

The official journal of

INTERNATIONAL FEDERATION OF PIGMENT CELL SOCIETIES · SOCIETY FOR MELANOMA RESEARCH

# PIGMENT CELL & MELANOMA Research

## Investigating DNA methylation as a potential mediator between pigmentation genes, pigmentary traits and skin cancer

Carolina Bonilla | Bernardo Bertoni | Josine L. Min |  
Gibran Hemani | Genetics of DNA Methylation Consortium |  
Hannah R. Elliott

**DOI: 10.1111/pcmr.12948**

**Volume 34, Issue 5, Pages 892–904**

If you wish to order reprints of this article,  
please see the guidelines [here](#)

Supporting Information for this article is freely available [here](#)

### EMAIL ALERTS

Receive free email alerts and stay up-to-date on what is published  
in Pigment Cell & Melanoma Research – [click here](#)

Submit your next paper to PCMR online at <http://mc.manuscriptcentral.com/pcmr>

Subscribe to PCMR and stay up-to-date with the only journal committed to publishing  
basic research in melanoma and pigment cell biology

As a member of the IFPCS or the SMR you automatically get online access to PCMR. Sign up as  
a member today at [www.ifpcs.org](http://www.ifpcs.org) or at [www.societymelanomaresarch.org](http://www.societymelanomaresarch.org)


To take out a personal subscription, please [click here](#)

More information about Pigment Cell & Melanoma Research at [www.pigment.org](http://www.pigment.org)



## ORIGINAL ARTICLE

# Investigating DNA methylation as a potential mediator between pigmentation genes, pigmentary traits and skin cancer

Carolina Bonilla<sup>1,2</sup>  | Bernardo Bertoni<sup>3</sup> | Josine L. Min<sup>2,4</sup> | Gibran Hemani<sup>2,4</sup> | Genetics of DNA Methylation Consortium | Hannah R. Elliott<sup>2,4</sup>

<sup>1</sup>Departamento de Medicina Preventiva, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil

<sup>2</sup>Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

<sup>3</sup>Departamento de Genética, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay

<sup>4</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

## Correspondence

Carolina Bonilla, Departamento de Medicina Preventiva, Faculdade de Medicina, Universidade de São Paulo, Av. Dr. Arnaldo 455, CEP 01246-903 São Paulo, Brazil. Email: cxbonilla@usp.br

## Funding information

23andme; Wellcome Trust, Grant/Award Number: 088806, 208806/Z/17/Z and 217065/Z/19/Z; Universidade de São Paulo, Grant/Award Number: 88887.160006/2017-00; Biotechnology and Biological Sciences Research Council, Grant/Award Number: BB/1025263/1 and BBI025751/1; Royal Society, Grant/Award Number: 208806/Z/17/Z; Medical Research Council, Grant/Award Number: MC-UU\_12013/2, MC\_UU\_00011/5, MC\_UU\_12013/1 and MC\_UU\_12013/8; University of Bristol; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Grant/Award Number: 88887.160006/2017-00

## Abstract

Pigmentation characteristics are well-known risk factors for skin cancer. Polymorphisms in pigmentation genes have been associated with these traits and with the risk of malignancy. However, the functional relationship between genetic variation and disease is still unclear. This study aims to assess whether pigmentation SNPs are associated with pigmentary traits and skin cancer via DNA methylation (DNAm). Using a meta-GWAS of whole-blood DNAm from 36 European cohorts ( $N = 27,750$ ; the Genetics of DNA Methylation Consortium, GoDMC), we found that 19 out of 27 SNPs in 10 pigmentation genes were associated with 391 DNAm sites across 30 genomic regions. We examined the effect of 25 selected DNAm sites on pigmentation traits, sun exposure phenotypes and skin cancer and on gene expression in whole blood. We uncovered an association of DNAm site cg07402062 with red hair in the Avon Longitudinal Study of Parents and Children (ALSPAC). We also found that the expression of *ASIP* and *CDK10* was associated with hair colour, melanoma and basal cell carcinoma. Our results indicate that DNAm and expression of pigmentation genes may play a role as potential mediators of the relationship between genetic variants, pigmentation phenotypes and skin cancer and thus deserve further scrutiny.

## KEYWORDS

ALSPAC, *ASIP*, DNA methylation, GoDMC, *MC1R*, pigmentation, skin cancer

## 1 | INTRODUCTION

The incidence of skin cancer, comprising malignant melanoma, basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), has increased rapidly in the past decades (Cahoon et al., 2015; Hu et al., 2009; Rees et al., 2014). Melanoma is the most aggressive of

skin cancers although it has a low incidence, while non-melanoma skin cancer (NMSC) shows high incidence yet considerably lower mortality rates compared to melanoma (Apalla et al., 2017). To date, recognized risk factors for melanoma and NMSC include fair skin, light-coloured eyes, red hair, freckles and melanocytic naevi as well as genetic variants, some of which underlie these skin pigmentation

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 John Wiley & Sons A/S. Published by John Wiley & Sons Ltd

and sun sensitivity phenotypes (Gordon, 2013). Even though there are missense and nonsense polymorphisms in pigmentation genes strongly associated with skin cancer, particularly in the gene *MC1R* (Nasti & Timares, 2015), it is not fully understood how non-coding variation in these genes relates to malignancy.

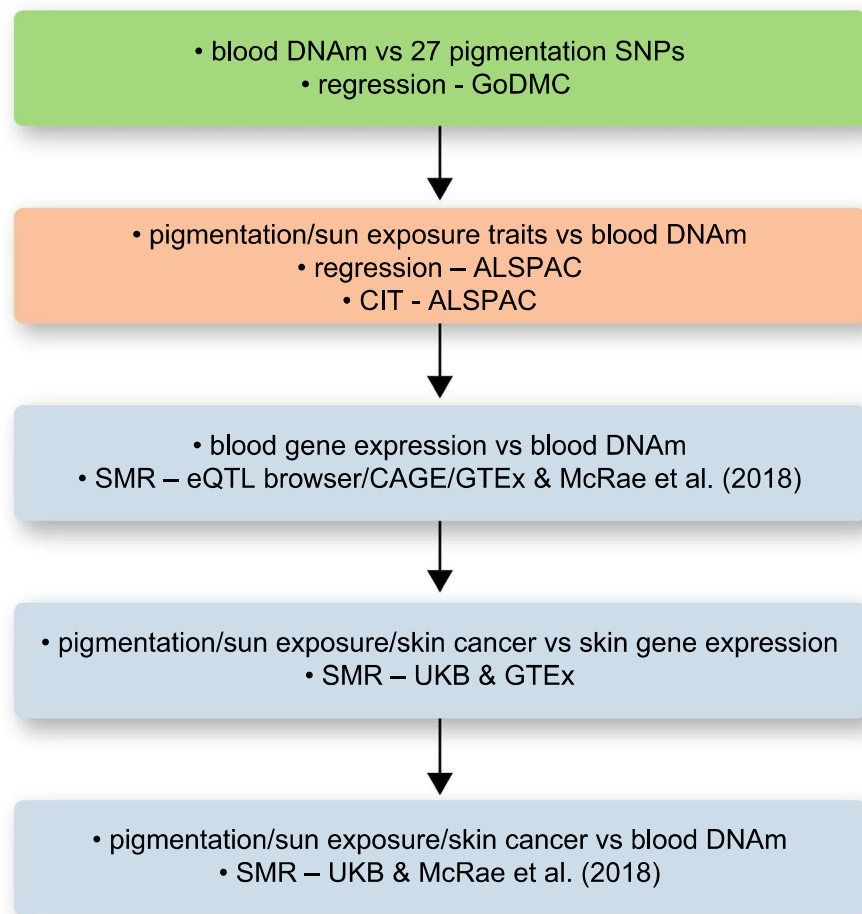
DNA methylation (DNAm) is an epigenetic modification with a prospective role in cancer aetiology. The association of global whole-blood DNA hypomethylation with cancer is well known and has also been described for melanoma (Cappetta et al., 2015; Shen et al., 2017). In addition, recent studies have shown the presence of hypomethylation in skin biopsy samples that were exposed to sunlight or artificial ultraviolet radiation (Grönniger et al., 2010; Holzschek et al., 2020; Vandiver et al., 2015). Inter-individual DNAm variation at specific sites, measured in peripheral blood, has been uncovered as a predictor of a number of complex trait risk factors as well as all-cause mortality (McCartney et al., 2018). DNAm has the potential to be used as a molecular biomarker to diagnose disease or to assess prognosis in those affected by disease and is a promising candidate leading towards the realization of personalized medicine (Nikolouzakis et al., 2020).

