Epigenetic Variability across Human Populations: A Focus on DNA Methylation Profiles of the *KRTCAP3*, *MAD1L1* and *BRSK2* Genes

Cristina Giuliani[1,*], Marco Sazzini[1], Maria Giulia Bacalini[2,3], Chiara Pirazzini[2,3], Elena Marasco[2,3], Elisa Fontanesi[2], Claudio Franceschi[2,3,4], Donata Luiselli[1,†], and Paolo Garagnani[2,3,5,†]

[1]Department of Biological Geological and Environmental Sciences, Laboratory of Molecular Anthropology & Centre for Genome Biology, University of Bologna, Italy

[2]Department of Experimental, Diagnostic and Specialty Medicine (DIMES), University of Bologna, Italy

[3]Interdepartmental Center "L. Galvani" (C.I.G.), University of Bologna, Italy

[4]IRCCS Institute of Neurological Sciences, Bologna, Italy

[5]Center for Applied Biomedical Research (CRBA), St. Orsola-Malpighi University Hospital, Bologna, Italy

*Corresponding author: E-mail: cristina.giuliani2@unibo.it.

†These authors contributed equally to this work.

Abstract

Natural epigenetic diversity has been suggested as a key mechanism in microevolutionary processes due to its capability to create phenotypic variability within individuals and populations. It constitutes an important reservoir of variation potentially useful for rapid adaptation in response to environmental stimuli. The analysis of population epigenetic structure represents a possible tool to study human adaptation and to identify external factors that are able to naturally shape human DNA methylation variability. The aim of this study is to investigate the dynamics that create epigenetic diversity between and within different human groups. To this end, we first used publicly available epigenome-wide data to explore population-specific DNA methylation changes that occur at macro-geographic scales. Results from this analysis suggest that nutrients, UVA exposure and pathogens load might represent the main environmental factors able to shape DNA methylation profiles. Then, we evaluated DNA methylation of candidate genes (*KRTCAP3*, *MAD1L1*, and *BRSK2*), emerged from the previous analysis, in individuals belonging to different populations from Morocco, Nigeria, Philippines, China, and Italy, but living in the same Italian city. DNA methylation of the *BRSK2* gene is significantly different between Moroccans and Nigerians (pairwise *t*-test: CpG 6 P-value = $5.2*10^{-3}$; CpG 9 P-value = $2.6*10^{-3}$; CpG 10 P-value = $3.1*10^{-3}$; CpG 11 P-value = $2.8*10^{-3}$). Comprehensively, these results suggest that DNA methylation diversity is a source of variability in human groups at macro and microgeographical scales and that population demographic and adaptive histories, as well as the individual ancestry, actually influence DNA methylation profiles.

Key words: DNA methylation variability, population epigenetics, human adaptation, environmental interaction.

Introduction

Epigenetics is a reversible molecular mechanism that is supposed to play a role in many adaptive processes, along with the classical ways to adapt represented by genetic and/or, in the case of humans, cultural changes (Giuliani et al. 2015). Several theories have been proposed to explain the role that

epigenetic variation may have played in human evolutionary history (Gluckman et al. 2005; Jablonka & Lamb 2005; Feinberg & Irizarry 2009; Shea et al. 2011; Klironomos et al. 2013), overall describing it as a sort of plastic interface between the genome and the environment. DNA methylation, the first epigenetic mechanism discovered, constitutes a source of phenotypic variability and several lines of evidence

