Forensic Science International: Genetics Supplement Series 8 (2022) 82–84

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/fsigss

Association between copy number variations in the *OCA2-HERC2* locus and human eye colour

Nina Mjølsnes Salvo [a,*], Kirstin Janssen [a], Gunn-Hege Olsen [a], Thomas Berg [a], Jeppe Dyrberg Andersen [b]

[a] Centre for Forensic Genetics, Department of Medical Biology, Faculty of Health Sciences, UiT – The Arctic University of Norway, Norway
[b] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

ARTICLE INFO

Keywords:
Next generation sequencing
Copy number variation
OCA2-HERC2
Eye colour
Forensic genetics
Forensic DNA phenotyping

ABSTRACT

Human eye colour variation is strongly associated with single nucleotide polymorphisms (SNPs) in the *OCA2-HERC2* locus, especially rs12913832 that is found in an enhancer element of *OCA2*. In a previous study we found that 43 out of 166 individuals in a Norwegian population with the brown eye colour genotype *HERC2* rs12913832:AA or AG, did not have the expected brown eye colour. To investigate if duplications or deletions in the *OCA2-HERC2* locus could explain the blue eye colour in these individuals, we analysed massively parallel sequencing (MPS) data for copy number variations (CNVs) in the *OCA2-HERC2* region. The ~500 kb long *OCA2-HERC2* locus was sequenced in 94 individuals with the rs12913832:AG and AA genotypes. Of these, 43 were observed to have blue eye colour and 51 were observed to have brown eye colour. CNVs were analysed using R and the R-package *panelcn.MOPS* - CNV detection tool for targeted NGS panel data. In rs12913832:AG individuals, CNVs in 32 regions were significantly associated with blue eye colour (Benjamini-Hochberg adjusted *p*-value ≤ 0.05). In rs12913832:AA individuals, CNVs in 14 regions were associated with blue eye colour using raw *p*-values ($p \leq 0.05$). The functional effects of these CNVs on *OCA2* expression are yet to be investigated. However, this study suggests that CNVs in the *OCA2-HERC2* locus might explain why some of the rs12913832:AG and AA individuals have unexpectedly blue eyes.

1. Introduction

Single nucleotide polymorphisms (SNPs) in the *OCA2-HERC2* locus have been extensively studied to explain human eye colour variations. The main predictor for eye colour is the SNP rs12913832 in *HERC2* [1]. This SNP is located ~21kb upstream to the promoter of the pigmentation gene *OCA2*, and acts as a distal *OCA2* enhancer [2]. Individuals with the genotypes rs12913832:AA and AG are predicted to have brown eye colour, whereas individuals with rs12913832:GG are predicted to have blue eye colour [3]. However, we found that 43/166 individuals in a Norwegian study population with the genotype rs12913832:AA or AG did not have brown eye colour [4].

The aim for this study was to analyse massively parallel sequencing (MPS) data for copy number variations (CNV) to investigate if duplications or deletions in the *OCA2-HERC2* locus could explain the unexpected blue eye colour in rs12913832:AA and AG individuals.

2. Materials and methods

The *OCA2-HERC2* locus (~500kb) was sequenced in 94 *HERC2* rs12913832:AG and AA individuals with SureSelectXT HS2 Target Enrichment System (Agilent Technologies) on an Illumina MiSeq with paired-end sequencing (2×150bp) using the MiSeq Reagent Kit V2 (300 cycles). Individuals were grouped into a control group with observed brown eye colour (21 AA and 30 AG individuals) and a case group with observed blue eye colour (three AA and 40 AG individuals). All samples were collected with fully informed consent and subsequently anonymised. The project was approved by the Faculty of Health Sciences, UiT - The Arctic University of Norway (reference number 2021/2034).

All samples were aligned to the GRCh38 human genome assembly using Burrows-Wheeler Aligner, BWA-MEM algorithm [5,6]. SAM files were converted into BAM files using SAMtools [7]. CNVs were detected in bins of 100 bp with 50 bp overlap, using R and the R-package *panelcn.MOPS* – CNV detection tool for targeted NGS panel data, with default

* Correspondence to: UiT – The Arctic University of Norway, Post Box 6050, 9037 Tromsø, Norway.
E-mail address: nina.mjolsnes@uit.no (N.M. Salvo).

https://doi.org/10.1016/j.fsigss.2022.09.030
Received 16 September 2022; Accepted 28 September 2022
Available online 29 September 2022

Table 1
CNVs detected in each CN class for selected regions that were significantly associated (adjusted p-value ≤ 0.05) with eye colour in rs12913832:AG individuals. The estimated CN classes are based on expected fold changes in RC: 0.025 (CN0), 0.57 (CN1), 1 (CN2), 1.46 (CN3) and 2 (CN4). All control samples (brown) were set to CN2.