This study investigated whether DNAm plays a role in the relationship between genetic variants, pigmentation-related skin cancer risk factors and skin cancer. First, we examined the effect of genetic variation on DNAm levels in peripheral blood across 10 regions

### Significance

Several genetic variants in pigmentation-related genes have been associated with skin cancer, particularly in populations of European ancestry. DNA methylation was investigated as a potential link between pigmentation SNPs, pigmentation phenotypes and skin cancer. DNA methylation sites near *MC1R* and *ASIP* were strongly associated with pigmentation SNPs, gene expression, hair colour, melanoma and basal cell carcinoma in an analysis that included data from the Genetics of DNA Methylation consortium, the Avon Longitudinal Study of Parents and Children and the UK Biobank. Further evaluation of the role played by DNA methylation in these relationships is warranted.

robustly associated with pigmentation traits and skin cancer, in 36 cohorts of European descent. We then analysed the association of pigmentation SNP-associated DNAm sites with sun exposure and pigmentation phenotypes in participants of the Avon Longitudinal Study of Parents and Children (ALSPAC). Finally, using summary data-based Mendelian randomization (SMR), we explored whether SNP-related DNAm was likely to causally underlie the expression



**FIGURE 1** Stages of analysis implemented and datasets used in this study. From the 391 unique DNA methylation (DNAm) sites associated with pigmentation SNPs, 25 were selected for further analysis as depicted. GoDMC, Genetics of DNA methylation consortium; ALSPAC, Avon Longitudinal Study of Parents and Children; CIT, causal inference test; SMR, summary data-based Mendelian randomization; eQTL browser, blood eQTL browser; CAGE, Cap Analysis of Gene Expression; GTEEx, Genotype-Tissue Expression consortium; UKB, UK Biobank

of genes associated with pigmentation phenotypes and skin cancer. The purpose of this work was to carry out an integrative analysis to evaluate the joint contribution of genetic and epigenetic risk factors to skin cancer susceptibility. As this is the first instance of such an analysis involving genes associated with pigmentation traits and skin cancer, the clinical relevance of our findings will depend on subsequent replication and further investigation of the genes and pathways that we described here.

## 2 | MATERIALS AND METHODS

The different stages of this study are depicted in Figure 1.

### 2.1 | Identification of DNAm sites associated with pigmentation SNPs

The genetics of DNA methylation consortium (GoDMC) was created to study the genetic basis of DNAm variation and bring together resources and researchers with expertise in the epigenetics field ([www.godmc.org.uk](http://www.godmc.org.uk)). One of its aims was to carry out a meta-GWAS of DNAm, as measured on Illumina 450k or EPIC Beadchips. Results from a meta-GWAS involving 36 European cohorts ( $N = 27,750$ ) were used here. We provide a brief description of the analysis in the Methods S1.

Using GoDMC data, we searched for DNAm sites that were strongly associated ( $p < 1 \times 10^{-5}$ ) with well-known pigmentation-related SNPs previously identified via genome-wide association studies (GWAS) or candidate gene studies (reviewed in reference Pavan & Sturm, 2019), located within the genes (or their surrounding regions) *ASIP* (rs1015362, rs4911414, rs619865), *BNC2* (rs2153271), *IRF4* (rs12203592, rs12210050), *HERC2* (rs12913832), *MC1R* (rs1110400, rs11547464, rs11648785, rs1805005, rs1805006, rs1805007, rs1805008, rs1805009, rs2228479, rs258322, rs4785763, rs885479), *OCA2* (rs1800401, rs1800407), *SLC24A4* (rs12896399), *SLC24A5* (rs1426654), *SLC45A2* (rs16891982, rs28777) and *TYR* (rs1042602, rs1393350).

We uncovered 874 strong SNP-DNAm associations across 30 genomic regions, which included 391 unique DNAm sites. No data were available for SNPs in *SLC24A5* and *SLC45A2*, neither for SNPs rs1110400, rs11547464, rs1805006 and rs1805009 in *MC1R* and rs1393350 in *TYR*. Some of these polymorphisms have low minor allele frequencies (MAF) and GoDMC only included SNPs with  $MAF > 1\%$ . The other SNPs may not have been part of the genotyping platforms, successfully imputed, or strongly associated with any DNAm site ( $p > 1 \times 10^{-5}$ ).

Variance in DNAm explained by SNPs was estimated as  $2 \times \beta^2 \times MAF \times (1 - MAF)$ , where  $\beta$  is the effect size and MAF is the minor allele frequency.

Linkage disequilibrium (LD)  $r^2$  values were obtained with LDlink (<https://ldlink.nci.nih.gov/>).

### 2.2 | DNA methylation and pigmentation/sun exposure phenotypes in the Avon Longitudinal Study of Parents and Children (ALSPAC)

Cohort description and information regarding the collection of DNAm data in ALSPAC can be found in the Methods S1.

#### 2.2.1 | Regression analysis

The pigmentation/sun exposure phenotypes evaluated included the following: skin reflectance, freckles, moles, sunburning, tanning ability, hair colour and eye colour. We examined DNAm at the nearest time point with respect to when the phenotypes were measured. We assessed the association of cord blood DNAm and pigmentation traits measured before age 7 [i.e. skin reflectance (49 months), freckles (49 and/or 61 months), red hair (15 and 54 months), eye colour (54 months), tanning ability (69 months)]; the association of DNAm in childhood (~7 years old) with sunburning from birth to age 12, and total number of moles at 15 years old; and the association of DNAm in adolescence (15–17 years old) with total number of moles at 15 years old, red hair at 18 years old and tanning ability at ~25 years old. Description of most of these phenotypes has been provided in an earlier work (Bonilla et al., 2014). Data on hair colour and tanning ability in young adults were collected in a recent ALSPAC questionnaire (Life@25+), using the same scales employed in past questionnaires (Bonilla et al., 2014).

We tested the association of SNP-associated DNAm sites with pigmentation and sun exposure phenotypes using  $t$  tests, one-way ANOVA tests, linear and logistic regression models. Regressions were adjusted for age, sex and the first 10 genetic principal components to account for population stratification. Pairwise correlation between DNAm sites was examined in ALSPAC children at age 7. All analyses were carried out with the statistical package Stata v15.

#### 2.2.2 | Mediation analysis

The causal inference test (CIT) approach (Millstein et al., 2009), implemented in the R package 'cit', was employed to investigate the causal direction between pigmentation SNPs, DNAm and red hair colour. Regressions were adjusted for sex, age and the top 10 genetic principal components.

### 2.3 | Summary data-based Mendelian randomization (SMR)

In order to investigate the potential for DNAm at selected sites to be causally related to gene expression and phenotypes such as pigmentation characteristics and skin cancer, we carried out a Mendelian randomization analysis with colocalization, implemented in the

summary data-based Mendelian randomization (SMR) method (Zhu et al., 2016). We used the platform Complex Trait Genetics Virtual Lab (CTG-VL, <https://genoma.io/>; Cuellar-Partida et al., 2019) to run the statistical package SMR (Lloyd-Jones et al., 2017; Qi et al., 2018; Zhu et al., 2016). This method uses summary data from GWAS, mQTL or eQTL studies to distinguish causal or pleiotropic associations between DNAm and gene expression, or between either of these and a phenotype, from a situation where the traits are caused by different genetic variants that are in strong LD (see figure 1b of Zhu et al., 2016). If the former is true, that is a single genetic variant underlies DNAm and the other tested phenotype, the traits are said to be colocalized. The different SMR analyses run are described below. Unless otherwise reported, the SMR analyses p-value thresholds were as follows:  $p_{SMR} < 5 \times 10^{-7}/5 \times 10^{-6}$  and  $p_{HEIDI} \geq 0.05$ .

### 2.3.1 | SMR of whole-blood DNA methylation and gene expression

Peripheral blood DNAm summary data were extracted from McRae et al. (McRae et al., 2018) while summary data on gene expression in blood were obtained from three different sources: the Genotype-Tissue Expression consortium (GTEx, <https://gtexportal.org/home/>), Cap Analysis of Gene Expression (CAGE; Lloyd-Jones et al., 2017), and the blood eQTL browser (<https://genenetwork.nl/bloodseqbrowser/>; Westra et al., 2013). All expression datasets used the hg19 genome assembly.

