Association Study Confirms the Role of Two OCA2 Polymorphisms in Normal Skin Pigmentation Variation in East Asian Populations

KATHERINE EATON,1 MELISSA EDWARDS,1 S. KRITHIKA,1 GILLIAN COOK,1 HEATHER NORTON,2 AND ESTEBAN J. PARRA1,3
1Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada
2Department of Anthropology, University of Cincinnati, Cincinnati, Ohio

Objectives: The main goal of the study was to test the association of 18 polymorphisms located within nine pigmentation candidate genes with quantitative skin pigmentation measures collected in a sample of individuals of East Asian ancestry living in Canada (N = 419).

Methods: The 18 polymorphisms are located within genes that show putative signatures of positive selection in East Asian populations. The genetic markers were selected for genotyping based on an annotation of common East Asian polymorphisms to predict potential functional effects. We restricted our attention to polymorphisms that have an allele frequency difference of at least 30% between East Asian populations and African and European populations, or have alleles that are present in East Asians, but are absent in Africans and Europeans.

Results: Two nonsynonymous variants selected within the OCA2 gene, rs1800414 (His615Arg) and rs74653330 (Ala481Thr), were significantly associated with melanin levels in the sample. Both single nucleotide polymorphisms (SNPs) are nonsynonymous polymorphisms located more than 30 kb apart on chromosome 15 and have very different frequencies in the East Asian sample. Additionally, both polymorphisms are predicted to have a deleterious effect on the protein. Linear regression analysis using an additive model indicate that each copy of the derived rs1800414 allele G decreases Melanin Index approximately 0.9 units and each copy of the derived rs74653330 allele A decreases Melanin Index approximately 1.9 units.

Conclusions: Two nonsynonymous OCA2 polymorphisms (rs1800414 and rs74653330) are independently associated with normal skin pigmentation variation in East Asian populations and have very different frequency distributions in East Asia. Am. J. Hum. Biol. 00:000–000, 2015. © 2015 Wiley Periodicals, Inc.
Herein, we present the results of the aforementioned study of individuals of East Asian ancestry living in Canada. Quantitative skin pigmentation measures collected in a sample of East Asian populations, we selected 18 polymorphisms for genotyping and evaluated the association of these loci with quantitative skin pigmentation measures collected in a sample of individuals of East Asian ancestry living in Canada. Present, the results of the aforementioned association study.

MATERIALS AND METHODS

Study population

Participants for this research were recruited at the University of Toronto, Mississauga Campus (UTM) from 2013 to 2014. The study cohort primarily consisted of students and staff of UTM who replied to advertisements that were distributed throughout the University of Toronto community online and in print. Individuals were asked to complete a questionnaire to evaluate their geographic ancestry through their parents and grandparents place of birth and native languages. A total of 461 participants reported their ancestors to be from East Asia. Of these, the majority of the participants were of Chinese ancestry (N = 325), with a much smaller proportion reporting ancestry from Korea (N = 69), Taiwan (N = 9), Japan (N = 7), and Vietnam (N = 24). A number of participants (N = 27) indicated ancestry from more than one East Asian country (e.g., China and Vietnam, China and Taiwan, Japan and Korea, etc.). For the statistical analysis, to minimize the potential impact of genetic heterogeneity, we only included individuals reporting ancestry from China, Taiwan, Korea, and Japan, or those indicating mixed ancestry from any of these four countries. The final sample included 423 individuals in total, with 284 female subjects and 139 male subjects. The age of the participants ranged between 19 and 34 years.

Ethics statement

Written informed consent was obtained from each participant and the study was approved by the University of Toronto Health Sciences Research Ethics Board.

Measurement of melanin levels

Melanin content was quantitatively measured from each individual’s inner arm with a handheld DSM II colormeter (Cortex Technology, Handsund, Denmark). This instrument offers measurements in different color systems. For our study, we reported melanin levels using the Melanin Index (M) which is a widely used measure in pigmentation studies (higher M levels correspond to higher melanin content in the skin). Measurements were carried out during the winter to minimize the influence of summer UV radiation. For statistical analysis, we used the average M value based on three independent measurements. Melanin index values were also measured with the Dermaspec, an older colormeter from Cortex Technology. The melanin index values obtained with both instruments were highly correlated (r = 0.788, P < 0.001).