Gene Intron[a]	Start	End	#Homozygous deletions (CN0)		#Heterozygous deletions (CN1)		#CN2		#Duplications (CN3)		#Amplifications (CN4)		Adjusted *p*-value value[b]	LowQual[c]
			Blue	Brown	Blue	Brown	Blue	Brown	Blue	Brown	Blue	Brown		
OCA2 23	27760520	27760619	0	0	13	0	25	30	2	0	0	0	5.0E-02	27/40
	27787649	27787748	0	0	13	0	26	30	1	0	0	0	5.0E-02	23/40
	27787699	27787798	0	0	13	0	25	30	2	0	0	0	4.8E-02	23/40
	27787799	27787898	1	0	11	0	25	30	3	0	0	0	4.9E-02	26/40
	27798376	27798475	0	0	18	0	22	30	0	0	0	0	4.9E-03	12/40
	27798426	27798525	0	0	14	0	24	30	2	0	0	0	2.8E-02	11/40
	27839862	27839961	0	0	18	0	20	30	2	0	0	0	2.1E-03	24/40
	27839912	27840011	0	0	18	0	18	30	4	0	0	0	9.2E-04	33/40
	27839962	27840061	0	0	16	0	23	30	1	0	0	0	1.2E-02	29/40
	27840012	27840111	0	0	17	0	21	30	2	0	0	0	4.4E-03	36/40
	27840062	27840161	0	0	15	0	23	30	2	0	0	0	1.5E-02	30/40
	27840162	27840261	0	0	14	0	26	30	0	0	0	0	4.6E-02	11/40
	27840212	27840311	0	0	15	0	25	30	0	0	0	0	2.9E-02	15/40
	27840312	27840411	0	0	15	0	25	30	0	0	0	0	2.7E-02	1/40
OCA2 21	27858553	27858652	0	0	16	0	24	30	0	0	0	0	1.9E-02	31/40
	27858653	27858752	0	0	15	0	24	30	1	0	0	0	2.1E-02	39/40
	27858703	27858802	0	0	14	0	25	30	1	0	0	0	3.3E-02	40/40
	27858753	27858852	0	0	14	0	26	30	0	0	0	0	4.5E-02	40/40
OCA2 19	27873420	27873519	0	0	14	0	26	30	0	0	0	0	4.3E-02	0/40
	27892556	27892655	0	0	13	0	26	30	1	0	0	0	4.8E-02	8/40
	27892606	27892705	0	0	13	0	26	30	1	0	0	0	4.7E-02	5/40
OCA2 14	27976556	27976655	0	0	11	0	25	30	4	0	0	0	4.7E-02	21/40
OCA2 2	28042549	28042670	0	0	3	0	23	30	12	0	2	0	2.6E-02	32/40
	28067438	28067559	0	0	18	0	21	30	1	0	0	0	4.6E-03	7/40
HERC2 82	28125690	28125789	0	0	14	0	26	30	0	0	0	0	4.1E-02	0/40

[a] Intron annotation from GRCh38, chr.15, NM_000275.3 (*OCA2*) and NP_000266.2 (*HERC2*), NCBI.
[b] Benjamini-Hochberg adjusted *p*-value.
[c] Number of test samples flagged for low quality by *panelcn.MOPS*.

settings [8]. The *panelcn.MOPS* assign all case samples to a CN class per region of interest based on fold changes in read counts (RC) relative to control samples: CN0 (homozygous deletion), CN1 (heterozygous deletion), CN2 (no change), CN3 (heterozygous duplication) and CN4 (homozygous duplication). All control samples were set to have two copies throughout the investigated region (CN2). Association with eye colour was tested using Fisher's exact test.

3. Results

When comparing CNV frequencies between the blue and brown eye colour category, CNVs in 32 regions were statistically significantly associated with blue eye colour in rs12913832:AG individuals (Benjamini-Hochberg adjusted *p*-value ≤ 0.05). A total of 25 of these regions were considered as candidate regions to explain eye colour variation as they were located in *OCA2* or close to the *OCA2* enhancer SNP, rs12913832 in *HERC2* (Table 1). Seven CNVs were in *HERC2,* ~100 kb upstream rs12913832. Notably, some regions in rs12913832:AG individuals, especially in *OCA2* intron 21 and 23, were flagged for low quality (possibly low read counts) in a substantial proportion of the samples (Table 1). Thus, results should be interpreted with care.

In rs12913832:AA individuals, CNVs in 14 regions were associated with blue eye colour using raw *p*-values ($p \leq 0.05$). Deletions were observed in two of these regions in *OCA2* (intron 23), whereas both deletions and duplications were observed in 12 regions in *HERC2,* ~100 kb upstream rs12913832.