### 2.3.2 | SMR of gene expression in skin, pigmentation/sun exposure traits and skin cancer

We ran SMR to assess the association of gene expression in sun-exposed and sun-unexposed skin, obtained from GTEx, with pigmentation and skin cancer phenotypes.

Data were obtained by CTG-VL from the UK Biobank and trait definitions can be found in the biobank website (<https://www.ukbiobank.ac.uk/>). Pigmentation characteristics analysed were skin colour (UKB ID#1717,  $N = 356,530$ ), ease of skin tanning (#1727,  $N = 353,697$ ), childhood sunburn occasions (#1737,  $N = 269,734$ ), and black ( $N = 15,809$ ) and blonde ( $N = 41,178$ ) hair colour (#1747,  $N = 360,270$ ). We also considered diagnosed malignant melanoma (#ICD10:C43,  $N = 1,672$ , total  $N = 361,194$ ), self-reported malignant melanoma ( $N = 2,898$ ), self-reported basal cell carcinoma ( $N = 3,441$ ), self-reported squamous cell carcinoma ( $N = 449$ ; all #20001, total  $N = 361,141$ ) and having melanocytic naevi (#ICD10:D22,  $N = 3,501$  cases, total  $N = 361,194$ ).

### 2.3.3 | SMR of whole-blood DNA methylation, pigmentation/sun exposure traits and skin cancer

We also investigated the potential colocalization of genetic variants underlying DNAm with pigmentation traits and skin cancer. DNAm utilized in this analysis was measured in blood (McRae et al., 2018).

## 2.4 | Heritability of the DNA methylation sites associated with pigmentation SNPs

We checked the heritability of DNAm sites using the resource provided by the Complex Disease Epigenetics Group ([www.epigenomicslab.com/online-data-resources/](http://www.epigenomicslab.com/online-data-resources/); Hannon, Knox, et al., 2018) that employs twin data to report the variance in whole-blood DNAm explained by an additive genetic component, a shared environmental component and a unique environmental component.

## 3 | RESULTS

### 3.1 | Identification of DNA methylation sites associated with pigmentation SNPs

We summarized available functional information on the analysed SNPs in Table 1.

Pigmentation SNPs that were strongly associated with DNAm sites in GoDMC are shown in Table S1.

We selected DNAm sites with the most reliable effects for further analysis as follows: sites associated with at least six of the *MC1R* region SNPs, three of the *ASIP* region SNPs or showing the strongest association with pigmentation SNPs in the other genes tested. Additionally, chosen DNAm sites had to display consistent associations with the allele that increased fair pigmentation for a minimum of five out of eight SNPs in *MC1R* and two out of three SNPs in *ASIP* (i.e. the allele increasing fair pigmentation had to always increase or always decrease DNAm at the same site for a majority of the SNPs tested in that gene). Finally, DNAm sites had to be associated with SNPs in ALSPAC participants consistently across the time points where DNAm was measured (i.e. birth, childhood, adolescence, pregnancy, middle age). Based on these criteria, we followed up a total of 25 DNAm sites: 17 in *MC1R*, 3 in *ASIP* and one each in *IRF4*, *BNC2*, *TYR*, *SLC24A4* and *HERC2*. DNAm sites that were selected for further analysis are shown in Table 2.

### 3.2 | MC1R

In this gene region, we obtained results for red hair colour loss-of-function variants (RHC) R151C (rs1805007) and R160W (rs1805008), and missense variants (rhc) V60L (rs1805005), V92M (rs2228479) and R163Q (rs885479) and for non-coding polymorphisms found in GWAS to be associated with pigmentation and melanoma risk (rs11648785, rs258322 and rs4785763).

*MC1R* non-coding polymorphisms and functional variants exhibited the largest number of associations with DNAm sites of all the pigmentation genes examined, ranging from 26 (rs1805005) to 163 (rs2228479; Table 1). Three hundred and twelve unique sites were strongly associated with at least one of the *MC1R* SNPs tested. Of these, 131 were associated with just one genetic variant. There was

**TABLE 1** Pigmentation SNPs investigated in GoDMC in association with DNA methylation (DNAm) changes

SNP	Chromosome	Position	Gene <sup>a</sup>	Effect allele <sup>b</sup>	Other allele	EAF <sup>c</sup>	DNAm sites associated	Other information <sup>d</sup>
rs12203592	6p25.3	396321	<i>IRF4</i>	T	C	0.188	3	eQTL (sun-exposed skin)
rs12210050	6p25.3	475489	<i>IRF4</i>	T	C	0.173	7	
rs2153271	9p22.2	16864521	<i>BNC2</i>	T	C	0.592	4	eQTL (whole blood)
rs1042602	11q14.3	89011046	<i>TYR</i>	A	C	0.262	5	NP_000363.1:p. Ser192Tyr
rs12896399	14q32.12	92773663	<i>SLC24A4</i>	T	G	0.468	6	
rs1800407	15q13.1	28230318	<i>OCA2</i>	T	C	0.081	1	NP_000266.2:p. Arg419Gln/ NP_001287913.1:p. Arg395Gln
rs1800401	15q13.1	28260053	<i>OCA2</i>	G	A	0.946	2	NP_000266.2:p. Arg305Trp
rs12913832	15q13.1	28365618	<i>HERC2</i>	G	A	0.744	3	eQTL (whole blood)
rs258322	16q24.3	89755903	<i>CDK10</i>	A	G	0.118	98	eQTL (sun-exposed and sun-unexposed skin and whole blood)
rs1805005	16q24.3	89985844	<i>MC1R</i>	T	G	0.121	26	rhc (NP_002377.4:p. Val60Leu)
rs2228479	16q24.3	89985940	<i>MC1R</i>	A	G	0.089	163	rhc (NP_002377.4:p. Val92Met)/eQTL (sun-exposed and sun- unexposed skin and whole blood)
rs1805007	16q24.3	89986117	<i>MC1R</i>	T	C	0.086	97	RHC (NP_002377.4:p. Arg151Cys)/ eQTL (sun-exposed and sun- unexposed skin and whole blood)
rs1805008	16q24.3	89986144	<i>MC1R</i>	T	C	0.083	97	RHC (NP_002377.4:p. Arg160Trp)/ eQTL (sun-exposed and sun- unexposed skin and whole blood)
rs885479	16q24.3	89986154	<i>MC1R</i>	A	G	0.105	110	rhc (NP_002377.4:p. Arg163Gln)/ eQTL (sun- exposed skin)
rs4785763	16q24.3	90066936	<i>AFG3L1P</i>	A	C	0.332	104	eQTL/ eQTL (sun- exposed and sun- unexposed skin and whole blood)
rs11648785	16q24.3	90084561	<i>DBNDD1</i>	C	T	0.689	67	
rs4911414	20q11.22	32729444	<i>ASIP</i>	T	G	0.342	24	
rs1015362	20q11.22	32738612	<i>ASIP</i>	C	T	0.727	27	eQTL (sun-exposed and sun-unexposed skin and whole blood)
rs619865	20q11.22	33867697	<i>EIF6</i>	A	G	0.109	29	
rs16891982	5p13.2	33951693	<i>SLC45A2</i>	G	C	0.938	n/a	NP_001012527.2:p. Leu374Phe
rs28777	5p13.2	33958959	<i>SLC45A2</i>	A	C	0.956	n/a	
rs1393350	11q14.3	89011046	<i>TYR</i>	A	G	0.244	n/a	
rs1426654	15q21.1	48426484	<i>SLC24A5</i>	A	G	0.997	n/a	NP_995322.1:p. Thr111Ala

(Continues)

TABLE 1 (Continued)

SNP	Chromosome	Position	Gene <sup>a</sup>	Effect allele <sup>b</sup>	Other allele	EAF <sup>c</sup>	DNAm sites associated	Other information <sup>d</sup>
rs1805006	16q24.3	89985918	MC1R	A	C	0.010	n/a	RHC (NP_002377.4:p. Asp84Glu)
rs11547464	16q24.3	89986091	MC1R	A	G	0.009	n/a	RHC (NP_002377.4:p. Arg142His)
rs1110400	16q24.3	89986130	MC1R	C	T	0.008	n/a	RHC (NP_002377.4:p. Ile155Thr)
rs1805009	16q24.3	89986546	MC1R	C	G	0.008	n/a	RHC (NP_002377.4:p. Asp294His)

<sup>a</sup>In this study, for simplicity, SNPs on chromosome 16q24.3 are considered part of the *MC1R* genetic region, and SNPs on chromosome 20q11.22 are considered part of the *ASIP* genetic region.