SNP selection

We recently carried out a study aimed at identifying putative signatures of positive selection in pigmentation candidate genes in populations of East Asian ancestry (Hider et al., 2013). For the genes showing putative signatures of selection in that study, we carried out detailed annotations of the common variants (frequencies >1%) present in the 1,000 Genomes Project (1KGP) Phase 1

RESULTS

DNA extraction and genotyping

A saliva sample was collected from each participant using Oragene DNA OG500 extraction Kits and DNA was extracted from the saliva samples using the protocols recommended by the manufacturer. Genotyping of samples was performed by LGC genomics (http://www.lgcgenomics.com, accessed September 2, 2014) using KASP genotyping chemistry. Genotyping quality was assessed through the inclusion of blind duplicates in the genotyping plates. The concordance rate for the genotype calls of the blind duplicate samples was 100%.

Statistical analysis

Statistical analysis was performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/, accessed September 2, 2014). The following analyses were carried out with this program: (1) Tests of deviations from Hardy–Weinberg; (2) Quantification of missing genotypes; (3) Linear regression analysis to evaluate the effect of each SNP on melanin level, using sex as covariate. For the linear regression analysis, we used additive tests, which report the effect of each extra minor allele, but also carried out additional analyses using a genotypic model in which the effects of heterozygotes and one of the homozygotes are reported with reference to the other homozygote; (4) Linear regression analysis using sex as covariate, and conditioning on one of the SNPs included in the analysis; and (5) Analysis of Linkage Disequilibrium (LD) between pairs of markers, reported as r^2 values. During the QC steps prior to statistical analysis, we excluded four individuals with genotype missingness rates higher than 15%, bringing the final sample to 419 individuals.
Index or M) in a sample of individuals of East Asian ancestry living in Canada. Figure 1 shows the distribution of M values in the sample. No significant pigmentation differences were observed between males and females in this sample (Average M value-males = 38.05, Average M value-females = 37.84, P = 0.485). Table 1 shows the genotype and allele frequencies observed for the 18 markers genotyped in the study. No significant deviations from Hardy–Weinberg proportions were observed for any of the loci. We used the program PLINK to evaluate the potential association of the 18 markers with M values, using sex as a covariate. The results of this linear regression using an additive model are reported in Table 2. The polymorphism rs1800414 located within the OCA2 gene reached significance after Bonferroni correction for multiple tests (P < 0.0028) (G allele, beta = -0.914, P = 1.37 × 10⁻⁵). Another polymorphism located within the OCA2 gene, rs74653330, was nominally significant (A allele, beta = -1.291, P = 0.025). Given that these two polymorphisms are located within the same gene, we carried out further LD and conditional regression analyses. These two polymorphisms have very different allele frequencies (rs1800414, G-allele frequency = 0.602, rs74653330, A-allele frequency = 0.033) and we observed an r² value between these markers of 0.052, indicating that these markers are not in LD. After conditioning for rs1800414, the P-value of rs74653330 decreases substantially, surpassing the Bonferroni’s threshold of significance (A allele, beta = -1.873, P = 0.0011). Figure 2 shows the distribution of skin pigmentation (expressed as melanin index) stratifying by rs1800414 and rs74653330 genotypes. None of the other variants included in this study showed nominally significant effects (P < 0.05) in the conditional analysis (data not shown).

DISCUSSION

We carried out an association study of 18 polymorphisms located within nine candidate pigmentation genes with skin pigmentation measured quantitatively in an East Asian sample. These nine genes show putative signatures of positive selection in East Asian populations (Hider et al., 2013), and genetic variants were selected for genotyping based on an annotation of common polymorphisms present in East Asian populations to predict potential functional effects. We restricted our attention to polymorphisms that have an allele frequency difference of at least 30% between East Asian populations and African and European populations, or have alleles that are present in the East Asians, but are absent in Africans and Europeans. Two of the variants selected within the OCA2 gene, rs1800414 and rs74653330, were significantly associated with melanin levels in the sample. Both SNPs are nonsynonymous polymorphisms, and are located more than 30 kb apart on chromosome 15. The frequencies of both variants are very different in our sample and a conditional analysis indicates that they have an independent effect on skin pigmentation.