4. Discussion

With current prediction models, blue and brown eye colours are predicted with high accuracies [3]. However, blue eye colour has frequently been observed in individuals with the non-blue rs12913832: AG and AA genotypes in Scandinavian populations [9]. *OCA2* SNPs such as rs1800407, rs75643330 and rs121918166 can potentially explain this lighter eye colour in some, but not all of these individuals [9,10]. In this study we observed intronic duplications and deletions in the *OCA2--HERC2* locus in blue-eyed individuals, including regions in *OCA2* intron 2 (upstream the missense mutations rs1800407, rs121918166 and rs75643330), and regions in *HERC2* intron 82 (upstream rs12913832). It is well known that CNVs can have significant impact on human traits, but copy-number-based genome-wide association studies (GWASs) from MPS data is scarce. One large-scale association study found CNV to be associated with hair colour at *HERC2* in a British biobank [11]. To the best of our knowledge, we suggest for the first time CNVs in these regions to be candidates to explain eye colour variation in rs12913832:AG or AA individuals with unexpected blue eye colour. However, other studies must confirm the results and it is yet to be functionally investigated if these structural variants have regulatory effects on *OCA2*.

5. Conclusion

Preliminary results from MPS data revealed CNVs in the *OCA2-HERC2* region in blue-eyed rs12913832:AG and rs12913832:AA individuals. Despite the small sample size, we identified candidate regions with significantly more deletions or duplications in blue-eyed compared to brown-eyed individuals. These could potentially explain the unexpected blue eyes in these individuals. However, the functional effects of these CNVs on *OCA2* expression are yet to be investigated.

Declaration of Competing Interest

None.

Acknowledgements

The authors would like to thank all participants. The project was funded by UiT-The Arctic University of Norway.

References

[1] R.A. Sturm, D.L. Duffy, Z.Z. Zhao, F.P.N. Leite, M.S. Stark, N.K. Hayward, N. G. Martin, G.W. Montgomery, A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color, Am. J. Hum. Genet. 82 (2008) 424, https://doi.org/10.1016/J.AJHG.2007.11.005.

[2] M. Visser, M. Kayser, R.-J. Palstra, HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter, Genome Res. 22 (2012) 446–455, https://doi.org/10.1101/gr.128652.111.

[3] S. Walsh, A. Wollstein, F. Liu, U. Chakravarthy, M. Rahu, J.H. Seland, G. Soubrane, L. Tomazzoli, F. Topouzis, J.R. Vingerling, J. Vioque, A.E. Fletcher, K. N. Ballantyne, M. Kayser, DNA-based eye colour prediction across Europe with the IrisPlex system, Forensic Sci. Int. Genet. 6 (2012) 330–340, https://doi.org/10.1016/J.FSIGEN.2011.07.009.

[4] N.M. Salvo, K. Janssen, M.K. Kirsebom, O.S. Meyer, T. Berg, G.H. Olsen, Predicting eye and hair colour in a Norwegian population using Verogen's ForenSeq™ DNA signature prep kit, Forensic Sci. Int. Genet. 56 (2022), 102620, https://doi.org/10.1016/J.FSIGEN.2021.102620.

[5] H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows–Wheeler transform, Bioinformatics 26 (2010) 589–595, https://doi.org/10.1093/BIOINFORMATICS/BTP698.

[6] H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, 2013. ⟨http://github.com/lh3/bwa⟩ (Accessed 31 January 2022).

[7] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform, Bioinformatics 25 (2009) 1754, https://doi.org/10.1093/BIOINFORMATICS/BTP324.

[8] G. Povysil, A. Tzika, J. Vogt, V. Haunschmid, L. Messiaen, J. Zschocke, G. Klambauer, S. Hochreiter, K. Wimmer, panelcn.MOPS: copy-number detection in targeted NGS panel data for clinical diagnostics, Hum. Mutat. 38 (2017) 889, https://doi.org/10.1002/HUMU.23237.

[9] O.S. Meyer, N.M. Salvo, A. Kjærbye, M. Kjersem, M.M. Andersen, E. Sørensen, H. Ullum, K. Janssen, N. Morling, C. Børsting, G.-H. Olsen, J.D. Andersen, Prediction of eye colour in scandinavians using the eyecolour 11 (EC11) SNP set, Genes 12 (2021) 821, https://doi.org/10.3390/GENES12060821.

[10] J.D. Andersen, C. Pietroni, P. Johansen, M.M. Andersen, V. Pereira, C. Børsting, N. Morling, Importance of nonsynonymous OCA2 variants in human eye color prediction, Mol. Genet. Genom. Med. 4 (2016) 420–430, https://doi.org/10.1002/mgg3.213.

[11] T. Fitzgerald, E. Birney, CNest: A novel copy number association discovery method uncovers 862 new associations from 200,629 whole-exome sequence datasets in the UK Biobank, Cell Genom. 2 (2022), 100167, https://doi.org/10.1016/J.XGEN.2022.100167.