<sup>b</sup>The effect allele is the allele associated with a lighter skin, hair and eye colour, and skin cancer susceptibility.

<sup>c</sup>Effect allele frequency from GoDMC for SNPs with associated DNAm sites. Effect allele frequency from 1,000 Genomes for SNPs not available in GoDMC.

<sup>d</sup>RHC = high penetrance red hair colour variant, rhc = low penetrance red hair colour variant. eQTL information obtained from the GTEx Portal v8 release.

overlap between DNAm sites associated with different *MC1R* SNPs (Table S2). Twenty-five DNAm sites were associated with a minimum of 6 of the SNPs with results, while 5 were associated with all 8 of them. However, not all of these DNAm sites showed a consistent direction of association with the alleles that increase the odds of having a fair skin phenotype and being affected by melanoma. For instance, full consistency was achieved for *MC1R* DNAm site cg01097406 (i.e. all alleles that are associated with fair skin decrease DNAm at this position) and cg07130392, both associated with 7 of the 8 SNPs tested, and for cg08845973 and cg09738481, associated with 6 SNPs.

Interestingly, DNAm sites associated with both RHC variants, which are located in the first exon of the gene *SPIRE2* (i.e. cg03388025, cg04287289, cg07402062, cg21285383, cg26513180), exhibited effects of similar magnitude and direction (Figure S1). However, for these same DNAm sites, the association with rhc variants rs2228479 and rs885479 occurred in the opposite direction to that of the RHCs (Table S1).

Although most associations uncovered were between SNPs and DNAm sites in *cis*, *MC1R* showed several strong associations in *trans* (Figure 2). Besides associations with more distant DNAm sites on chromosome 16 (>1 Mb), we identified associations with sites on chromosomes 1, 2, 3, 4, 5, 8, 10, 11, 12 and 13 (Table S1, Figure 2). One of the *trans* sites most frequently associated with *MC1R* SNPs (7 of them, largely in low LD) was cg26114043 located in the gene *inturned planar cell polarity protein (INTU)* on chromosome 4 position 128,544,375. This gene is related to the processes of keratinocyte differentiation and hair follicle morphogenesis (<https://www.ncbi.nlm.nih.gov/gene/27152>).

The maximum variance in DNAm explained by a pigmentation SNP was 24.4% shown by the pair rs2228479-cg09569215. Amongst the RHC, the highest  $R^2$  was found for rs1805007, which explained 17.8% of variation at cg08854185 (Table S1).

### 3.3 | Other pigmentation genes

The number of SNP-DNAm site associations for the other pigmentation genes examined ranged from 1 (rs1800407 in *OCA2*) to 29 (rs619865 in *ASIP*; Table 1). There were 127 unique DNAm sites associated with at least one SNP in these 7 genes, while 102 DNAm sites were associated with only one polymorphism.

Five different DNAm sites in the *ASIP* region were associated with three SNPs each, whereas one site was associated with 3 SNPs in the *OCA2/HERC2* region (Table S2). There was no DNAm site overlap between SNPs rs12210050 and rs12203592 in *IRF4*. The only fully consistent association between SNPs and DNAm sites was with cg01901788, located near *ASIP* (i.e. for all 3 SNPs the alleles associated with a fair skin phenotype increased DNAm at this site).

As expected, most associations were in *cis*, although there were three associations in *trans* in *ASIP* (distant chromosome 20) and one in *BNC2* (chromosome 1; Table S1, Figure 2). Maximum variance in DNAm explained by SNPs was: 8.9% (*ASIP* rs4911414-cg08999081), 8.6% (*BNC2* rs2153271-cg03291755), 5.3% (*OCA2* rs1800407-cg20906524), 3.0% (*TYR* rs1042602-cg25941151), 2.7% (*IRF4* rs12210050-cg26187313), 1.7% (*HERC2* rs12913832-cg20906524) and 1.3% (*SLC24A4* rs12896399-cg16993582; Table S1).

### 3.4 | DNA methylation sites and pigmentation/sun exposure phenotypes in ALSPAC

Pairwise correlation coefficients (Pearson's) and p-values for DNAm sites measured in ALSPAC children at age 7 are shown in Table S3a-c. There was moderate to strong correlation between DNAm sites in the region that extends from cg04287289 to cg03388025 in *MC1R* (Table S3b), and some moderate correlation was detected between DNAm sites in unlinked genetic regions (Table S3a).

**TABLE 2** Pigmentation SNP-associated DNA methylation (DNAm) sites selected for follow-up

DNAm site	Gene region	Gene	chr	chr position	# Associated SNPs	Associated in ALSPAC at all time points?
cg26114043	MC1R <sup>a</sup>	INTU	4	128544375	7	Y
cg09806625	IRF4	EXOC2	6	611523	1	Y
cg03291755	BNC2	BNC2	9	16868891	1	Y
cg05041596	TYR	NAALAD2	11	89867385	1	Y
cg04136915	SLC24A4	intergenic	14	92721383	1	Y
cg14091419	HERC2	HERC2	15	28356810	1	N <sup>b</sup>
cg05504729	MC1R	ANKRD11	16	89399490	6	Y
cg08845973	MC1R	SPG7	16	89592071	6	Y
cg09738481	MC1R	CPNE7	16	89653835	6	Y
cg01097406	MC1R	intergenic	16	89675127	7	Y
cg04240660	MC1R	CHMP1A	16	89714849	6	Y
cg03605463	MC1R	LOC284241	16	89740564	6	Y
cg04287289	MC1R	FANCA	16	89883240	6	Y
cg26513180	MC1R	FANCA	16	89883248	7	Y
cg07402062	MC1R	SPIRE2	16	89894098	7	Y
cg21285383	MC1R	SPIRE2	16	89894308	7	Y
cg03388025	MC1R	SPIRE2	16	89894329	7	Y
cg01768446	MC1R	MC1R	16	89982419	8	Y
cg01023759	MC1R	MC1R	16	89983766	8	Y
cg01883217	MC1R	DEF8	16	90015832	8	Y
cg05185784	MC1R	DEF8	16	90016020	7	Y
cg15936718	MC1R	GAS8	16	90088801	7	Y
cg01901788	ASIP	MAP1LC3A	20	33145848	3	Y
cg08999081	ASIP	PIGU	20	33150536	3	Y
cg21132536	ASIP	ACSS2	20	33465180	3	Y

<sup>a</sup>cg26114043 is located on chromosome 4 but it is strongly associated with SNPs on chromosome 16 (MC1R region).

<sup>b</sup>cg14091419 was associated with HERC2 SNP rs12913832 in pregnancy and middle-age only.

### 3.4.1 | Regression analysis

When we investigated the association of the selected DNAm sites with pigmentation and sun exposure traits in ALSPAC children, we found the strongest associations with red hair colour. The most robust association with this phenotype was that of cg07402062, where redhead children showed lower levels of DNAm at this site than non-redhead children (Table S4). This effect was seen with cord blood DNAm and red hair at 15 and 54 months (after correction for multiple testing), as well as with DNAm in adolescence and red hair at 18 years old (before correction for multiple testing). Associations with the other traits tested were either weaker or absent.