The SNP rs1800414 is a nonsynonymous polymorphism (His615Arg) that is present in high frequencies in East Asians. The frequency of the allele associated with lower melanin levels (allele G) is 60.2% in our East Asian sample and ranges between 53% and 63% in the East Asian HapMap samples. This allele is not present in African or European populations. This SNP is classified as deleterious by the annotation portal SNVrap (http://jjwanglab.org/snvrap/, accessed September 2, 2014), which provides functional prediction scores that are based on multiple annotation sources.

The SNP rs1800414 has been associated with melanin levels in two previous studies focused on East Asian populations (Abe et al., 2013; Edwards et al., 2010). Our linear regression analysis using an additive model estimated that each copy of the G allele decreases the Melanin Index (obtained with the DSM II colorimeter) approximately 0.9 units. Using the Dermaspec, we obtained similar results (each copy of the G allele decreases Melanin Index approximately 0.75 units, P = 1.96 × 10⁻³). This result is in good agreement with the estimates that were obtained in a previous study using the Dermaspec, which were 1.2 units in the original sample and 0.85 units in the replication sample (Edwards et al., 2010). Using an unconstrained genotypic model, we estimated that homozygotes GG are approximately 1.93 melanin units lighter (P = 1.15 × 10⁻⁵) and heterozygotes 0.73 melanin units lighter (P = 0.020) than homozygotes for the ancestral allele A.

The SNP rs74653330 is another nonsynonymous polymorphism (Ala481Thr) located within the OCA2 gene. In our sample, the frequency of the minor allele (allele A) is 3.3% and we did not observe any homozygote for the A allele. Yuasa et al. (2011) studied the distribution of this variant in several East Asian populations and found that the frequency of the derived A allele ranged between 0.0% and 7.4% in Han Chinese, Japanese, and Korean populations. However, the A allele reached much higher frequencies in Mongolia and nearby regions (from 13% in the Khalkha from Ulaan Baator to 51.85% in the Oroqen from the Heilongjiang province in Northern China). By contrast, the A allele was not observed in European, South Asian, and African samples (Yuasa et al., 2011).

The SNP rs74653330 is classified as highly deleterious by SNVrap, and according to our conditional regression analysis, has a stronger effect on melanin levels than rs1800414. The decrease in Melanin Index associated with each copy of the A allele is approximately 1.9 units (after conditioning for rs1800414). This variant was