suggest that epigenetic modifications have the potential to influence evolutionary processes (Szyf 2015). In addition, many studies based on animal models observed that DNA methylation patterns can be transmitted for generations after the exposure to an environmental perturbation (i.e., toxins, hormone exposure, etc.), thus escaping transgenerational erasure mechanisms (Lane et al. 2003; Seisenberger et al. 2012). Changes in DNA methylation can be also lost more rapidly than genetic ones and, therefore, they require stronger selective pressures to be fixed in a population (Klironomos et al. 2013). According to this view, population epigenetics studies might represent a suitable tool for the study of human adaptation and for the identification of environmental factors able to shape DNA methylation variability across human groups. In the last years, studies that simultaneously investigated patterns of both population epigenetic and genetic structure have demonstrated that the individual's genetic background at specific loci influences DNA methylation profiles of distal and nearby genomic regions by remodelling the conformation of chromatin itself (Lemire et al. 2015). Therefore, the identification of genes and pathways whose DNA methylation profiles change across populations, together with the analysis of their genetic variability, could represent a valuable approach for pinpointing strong environmental pressures that acted (or that are still acting) on different human populations (Fraser et al. 2012).

A recent epigenome-wide study conducted by using Illumina 450k BeadChip investigated DNA methylation variability across human populations at a macro-geographical level (i.e., by considering human groups belonging to different continents, such as Africans, Europeans and Asians). From this study emerged that selective pressures not only shaped the genetic background of these populations, but also their DNA methylation profiles (Heyn et al. 2013). In this article, the authors identified population-specific DNA methylation patterns using a single CpG analysis. However, the biological relevance of fluctuations in DNA methylation levels of individual CpGs is still debated (Wessely & Emes 2012), although changes in DNA methylation in groups of adjacent CpG site—especially in the CpG islands where CpGs methylation levels are correlated—are more likely to have "a biological role", because they potentially affect chromatin structure. To our knowledge, no replication studies have been performed on different cohorts to confirm signals identified in such epigenome-wide analyses, as well as to provide further information on the surrounding CpGs sites.

Moreover, to date DNA methylation variability at a microgeographical level—considering variability inside each continent—has been poorly investigated and, to our knowledge, only one study evaluated the DNA methylation variability across rainforest hunter-gatherers and sedentary farmers from Central Africa identifying genomic regions whose methylation is affected by recent changes in habitat and by historical lifestyle (Fagny et al. 2015).

On the basis of the above considerations, the aim of the present study is 4-fold:

1) to exploit a new approach for the analysis of public epigenome-wide datasets based on blocks of adjacent CpGs (Bacalini et al. 2015) to identify hotspots of variability—that likely have a functional role—between different populations;

2) to confirm the loci identified in the previous step in an independent sample of individuals belonging to different macro-geographical areas and living in Italy, by means of a different technique (MALDI-TOF);

3) to assess DNA methylation variability at a microgeographical level, by considering variations among populations belonging to the same macro-geographic area (e.g., African and Asian groups);

4) to deepen the understanding of the genetic structure of selected candidate genes in order to identify factors that could impact on their epigenetic profiles.

For this purpose, firstly we performed a new analysis on the existing human populations data (i.e., of African, Asian and European ancestry) published by Heyn and colleagues (2013) to identify genomic regions and associated pathways with a high level of plasticity in terms of DNA methylation. These genomic regions were then analyzed at higher resolution by means of a MALDI-TOF technology (Sequenom, EpiTYPER protocol). A total of 90 individuals of European, Asian, and African ancestry, but all living in the same place (Bologna, Italy), were typed and, in particular, we selected 17 individuals from China, 13 from Philippines, 16 from Morocco, and 14 from Nigeria to investigate methylation variability at a microgeographical level. Patterns of population genetic structure for the selected candidate genes were also investigated using data from the 1000 Genomes Project (1000 Genomes Project Consortium 2010) to check the correlation between genetic and epigenetic population structures at these loci.

Materials and Methods

Samples Description

Already published data available in the public Gene Expression Omnibus (GEO) database under accession number GSE36369 were used for a meta-analysis. These data include DNA methylation levels measured in lymphoblastoid cell lines (LCLs) of 96 Americans of European ancestry (EU), 96 Americans of African ancestry (AFR) and 96 Han Chinese Americans (ASN). Mean age was 37.3 ± 16.2 years for EU, 29.4 ± 9.9 for AFR and 36.2 ± 15.7 for ASN. In order to replicate the obtained results, methylation data available from whole blood of 30 individuals of different ancestry were considered. We used the methylation data from whole blood samples of 10 Americans of European ancestry and 10 Asian Americans described in the study of Heyn and colleagues (2013) (GSE36369), and of 10 Americans of African ancestry retrieved from Alisch et al. (2012) (GSE36064).