### 3.4.2 | Mediation analysis

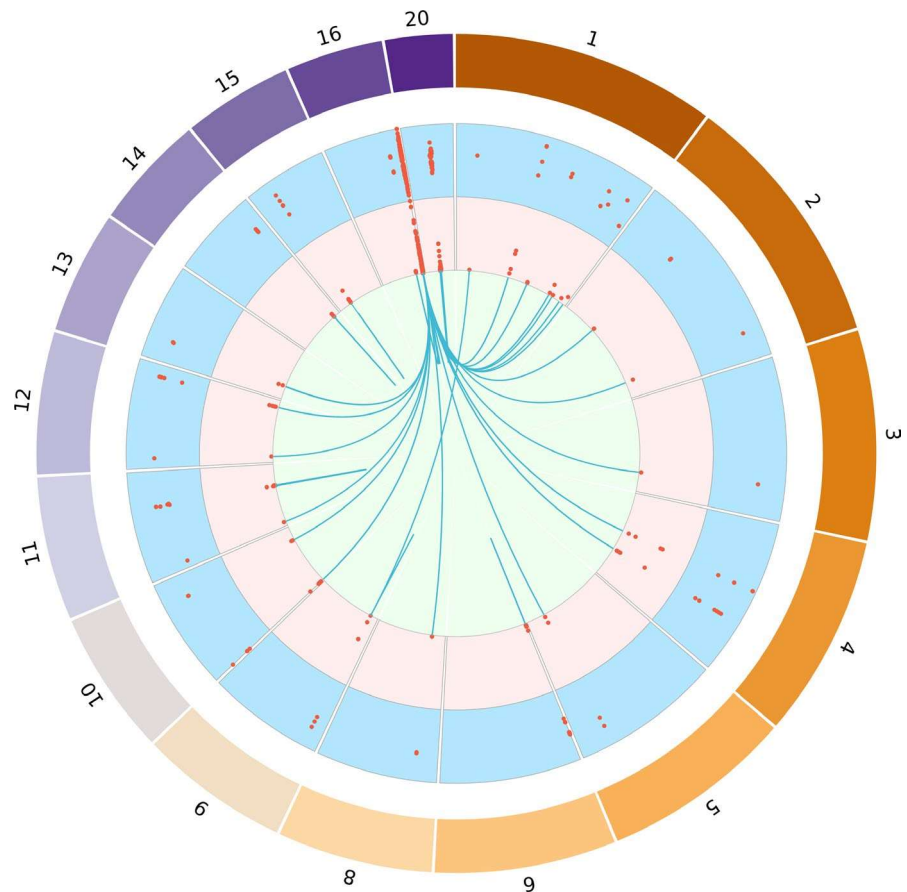
The causal inference test (CIT) showed inconclusive results for SNPs rs1805007, rs1805008, rs258322, and rs4785763 (Table S5). There

was evidence for a mediating effect of DNAm at cg07402062 between these pigmentation SNPs and red hair colour, as well as for red hair as a mediator of the relationship between the pigmentation SNPs and DNAm at cg07402062. On the other hand, for variants rs2228479 and rs885479, we found evidence of a causal effect of DNAm at cg07402062 on red hair colour, but not the other way around. Finally, rs11648785 appeared to influence both traits independently.

### 3.5 | Summary data-based Mendelian randomization (SMR) of whole-blood DNA methylation and gene expression

Out of the 25 DNAm sites under analysis, four (cg01023759, cg09738481, cg14091419 and cg26114043) were not present in any of the expression datasets used. Since the data sources only included variation in *cis*, cg26114043, located on chromosome 4, was left out. The 21 remaining sites showed associations with gene expression to varying degrees which were, for the most part, consistent

**FIGURE 2** Circos plot showing the relationship between SNPs in pigmentation genes and DNA methylation (DNAm) sites in cis and trans. The external track (light blue) depicts the effect of each pigmentation SNP on DNAm (negative in red, positive in blue). The internal track (pink) depicts the  $R^2$  (variance explained) values. Information extracted from Table S1



between datasets. Of note were the associations of correlated DNAm sites cg03388025, cg04287289, cg07402062, cg21285383 and cg26513180 with the expression of *SPIRE2*, cg08845973 with *SPG7*, cg01901788 with *NCOA6* and cg21132536 with *GGT7*. Results of this analysis, including total numbers of associations per DNAm site and dataset, are shown in Table S6.

### 3.6 | Summary data-based Mendelian randomization (SMR) of gene expression in skin, pigmentation/sun exposure traits and skin cancer

We found no evidence of a pleiotropic or causal association between gene expression at 16q24.3, 20q11.22, or any of the other pigmentation regions included in this study, and ease of skin tanning, skin colour, childhood sunburn occasions, melanocytic naevi and SCC, in sun-exposed and sun-unexposed skin.

There was evidence of an association with the same genetic variant between black hair colour and the expression of genes *ASIP* and *FAM83C* (20q11.22) in sun-exposed skin and of *CDK10* (16q24.3), *ASIP*, *NCOA6* (20q11.22) and *EXOC2* (6p25.3) in sun-unexposed skin, whereas blonde hair colour was associated with the expression of *ASIP* and *FAM83C* in sun-exposed skin and with *NCOA6* and *EXOC2* in sun-unexposed skin. In addition, *ASIP* gene expression was associated with self-reported and diagnosed malignant melanoma, and

BCC in both skin tissues, and *CDK10* was associated with self-reported and diagnosed malignant melanoma, and BCC in unexposed skin (Table S7).

When the HEIDI p-value threshold was Bonferroni-corrected for multiple testing (instead of being set as  $\geq 0.05$ ), additional genes showed potential pleiotropic/causal associations with the traits of interest, including *CPNE7*, *DBNDD1*, *DEF8*, *MC1R*, *SNAI3-AS1*, *SPATA2L*, *URAHP* in 16q24.3, and *EIF6*, *GGT7*, *UQCC1* in 20q22.11.

### 3.7 | Additional summary data-based Mendelian randomization (SMR) of whole-blood DNA methylation and gene expression

Because the DNAm sites we selected were associated with the expression of just 5 of the genes identified in the previous step (i.e. *DBNDD1*, *DEF8*, *GGT7*, *NCOA6*, *SPATA2L*), we searched for additional DNAm sites that showed an association with the expression of the genes reported above. For *CDK10*, there was no common colocalized DNAm site across all the datasets, rather associations were evident with several DNAm sites in the 16q24.3 region. For *ASIP* and *FAM83C*, there were no data available or no evidence of association with DNAm sites, and *EXOC2* showed two overlapping associated DNAm sites between the CAGE and GTEx datasets. Results for these and other genes are shown in Table S8.

### 3.8 | Summary data-based Mendelian randomization (SMR) of whole-blood DNA methylation, pigmentation/sun exposure traits and skin cancer

We uncovered 696 DNAm sites associated and colocalized with pigmentary and skin cancer traits (Table S9), but of these, only a reduced number ( $n = 18$ ) were associated with gene expression colocalizing with these phenotypes (see Table S8). From this set of 18 DNAm sites, 9 were identified in our original analysis in GoDMC but only one (cg01901788) was on the list we followed up (the others were associated with three SNPs or less, Table S2). SNPs associated with these “mediating” 9 DNAm sites were as follows: rs11648785, rs1805008 (R160W), rs2228479 (V92M) and rs4785763 in the *MC1R* region; and rs1015362, rs4911414 and rs619865 in the *ASIP* region (Table 3).

### 3.9 | Heritability

Nine (38%) of 24 DNAm sites followed up showed a heritability value of 80% or more (additive genetic contribution), almost all of these are located in *MC1R*. DNAm site cg01023759 was unavailable in this dataset. Median heritability was 63% (IQR 31%, 85%; Table S10). The five correlated DNAm sites associated with *SPIRE2* expression and also with SNPs rs1805007, rs1805008, rs2228479 and rs885479 showed a heritability higher than 63%.

## 4 | DISCUSSION

In this study, we reported the existence of strong associations between pigmentation/sun exposure/skin cancer-associated SNPs and DNAm variation at specific sites, mostly in *cis*, with a few cases in *trans*. This appears to be especially marked in the *MC1R* gene region, which could be related to the high CpG content present in this region (Martinez-Cadenas et al., 2013). The SNP-DNAm site associations were observed for functional variants (RHC and rhc) as well as for non-coding polymorphisms within *MC1R*.

For a set of selected DNAm sites (those associated with several pigmentation SNPs in the same direction), we assessed their effect on pigmentation/sun exposure phenotypes measured in ALSPAC participants. Although sample sizes were modest, we detected a strong inverse association of red hair with cg07402062, which is located in the gene *SPIRE2* (*MC1R* region). Remarkably, a recent GWAS of hair colour has identified the strongest genetic signal of association with red hair (versus brown/black hair) at SNP rs34357723 (T allele associated with red hair; LD with rs1805007,  $r^2 = .27$ ), a signal that remained relevant even after adjustment for *MC1R* coding variants (Morgan et al., 2018). Rs34357723 is an intronic variant of *SPIRE2*, and GoDMC data indicate that it is also a strong mQTL for cg07402062, explaining ~16% of DNAm at this site (T allele beta = -0.82, SE = 0.01). The CIT analysis suggested that a

mediation effect of cg07402062 DNAm between rs34357723 and red hair colour is compatible with ALSPAC data (see Table S5), although replication of these findings is needed.

Galanter and colleagues reported that DNAm at cg07402062 and cg21285383 differed between ethnic groups and that these differences could be accounted for by differences in genetic ancestry (Galanter et al., 2017), which would also be consistent with hair colour variation between groups.

The high heritability detected for these two DNAm sites, along with cg03388025, cg04287289 and cg26513180, and their association with RHC and rhc variants suggests a close relationship between *SPIRE2* DNAm status and the *MC1R* gene structure.