<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr (pos)</th>
<th>SNP (alleles)</th>
<th>Genotype frequencies</th>
<th>Allele frequencies</th>
<th>HW p</th>
</tr>
</thead>
<tbody>
<tr>
<td>LYST</td>
<td>1 (235910542)</td>
<td>rs3754234 (C/T)</td>
<td>C: 379, C: 40, C: 0</td>
<td>C = 0.953, C = 0.048</td>
<td>0.614</td>
</tr>
<tr>
<td>LYST</td>
<td>1 (235976651)</td>
<td>rs7522053 (T/C)</td>
<td>T: 271, T: 39, T: 0</td>
<td>T = 0.953, T = 0.047</td>
<td>1.000</td>
</tr>
<tr>
<td>LYST</td>
<td>1 (236031001)</td>
<td>rs4659610 (A/G)</td>
<td>A: 153, A: 39, A: 0</td>
<td>A = 0.818, A = 0.182</td>
<td>0.069</td>
</tr>
<tr>
<td>MLPH</td>
<td>2 (238434249)</td>
<td>rs2292881 (C/T)</td>
<td>C: 260, C: 143, C: 13</td>
<td>C = 0.797, C = 0.203</td>
<td>0.288</td>
</tr>
<tr>
<td>OPRM1</td>
<td>6 (154360797)</td>
<td>rs179997 (A/G)</td>
<td>A: 195, A: 47, A: 0</td>
<td>A = 0.651, A = 0.349</td>
<td>0.517</td>
</tr>
<tr>
<td>OPRM1</td>
<td>6 (154721557)</td>
<td>rs6917661 (C/T)</td>
<td>T: 181, T: 61, T: 0</td>
<td>T = 0.363</td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>7 (55229255)</td>
<td>rs2227983 (A/G)</td>
<td>A: 138, A: 195, A: 0</td>
<td>A = 0.565, A = 0.435</td>
<td>0.320</td>
</tr>
<tr>
<td>BNC2</td>
<td>9 (16407821)</td>
<td>rs9406647 (C/A)</td>
<td>C: 108, C: 195, C: 0</td>
<td>C = 0.504, C = 0.496</td>
<td>0.490</td>
</tr>
<tr>
<td>BNC2</td>
<td>9 (16435848)</td>
<td>rs3739714 (A/G)</td>
<td>A: 195, A: 47, A: 0</td>
<td>A = 0.933, A = 0.067</td>
<td>0.418</td>
</tr>
<tr>
<td>BNC2</td>
<td>9 (16673620)</td>
<td>rs10756778 (C/T)</td>
<td>C: 149, C: 30, C: 0</td>
<td>C = 0.750, C = 0.250</td>
<td>0.299</td>
</tr>
<tr>
<td>BNC2</td>
<td>9 (16786281)</td>
<td>rs10962591 (G/A)</td>
<td>A: 202, A: 164, A: 0</td>
<td>A = 0.363</td>
<td></td>
</tr>
<tr>
<td>TH</td>
<td>11 (2197148)</td>
<td>rs4930046 (A/G)</td>
<td>A: 207, A: 270, A: 0</td>
<td>A = 0.804, A = 0.196</td>
<td>0.348</td>
</tr>
<tr>
<td>OCA2</td>
<td>15 (28197037)</td>
<td>rs1800414 (G/A)</td>
<td>G: 148, G: 124, G: 19</td>
<td>G = 0.602, G = 0.196</td>
<td>0.541</td>
</tr>
<tr>
<td>OCA2</td>
<td>15 (28228553)</td>
<td>rs74653330 (G/A)</td>
<td>A: 28, A: 63, A: 0</td>
<td>A = 0.967, A = 0.033</td>
<td>1.000</td>
</tr>
<tr>
<td>OCA2</td>
<td>15 (28344328)</td>
<td>rs7497270 (T/C)</td>
<td>T: 185, T: 184, T: 0</td>
<td>T = 0.666, T = 0.334</td>
<td>0.912</td>
</tr>
<tr>
<td>TRPM1</td>
<td>15 (31394537)</td>
<td>rs3809578 (T/C)</td>
<td>T: 265, T: 133, T: 0</td>
<td>T = 0.799, T = 0.201</td>
<td>1.000</td>
</tr>
<tr>
<td>MCIR</td>
<td>16 (89986025)</td>
<td>rs33962559 (T/C)</td>
<td>C: 17, C: 269, C: 0</td>
<td>C = 0.944, C = 0.056</td>
<td>0.631</td>
</tr>
<tr>
<td>MCIR</td>
<td>16 (89986154)</td>
<td>rs885479 (A/G)</td>
<td>A: 166, A: 178, A: 0</td>
<td>A = 0.622, A = 0.378</td>
<td>0.116</td>
</tr>
</tbody>
</table>

*Ancestral alleles are indicated in bold.*

Previously associated with melanin levels in a Japanese study (Abe et al., 2013). It is important to note that both OCA2 variants (rs1800414 and rs74653330) have very different allele frequencies and are not in LD in our East Asian sample. Based on conditional analyses, both SNPs have independent effects on skin pigmentation (e.g., considering on rs1800414, the polymorphism rs74653330 has a significant effect on melanin index, \( P = 0.0011 \), and conversely, considering on rs74653330, the polymorphism rs1800414 remains significant, with a lower \( P \)-value than in the original analysis, \( \beta = 1.05, \ P = 7.45 \times 10^{-10} \)).