Whole blood of 90 individuals from Morocco ($N = 16$), Nigeria ($N = 14$), Philippines ($N = 13$), China ($N = 17$) and Italy ($N = 30$) collected in Bologna (Italy) before 1997 and stored at $-20\,^{\circ}\text{C}$ was analysed. Ethical approval for this experiment was obtained from the ethics committee of the University of Bologna (permission to DL in date 11/11/2015). No data regarding health status or the time of migration to Italy are available for the subjects. The only available information is that they live in Bologna (Italy).

New Analysis of Public Datasets

To identify differentially methylated regions (DMRs) between the three examined populations we applied the bioinformatics pipeline described in Bacalini et al. (2015). In brief, this approach is based on the analysis of groups of adjacent CpGs because DNA methylation patterns of groups of CpGs are more likely associated with functional variation than DNA methylation changes of single CpGs. Accordingly, we divided the Infinium 450k probes into four classes: 1) Class A: probes in CpG islands and CpG islands-surrounding sequences (i.e., shores and shelves) that map in regions near genes, 2) Class B: probes in CpG islands and CpG islands-surrounding regions (i.e., shores and shelves) which do not map in genic regions, 3) Class C: probes in genic regions which are not CpG rich 4) Class D: probes in nongenic regions which are not CpG rich. With the term "blocks of probes" (BOPs), we refer to groups of Class A and Class B CpG probes localized in the same island/shore/shelf. Class C and Class D were instead analysed by a single CpG approach because these CpGs are too distant to support an analysis based on groups of adjacent probes.

Ingenuity Pathway and Network Analyses

Pathway and network analyses were performed using Ingenuity Pathway Analysis (IPA, Qiagen) by considering the list of genes associated with Class A BOPs for each pairwise comparison (i.e., AFR-EU, AFR-ASN, ASN-EU). We decided to use these lists because Class A probes are located in genic regions and are more likely associated to phenotypic changes. For all the analyses conducted using the IPA software, we reported statistically significantly enriched canonical pathways together with the related P-values and the number of genes involved in each pathway. Canonical pathways give a wealth of information regarding what is known to occur at the cellular level. In supplementary materials, molecular and cellular functions were also reported, as well as the most significant networks with their scores. The score is a numerical value used to rank networks according to their degree of relevance to the Network Eligible molecules in the dataset. The network Score is based on the hypergeometric distribution and is calculated with the right-tailed Fisher's Exact Test (Calvano et al. 2005). Comparison analyses were performed using IPA and by considering the aforementioned pairwise population groups comparisons. This approach allowed the identification of pathways that are unique or shared between each comparison and Fisher's Exact Test P-values were reported for each of them. The x-axis represents negative log P-values based on the probability that molecules identified were included in the predefined IPA canonical pathways by true association as opposed to inclusion of molecules based on chance alone.

Selection of Candidate Genes

BOPs of Class A identified in the three performed pairwise comparisons were then analysed by considering a dataset composed of 10 individuals of African origin, 10 individuals of Asian origin and 10 individual of European origin. Absolute pairwise differences between whole blood DNA methylation of these three groups were calculated and the ANOVA test was performed. The most significant CpGs selected according to the calculated P-values, to the DNA methylation differences in LCLs and in the dataset of whole blood samples described in the study of Heyn and colleagues (Heyn et al. 2013) were thus selected for further analyses. These CpGs sites are cg21248554 (grch37/hg19 chr2: 27,665,151), cg16658412 (grch37/hg19 chr7:1,883,420) and cg15465743 (grch37/hg19 chr11: 1,413,145) that mapped on the KRTCAP3, MAD1L1 and BRSK2 genes, respectively. These sites are also included in the list of "pop-CpGs" described by Heyn and colleagues (Heyn et al.2013).