*SPIRE2* was recently shown to participate in myosin Va-dependent transport of melanosomes in melanocytes via the generation of actin tracks (Alzahofi et al., 2018). However, despite *SPIRE2* expression colocalizing with DNAm at cg07402062 and correlated sites, we detected no evidence of colocalization with hair colour GWAS signals, and thus, we cannot yet conclude that this gene causally underlies the phenotype.

Using summary data and two-sample Mendelian randomization implemented in SMR, we determined that the selected DNAm sites affected expression of a number of genes in *cis*, but there was limited overlap between this set of genes and those that showed an association of expression with pigmentation traits and skin cancer. Given that the DNAm expression analysis was carried out using whole-blood data, and the expression-phenotype analysis used skin data, the differences observed may be due to this discrepancy.

Previous reports on DNAm in pigmentation/skin cancer-associated regions include a study by Roos and colleagues (Roos et al., 2017) who showed that DNAm in healthy human skin was associated in *cis* with SNPs underlying melanoma risk discovered by GWAS, including two polymorphisms in *MC1R* (rs258322, rs4785763, which we tested) and one in the *ASIP* region (rs910873, not considered by us but in LD with rs619865). SNPs in *SLC45A2* (rs35390) and *TYR* (rs1393350, rs1847134) were also investigated but yielded no evidence of association with DNAm, a finding similar to ours. Roos et al. also found that none of the DNAm sites associated with these SNPs was associated with naevus count in their epigenome-wide association study (EWAS). Likewise, in our SMR analysis presence/absence of melanocytic naevi failed to colocalize with any DNAm site.

Another study by Budden and Bowden described the regulation of *MC1R* expression by a CpG island overlapping a potential enhancer in the *MC1R* gene (Budden & Bowden, 2018). These authors found that this CpG island is unmethylated in melanocytes but highly methylated in other skin cell types. In addition, most of the melanoma cell lines and tumours examined in their study were methylated in this region. Further analysis indicated that differential methylation of the CpG island in tumours mainly affected the expression of *MC1R*, and less so of *TUBB3*. Budden and Bowden also reported that there was weak evidence of increasing CpG island methylation levels with increasing numbers of red hair colour alleles (R) in melanoma

**TABLE 3** DNA methylation sites identified in GoDMC that colocalized with the expression of genes which, in turn, colocalized with pigmentation/sun exposure/skin cancer phenotypes. These DNAm sites also colocalized with pigmentation/sun exposure/skin cancer phenotypes in an independent test

DNAm sites	Chromosome	Associated pigmentation SNPs	Colocalized genes	Colocalized phenotypes <sup>a</sup>
<b>cg01901788</b>	20q11.22	rs1015362, rs4911414, rs619865	NCOA6	skin colour, ease of skin tanning, blonde hair, black hair
cg04270048	20q11.22	rs4911414, rs619865	NCOA6	skin colour, ease of skin tanning, childhood sunburn, blonde hair, black hair
cg16810054	20q11.22	rs619865	NCOA6	blonde hair, black hair
cg04378830	16q24.3	rs2228479	DBNDD1, MC1R	blonde hair, black hair, self-reported BCC
cg04752812	16q24.3	rs2228479	DBNDD1, DEF8	blonde hair, skin colour, ease of skin tanning, self-reported BCC
cg09569215	16q24.3	rs11648785, rs2228479, rs4785763	MC1R	blonde hair, black hair
cg10062109	16q24.3	rs1805008, rs4785763	CDK10	blonde hair, black hair, self-reported melanoma, skin colour, ease of skin tanning, childhood sunburn, diagnosed melanoma, self-reported BCC
cg11215013	16q24.3	rs2228479	SPATA2L	skin colour, black hair, self-reported melanoma
cg26598918	16q24.3	rs1805008, rs4785763	CDK10	blonde hair, black hair, skin colour, ease of skin tanning, childhood sunburn, diagnosed melanoma, self-reported BCC

Note: In bold: DNAm site followed up in this study.

<sup>a</sup>Phenotypes in red colocalized with the DNAm site shown in the first column, phenotypes in blue colocalized with the expression of the gene/s listed in the fourth column, and phenotypes in green colocalized with both, the DNAm site and gene expression. See Tables S6–S9 for the colocalization results.

tumours from The Cancer Genome Atlas (TCGA) dataset (Budden & Bowden, 2018).

While seven of the 12 *MC1R* SNPs that we evaluated were located within the CpG island identified by Budden and Bowden, they were strongly associated with DNAm sites located nearby, although outside the island/enhancer region, making direct comparisons of results difficult.

Lastly, Martinez-Cadenas and colleagues (Martinez-Cadenas et al., 2013) postulated that the considerable nucleotide diversity evident at *MC1R* was probably the result of a high mutation rate that occurred as a consequence of the elevated CpG concentration present in the region. Chromosome 16q is also one of the richest regions in gene and GC content, which has been found to be correlated (Zeng et al., 2018). In this study, polymorphisms in *MC1R* were associated with DNAm variation at substantially more sites than is the case in other pigmentation genes, with lower GC content, tested (for instance, *MC1R* region SNPs are associated with 95 DNAm sites each on average, whereas *ASIP* region SNPs were associated with an average of 27 DNAm sites each, see Table 1).

From an evolutionary point of view, it is interesting to consider the potential for epigenetic modifications in pigmentation genes as a mechanism of adaptation to variable levels of ultraviolet radiation exposure, providing phenotypic flexibility in a changing environment, similar to what has been suggested for adaptation to high altitude (Julian, 2017). Since there is evidence that pigmentation genes, such

as *MC1R* and *ASIP*, have been shaped by natural selection (Martinez-Cadenas et al., 2013; Norton et al., 2007; Quillen et al., 2019), polymorphic variants in them may act, in part, by affecting how they are epigenetically regulated to optimize a response to an environment with reduced sunlight.

#### 4.1 | Limitations

Our analysis was conducted on DNA derived from whole-blood rather than skin tissue, and any pigmentation SNP-DNAm site associations may not be as strong or present in skin. We used heritability estimates provided by Hannon, Gorrie-Stone, et al. (2018) to assess the likelihood of the DNAm sites examined to vary in a similar way in different tissues. These authors reported a low contribution on average of additive genetic effects on DNAm across the genome (~16%) and a higher heritability of the sites that were correlated across tissues (median 71%). Since most of the DNAm sites (16/24) we analysed had heritability values > 50%, it is possible that the associations that we detected with pigmentation SNPs are also there in skin. Moreover, for two *MC1R* polymorphisms, Roos et al. (2017) already described an association in healthy skin with some of the same DNAm sites identified in our study. Although not examining mQTLs, Yuan et al. (2019) also reported an association of DNAm sites in 16q24.3 within the genes *SPG7*, *DPEP1*, *FANCA* and *SPIRE2*,

which included cg04287289 and cg26513180, with ethnicity in placental tissue. However, the small overlap between the genes whose whole-blood expression was affected by the DNAm sites we tested and those whose expression in skin was associated with pigmentation and skin cancer phenotypes could be suggesting that there are in fact differences between tissues.

Unfortunately, mQTL data on populations of non-European ancestry are limited. In the only study involving one of these populations that we identified, Pierce and colleagues reported an association of 5 of our pigmentation DNAm sites with SNPs in *cis* in the population of Bangladesh (Pierce et al., 2018). While this is an indication that similar genetic–epigenetic relationships are likely to occur in other groups, knowledge of the role of these chromosomal regions on pigmentation in non-White populations is still scarce (Pavan & Sturm, 2019).

Each of the methods we used has their own particular limitations, so it is advisable to be cautious in the interpretation of our findings.

The CIT method requires that the analysis is carried out in one sample where the SNP, exposure and outcome have all been measured, and therefore, statistical power is constrained by the size of this sample. It may also be affected by collider bias and phenotype measurement error. In order to protect against collider bias, it is essential to adjust for all confounders of the exposure–outcome relationship. While we adjusted the regression models for age, sex and genetic principal components, unknown confounders may be still playing a role (Hemani et al., 2017).

The SMR method was conceived to try to tell apart pleiotropy from LD in the event of an association between gene expression and a complex trait, and thus to interpret results under a model of pleiotropy rather than causality, even though it could be consistent with the latter as well (Zhu et al., 2016). To tease out a causal effect, additional analyses are necessary, such as two-step Mendelian randomization (Relton & Davey Smith, 2012) using independent (in *cis* or *trans*) mQTLs. However, independent instrumental variables for the DNAm sites in this study have been hard to find since there is usually considerable LD with the original SNPs. As a consequence, SMR results are mostly intended to be used to highlight genes for follow-up studies and not to establish unequivocal causal relationships (Zhu et al., 2016).

