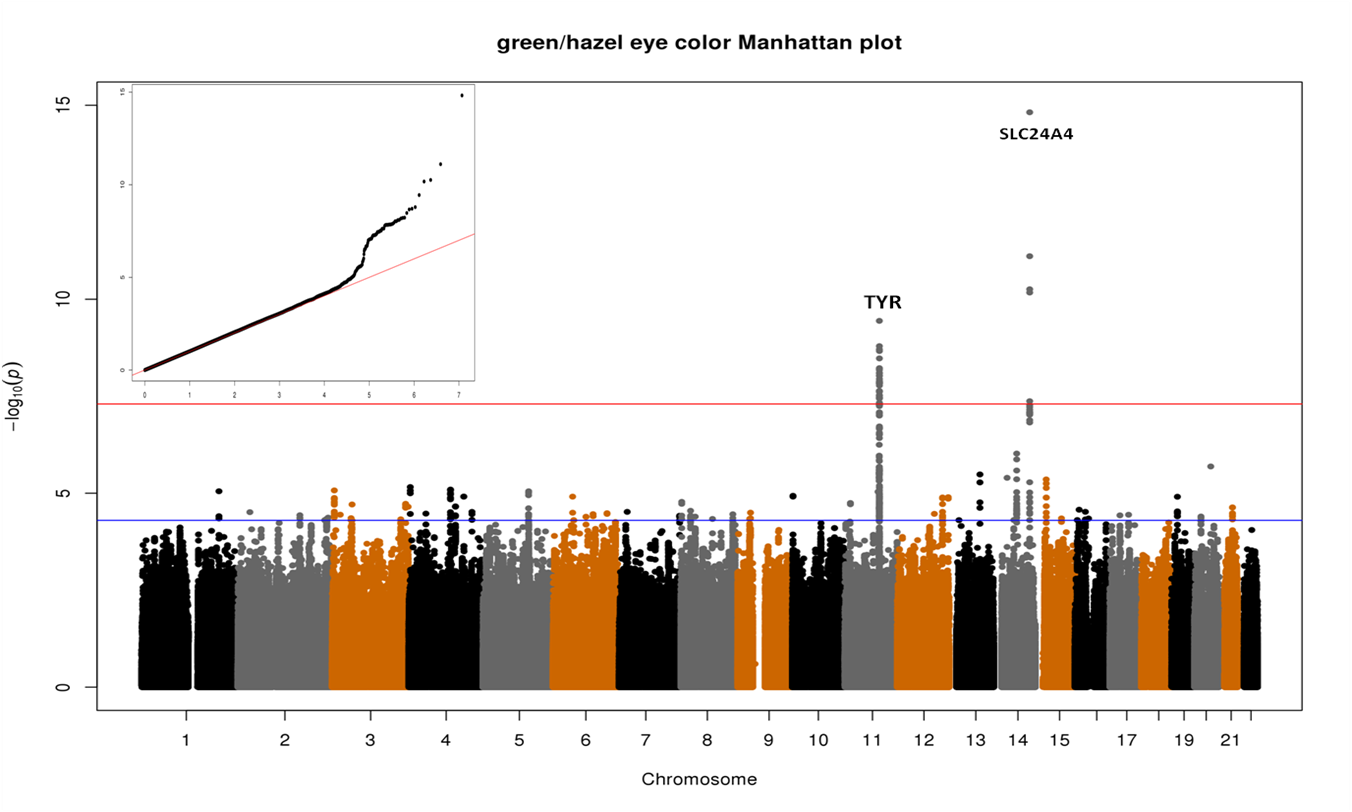


Figure 3: Manhattan and QQ plot for green/hazel eye color (MAF > 0.01, λ = 1.02163).



Boomsma, D. I., Willemsen, G., Sullivan, P. F., Heutink, P., Meijer, P., & Sondervan, D. (2008). Genome-wide association of major depression: Description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *European Journal of Human Genetics, 16*, 335–342.

Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C., Qureshi, A., … Hunter, D. J. (2008). A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *Plos Genetics, 4*, e1000074.

Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics, 44*, 955-959.

Kayser, M., Liu, F., Janssens, A. C., Rivadeneira, F., Lao, O., van Duijn, K., … van Duijn, C. M. (2008). Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *American Journal of Human Genetics, 82*, 411–423.

Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010). MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology, 34*, 816–834.

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, E. Y., Li, M., Wang, W., & Li, Y. (2013). MaCH-admix: Genotype imputation for admixed populations. *Genetic Epidemiology, 37*, 25–37.

Liu, J., Choy, K. W., Chan, L. W. L., Leung, T. Y., Tam, P. O. S., S. … Lai, T. Y. Y. (2010). Tyrosinase gene (TYR) mutations in Chinese patients with oculocutaneous albinism type 1. *Clinical and Experimental Ophthalmology, 38*, 37–42.

Morice-Picard, F., Lasseaux, E., Cailley, D., Gros, A., Toutain, J., Plaisant, C., … Arveiler, B. (2014). High-resolution array-CGH in patients with oculocutaneous albinism identifies new deletions of the TYR, OCA2, and SLC45A2 genes and a complex rearrangement of the OCA2 gene. *Pigment Cell & Melanoma Research, 27*, 59–71.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*, 559–575.

Sturm, R. A., Duffy, D. L., Zhao, Z. Z., Leite, F. P. N., Stark, M. S., Hayward, N. K., … Montgomery, G. W. (2008). A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *American Journal of Human Genetics, 82*, 424–431.

Sulem, P., Gudbjartsson, D. F., Stacey, S. N., Helgason, A., Rafnar, T., Magnusson, K. P., … Stefansson, K.. (2007). Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics, 39*, 1443–1452.

Willemsen, G., de Geus, E. J., Bartels, M., van Beijsterveldt, C. E., Brooks, A. I., & Estourgie-van Burk, G. F. (2010). The Netherlands Twin Register biobank: A resource for genetic epidemiological studies. *Twin Research and Human Genetics, 13*, 231–245.

Zhang, M. F., Song, F. J., Liang, L. M., Nan, H. M., Zhang, J. W., … Han, J. L. (2013). Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Human Molecular Genetics, 22*, 2948–2959.