In principle, the presence of cryptic genetic structure in the sample could potentially have an effect on our results. However, several lines of evidence indicate that this is not the case: (1) There are no significant deviations from Hardy–Weinberg proportions in any of the markers studied. The presence of genetic structure would be reflected in an excess of homozygotes in the sample; (2) Data for East Asian populations (Yuasa et al., 2011) indicate that there are no large frequency differences for the relevant markers (rs1800414 and rs74653330) in East Asian populations including Han Chinese from different geographic areas, Japanese and Koreans. An exception is the region around Mongolia, which is not represented in our study; (3) There are no significant differences in melanin levels between the individuals of our sample classified according to geographic origin (China, Taiwan, Japan, and Korea, \( P = 0.263 \); (4) Both markers (rs1800414 and rs74653330) have been previously...
associated with pigmentation in studies in East Asia, including samples from Japan and Han Chinese from Fudan; and (5) Both markers are nonsynonymous and predicted to be deleterious by SNVrap.

It is of interest to compare the effects of the two OCA2 polymorphisms with those of other variants associated with normal pigmentation variation in human populations. Two of the most relevant SNPs are SLC24A5 rs1426654 and SLC45A2 rs16891982, which are the polymorphisms with the strongest effects on melanin levels described in humans (Beleza et al., 2013; Lamason et al., 2005; Marcheco-Teruel et al., 2014; Norton et al., 2007). The derived alleles of rs1426654 (allele A) and rs16891982 (allele G) have reached extremely high frequencies in some European populations through a recent selective sweep (Beleza et al., 2013; Lamason et al., 2005; Hider et al., 2013). The distribution of allele frequencies in East Asian populations (Donelly et al., 2012; Norton et al., 2007) indicates that these two variants may have been favored by selection in different regions. The derived allele of OCA2 rs1800414 has reached very high frequencies in Chinese, Japanese, and Korean populations. On the other side, the derived allele of OCA2 rs74653330, which has a stronger effect on pigmentation levels than rs1800414, has the highest frequencies in East Asian populations living at the highest latitudes (northern China and Mongolia). Neither of these variants is present in African or European populations.

![Fig. 2. Boxplots showing the distribution of skin pigmentation (expressed as melanin index values) stratifying by rs1800414 and rs74653330 genotypes. (A) Boxplots for rs1800414. (B) Boxplots for rs74653330. The boxplots present five statistics: the top of the box represents the 75th percentile, the line within the box represents the median, and the bottom of the box represents the 25th percentile, while the whiskers correspond to the minimum and maximum values that are not outliers. We also show the mean melanin values for each genotype as numbers located on top of each Boxplot.](image_url)

Both of them are predicted to have deleterious effects on the protein. It seems clear that OCA2 has been one of the major targets of selection driving the reduction in melanin levels in East Asian populations (Donelly et al., 2012; Hider et al., 2013). The distribution of allele frequencies of rs1800414 and rs74653330 in East Asia indicate that these two variants may have been favored by selection in different regions. The derived allele of OCA2 rs1800414 has reached very high frequencies in Chinese, Japanese, and Korean populations. On the other side, the derived allele of OCA2 rs74653330, which has a stronger effect on pigmentation levels than rs1800414, has the highest frequencies in East Asian populations living at the highest latitudes (northern China and Mongolia). Neither of these variants is present in African or European populations.
suggesting that these mutations arose in East Asia, and then increased in frequency in different geographic areas. There is evidence indicating that the OCA2 gene has also been under positive selection in European populations but different variants have been favored in East Asian and European populations (Donnelly et al., 2012; Edwards et al., 2010). Our knowledge about the polymorphisms involved in normal pigmentation variation in non-European populations is still in its infancy, and further studies focused on these populations are needed to get a global perspective of the genetic architecture and evolutionary history of pigmentation in our species.

ACKNOWLEDGMENTS

The authors would like to thank all the individuals who participated in this study.

REFERENCES