In summary, our results suggest that DNAm may lie on the causal pathway leading from pigmentation-associated SNPs to pigmentation/sun exposure traits and eventually skin cancer, possibly via regulation of gene expression, although more evidence to support these findings is needed. This study puts forward a set of genes that could be prioritized in upcoming investigations (e.g. *CDK10* or *SPIRE2*) to delve deeper into the role played by the variability in DNAm as a potentially causal risk factor for skin cancer. We have reported for the first time the existence of DNAm sites strongly influenced by genetic variants across pigmentation genes and demonstrated the importance of integrating different sources of variation and of uncovering the interaction of markers within and between chromosomal regions (for instance, on chromosome 16 and chromosome 4)

to obtain a better picture of pathways leading to disease. However, there is still some way to go before human genetic and epigenetic information can be introduced fully into the dermatology clinic for the provision of personalized medicine.

## ACKNOWLEDGEMENTS

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. CB was supported by a Universidade de São Paulo/ Coordenação de Aperfeiçoamento de Pessoal de Nível Superior fellowship (process no. 88887.160006/2017-00). JLM, GH and HRE are currently supported by the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, UK (MC\_UU\_00011/5). GH is funded by the Wellcome Trust and Royal Society [208806/Z/17/Z]. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. GWAS data for the children were generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. GWAS data for the mothers were generated with support from the Wellcome Trust (WT088806). Methylation data were generated with support from BBSRC (BBI025751/1 and BB/1025263/1) and from MRC (MC\_UU\_12013/1, MC-UU\_12013/2 and MC\_UU\_12013/8). A comprehensive list of grant funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>).

## CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

## ORCID

Carolina Bonilla  <https://orcid.org/0000-0002-9972-6383>

## REFERENCES

- Alzahofi, N., Robinson, C. L., Welz, T., Page, E. L., Briggs, D. A., Stainthorp, A. K., Reekes, J., Elbe, D. A., Straube, F., Tate, E. W., Goff, P. S., Sviderskaya, E. V., Cantero, M., Montoliu, L., Bailly, M., Kerkhoff, E., & Hume, A. N. (2018). Rab27a co-ordinates actin-dependent transport by controlling organelle-associated motors and track assembly proteins. *bioRxiv*, 11–43. <https://doi.org/10.1101/314153>
- Apalla, Z., Nashan, D., Weller, R. B., & Castellsagué, X. (2017). Skin cancer: Epidemiology, disease burden, pathophysiology, diagnosis, and therapeutic approaches. *Dermatology and Therapy*, 7(S1), 5–19. <https://doi.org/10.1007/s13555-016-0165-y>
- Bonilla, C., Ness, A. R., Wills, A. K., Lawlor, D. A., Lewis, S. J., & Davey Smith, G. (2014). Skin pigmentation, sun exposure and vitamin D levels in children of the Avon Longitudinal Study of Parents and Children. *BMC Public Health*, 14(1), 597. <https://doi.org/10.1186/1471-2458-14-597>
- Budden, T., & Bowden, N. A. (2018). MC1R CpG island regulates MC1R expression and is methylated in a subset of melanoma tumours. *Pigment Cell & Melanoma Research*, 32, 320–325. <https://doi.org/10.1111/pcmr.12739>

- Cahoon, E. K., Kitahara, C. M., Ntowe, E., Bowen, E. M., Doody, M. M., Alexander, B. H., Lee, T., Little, M. P., Linet, M. S., & Freedman, D. M. (2015). Female estrogen-related factors and incidence of basal cell carcinoma in a nationwide US cohort. *Journal of Clinical Oncology*, 33(34), 4058–4065. <https://doi.org/10.1200/JCO.2015.62.0625>
- Cappetta, M., Berdasco, M., Hochmann, J., Bonilla, C., Sans, M., Hidalgo, P. C., Artagaveytia, N., Kittles, R., Martínez, M., Esteller, M., & Bertoni, B. (2015). Effect of genetic ancestry on leukocyte global DNA methylation in cancer patients. *BMC Cancer*, 15(1), 434. <https://doi.org/10.1186/s12885-015-1461-0>
- Cuellar-Partida, G., Lundberg, M., Kho, P. F., D'Urso, S., Gutierrez-Mondragon, L. F., & Hwang, L.-D. (2019). Complex-Traits Genetics Virtual Lab: A community-driven web platform for post-GWAS analyses. *bioRxiv*, 518027. <https://doi.org/10.1101/518027>
- Galanter, J. M., Gignoux, C. R., Oh, S. S., Torgerson, D., Pino-Yanes, M., Thakur, N., Eng, C., Hu, D., Huntsman, S., Farber, H. J., Avila, P. C., Brigino-Buenaventura, E., LeNoir, M. A., Meade, K., Serebrisky, D., Rodríguez-Cintrón, W., Kumar, R., Rodríguez-Santana, J. R., Seibold, M. A., ... Zaitlen, N. (2017). Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife*, 6, 1–24. <https://doi.org/10.7554/eLife.20532>
- Gordon, R. (2013). Skin cancer: An overview of epidemiology and risk factors. *Seminars in Oncology Nursing*, 29(3), 160–169. <https://doi.org/10.1016/j.soncn.2013.06.002>
- Grönniger, E., Weber, B., Heil, O., Peters, N., Stüb, F., Wenck, H., Korn, B., Winnefeld, M., & Lyko, F. (2010). Aging and chronic sun exposure cause distinct epigenetic changes in human skin. *PLoS Genetics*, 6(5), e1000971. <https://doi.org/10.1371/journal.pgen.1000971>
- Hannon, E., Gorrie-Stone, T. J., Smart, M. C., Burrage, J., Hughes, A., Bao, Y., Kumari, M., Schalkwyk, L. C., & Mill, J. (2018). Leveraging DNA-methylation quantitative-trait loci to characterize the relationship between methylomic variation, gene expression, and complex traits. *The American Journal of Human Genetics*, 103(5), 654–665. <https://doi.org/10.1016/j.ajhg.2018.09.007>
- Hannon, E., Knox, O., Sugden, K., Burrage, J., Wong, C. C. Y., Belsky, D. W., Corcoran, D. L., Arseneault, L., Moffitt, T. E., Caspi, A., & Mill, J. (2018). Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLOS Genetics*, 14(8), e1007544. <https://doi.org/10.1371/journal.pgen.1007544>
- Hemani, G., Tilling, K., & Davey Smith, G. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genetics*, 13(11), 1–22. <https://doi.org/10.1371/journal.pgen.1007081>
- Holzschek, N., Söhle, J., Schläger, T., Falckenhayn, C., Grönniger, E., Kolbe, L., Wenck, H., Terstegen, L., Kaderali, L., Winnefeld, M., & Gorges, K. (2020). Concomitant DNA methylation and transcriptome signatures define epidermal responses to acute solar UV radiation. *Scientific Reports*, 10(1), 12918. <https://doi.org/10.1038/s41598-020-69683-8>
- Hu, S., Parmet, Y., Allen, G., Parker, D. F., Ma, F., Rouhani, P., & Kirsner, R. S. (2009). Disparity in Melanoma. *Archives of Dermatology*, 145(12), 1369–1374. <https://doi.org/10.1001/archdermatol.2009.302>
- Julian, C. G. (2017). Epigenomics and human adaptation to high altitude. *Journal of Applied Physiology*, 123(5), 1362–1370. <https://doi.org/10.1152/jappphysiol.00351.2017>
- Lloyd-Jones, L. R., Holloway, A., McRae, A., Yang, J., Small, K., Zhao, J., Zeng, B., Bakshi, A., Metspalu, A., Dermitzakis, M., Gibson, G., Spector, T., Montgomery, G., Esko, T., Visscher, P. M., & Powell, J. E. (2017). The Genetic Architecture of Gene Expression in Peripheral Blood. *The American Journal of Human Genetics*, 100(2), 228–237. <https://doi.org/10.1016/j.ajhg.2016.12.008>
- Martínez-Cadenas, C., López, S., Ribas, G., Flores, C., García, O., Sevilla, A., Smith-Zubiaga, I., Ibarrola-Villaba, M., Pino-Yanes, M. D. M., Gardeazabal, J., Boyano, D., García de Galdeano, A., Izagirre, N., de la Rúa, C., & Alonso, S. (2013). Simultaneous purifying selection on the ancestral MC1R allele and positive selection on the melanoma-risk allele V60L in South Europeans. *Molecular Biology and Evolution*, 30(12), 2654–2665. <https://doi.org/10.1093/molbev/mst158>
- McCartney, D. L., Hillary, R. F., Stevenson, A. J., Ritchie, S. J., Walker, R. M., Zhang, Q., Morris, S. W., Birmingham, M. L., Campbell, A., Murray, A. D., Whalley, H. C., Gale, C. R., Porteous, D. J., Haley, C. S., McRae, A. F., Wray, N. R., Visscher, P. M., McIntosh, A. M., Evans, K. L., ... Marioni, R. E. (2018). Epigenetic prediction of complex traits and death. *Genome Biology*, 19(1), 136. <https://doi.org/10.1186/s13059-018-1514-1>
- McRae, A. F., Marioni, R. E., Shah, S., Yang, J., Powell, J. E., Harris, S. E., Gibson, J., Henders, A. K., Bowdler, L., Painter, J. N., Murphy, L., Martin, N. G., Starr, J. M., Wray, N. R., Deary, I. J., Visscher, P. M., & Montgomery, G. W. (2018). Identification of 55,000 Replicated DNA Methylation QTL. *Scientific Reports*, 8(1), 17605. <https://doi.org/10.1038/s41598-018-35871-w>
- Millstein, J., Zhang, B., Zhu, J., & Schadt, E. E. (2009). Disentangling molecular relationships with a causal inference test. *BMC Genetics*, 10, 1–15. <https://doi.org/10.1186/1471-2156-10-23>
- Morgan, M. D., Pairo-Castineira, E., Rawlik, K., Canela-Xandri, O., Rees, J., Sims, D., Tenesa, A., & Jackson, I. J. (2018). Genome-wide study of hair colour in UK Biobank explains most of the SNP heritability. *Nature Communications*, 9(1), 5271. <https://doi.org/10.1038/s41467-018-07691-z>
- Nasti, T. H., & Timares, L. (2015). MC1R, eumelanin and pheomelanin: Their role in determining the susceptibility to skin cancer. *Photochemistry and Photobiology*, 91(1), 188–200. <https://doi.org/10.1111/php.12335>
- Nikolouzakis, T. K., Falzone, L., Lasithiotakis, K., Krüger-Krasagakis, S., Kalogeraki, A., Sifaki, M., Spandidos, D. A., Chrysos, E., Tsatsakis, A., & Tsiaoussis, J. (2020). Current and future trends in molecular biomarkers for diagnostic, prognostic, and predictive purposes in non-melanoma skin cancer. *Journal of Clinical Medicine*, 9(9), 2868. <https://doi.org/10.3390/jcm9092868>
- Norton, H. L., Kittles, R. A., Parra, E., McKeigue, P., Mao, X., Cheng, K., Canfield, V. A., Bradley, D. G., McEvoy, B., & Shriver, M. D. (2007). Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Molecular Biology and Evolution*, 24(3), 710–722. <https://doi.org/10.1093/molbev/msl203>
- Pavan, W. J., & Sturm, R. A. (2019). The genetics of human skin and hair pigmentation. *Annual Review of Genomics and Human Genetics*, 20(1), 41–72. <https://doi.org/10.1146/annurev-genom-083118-015230>
- Pierce, B. L., Tong, L., Argos, M., Demanelis, K., Jasmine, F., Rakibuz-Zaman, M., Sarwar, G., Islam, M. T., Shahriar, H., Islam, T., Rahman, M., Yunus, M. D., Kibriya, M. G., Chen, L. S., & Ahsan, H. (2018). Co-occurring expression and methylation QTLs allow detection of common causal variants and shared biological mechanisms. *Nature Communications*, 9(1), 1–12. <https://doi.org/10.1038/s41467-018-03209-9>
- Qi, T., Wu, Y., Zeng, J., Zhang, F., Xue, A., Jiang, L., Zhu, Z., Kemper, K., Yengo, L., Zheng, Z., Marioni, R. E., Montgomery, G. W., Deary, I. J., Wray, N. R., Visscher, P. M., McRae, A. F., & Yang, J. (2018). Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nature Communications*, 9(1), 2282. <https://doi.org/10.1038/s41467-018-04558-1>
- Quillen, E. E., Norton, H. L., Parra, E. J., Lona-Durazo, F., Ang, K. C., Illiescu, F. M., Pearson, L. N., Shriver, M. D., Lasisi, T., Gokcumen, O., Starr, I., Lin, Y.-L., Martin, A. R., & Jablonski, N. G. (2019). Shades of complexity: New perspectives on the evolution and genetic architecture of human skin. *American Journal of Physical Anthropology*, 168(S67), 4–26. <https://doi.org/10.1002/ajpa.23737>
- Rees, J. R., Zens, M. S., Gui, J., Celaya, M. O., Riddle, B. L., & Karagas, M. R. (2014). Non melanoma skin cancer and subsequent cancer risk. *PLoS One*, 9(6), e99674. <https://doi.org/10.1371/journal.pone.0099674>

- Relton, C. L., & Davey Smith, G. (2012). Two-step epigenetic Mendelian randomization: A strategy for establishing the causal role of epigenetic processes in pathways to disease. *International Journal of Epidemiology*, 41(1), 161–176. <https://doi.org/10.1093/ije/dyr233>
- Roos, L., Sandling, J. K., Bell, C. G., Glass, D., Mangino, M., Spector, T. D., Deloukas, P., Bataille, V., & Bell, J. T. (2017). Higher nevus count exhibits a distinct DNA methylation signature in healthy human skin: Implications for melanoma. *Journal of Investigative Dermatology*, 137(4), 910–920. <https://doi.org/10.1016/j.jid.2016.11.029>
- Shen, J., Song, R., Wan, J., Huff, C., Fang, S., Lee, J. E., & Zhao, H. (2017). Global methylation of blood leukocyte DNA and risk of melanoma. *International Journal of Cancer*, 140(7), 1503–1509. <https://doi.org/10.1002/ijc.30577>
- Vandiver, A. R., Irizarry, R. A., Hansen, K. D., Garza, L. A., Runarsson, A., Li, X., Chien, A. L., Wang, T. S., Leung, S. G., Kang, S., & Feinberg, A. P. (2015). Age and sun exposure-related widespread genomic blocks of hypomethylation in nonmalignant skin. *Genome Biology*, 16(1), 1–15. <https://doi.org/10.1186/s13059-015-0644-y>
- Westra, H.-J., Peters, M. J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M. W., Fairfax, B. P., Schramm, K., Powell, J. E., Zhernakova, A., Zhernakova, D. V., Veldink, J. H., Van den Berg, L. H., Karjalainen, J., Withoff, S., Uitterlinden, A. G., Hofman, A., Rivadeneira, F., ... Franke, L. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genetics*, 45(10), 1238–1243. <https://doi.org/10.1038/ng.2756>
- Yuan, V., Price, E. M., Del Gobbo, G., Mostafavi, S., Cox, B., Binder, A. M., Michels, K. B., Marsit, C., & Robinson, W. P. (2019). Accurate ethnicity prediction from placental DNA methylation data. *Epigenetics & Chromatin*, 12(1), 51. <https://doi.org/10.1186/s13072-019-0296-3>
- Zeng, J., de Vlaming, R., Wu, Y., Robinson, M. R., Lloyd-Jones, L. R., Yengo, L., Yap, C. X., Xue, A., Sidorenko, J., McRae, A. F., Powell, J. E., Montgomery, G. W., Metspalu, A., Esko, T., Gibson, G., Wray, N. R., Visscher, P. M., & Yang, J. (2018). Signatures of negative selection in the genetic architecture of human complex traits. *Nature Genetics*, 50(5), 746–753. <https://doi.org/10.1038/s41588-018-0101-4>
- Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., Montgomery, G. W., Goddard, M. E., Wray, N. R., Visscher, P. M., & Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics*, 48(5), 481–487. <https://doi.org/10.1038/ng.3538>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Bonilla C, Bertoni B, Min JL, Hemani G, Elliott HR. Investigating DNA methylation as a potential mediator between pigmentation genes, pigimentary traits and skin cancer. *Pigment Cell Melanoma Res.* 2021;34:892–904. <https://doi.org/10.1111/pcmr.12948>