DNA Extraction and Bisulfite Treatment

Genomic DNA was extracted from 150 µl of whole peripheral blood using the QIAamp 96 DNA Blood Kit (QIAGEN, Hilden, Germany). For Sequenom EpiTYPER assay, 1,000 ng of DNA were bisulfite-converted using the EZ-96 DNA Methylation Kit (Zymo Research Corporation, Orange, CA) with the following modifications of the manufacture's protocol: bisulfite conversion was performed with thermal conditions that repeatedly varied between $55\,^{\circ}\text{C}$ for 15 min. and $95\,^{\circ}\text{C}$ for 30 s for a total of 21 cycles; after the desulfonation and the cleaning steps, bisulfite-treated DNA was eluted in 100 µl of water.

EpiTYPER Assay on MALDI-TOF Platform

Quantitative analysis of methylation status of CpG sites in candidate genes was performed by the EpiTYPER assay (Agena Bioscience Inc., San Diego, CA previously Sequenom Inc.), a MALDI-TOF mass spectrometry-based method. Bisulfite-treated DNA was PCR-amplified and then processed following manufacturer's instructions. DNA methylation levels of 16 CpG sites nearby cg 21248554 (GRCh37/hg19 chr2: 27,664,896-27,665,359), 13 CpG sites nearby cg 16658412 (GRCh37/hg19 chr7: 1,883,322–1,883,759) and six CpG sites nearby cg 15465743 (GRCh37/hg19 chr11:1,413,109–1,413,442) were measured. The following bisulfite specific primers were used:

F-AGGAAGAGAGTTTGGTATTTGGTGTTAAGTGGTTT and R-CAGTAATACGACTCACTATAGGGAGAAGGCTAAAAACTA

ATCTCCACTCTTCATAACA for the *BRSK2* gene; F-AGGAAGA
GAGGGTAAGGGTAGTTTTAGGGTAAGGA and R- CAGTAATA
CGACTCACTATAGGGAGAAGGCTAAACCTAAACCTTCTCAA
CAACC for *KRTCAP3* and F-AGGAAGAGAGTGAAGATTTATTT
TTGGAGTGGGTA and R- CAGTAATACGACTCACTATAGGGA
GAAGGCTTAACACCAACCAAAACACACCTAA for the
MAD1L1 gene.

Statistical Analyses

Genome-wide DNA methylation values were in part obtained
by public database (GSE36369 and GSE36064). Color bias
adjustment, background level adjustment and quantile nor-
malization across arrays were performed using the *lumi* pack-
age (Du et al. 2008). Probes on chromosome X and Y, probes
containing missing β-values and those showing SNPs with a
frequency > 1% within their sequence were removed, as sug-
gested in Heyn et al. (2013). For Class A and Class B, BOPs
methylation values were compared between groups by using
the MANOVA function implemented in the *car* R package.
BOPs containing one or two CpG probes were excluded
from the analysis and MANOVA was applied on sliding win-
dows of three consecutive CpGs within the same BOP. For
each BOP, the lowest P-value among those calculated for the
different sliding windows was retained. For Class C and Class
D, the methylation values of single CpG site was compared
between groups using the ANOVA function implemented in
the *car* R package. Benjamini–Hochberg False Discovery Rate
was computed to account for the adopted multiple testing
procedures and by using the function mt.rawp2adjp imple-
mented in the *multtest* R package.

For the candidate genes analysis, the *MassArray* R package
was used to test whether bisulfite conversion reactions run to
completion (Thompson et al. 2009). CpG sites with missing
values in more than 20% of samples were filtered out, to-
gether with samples with missing values in more than 20% of
CpG sites.

For the analysis of genetic structure at the *BRSK2*,
MAD1L1 and *KRTCAP3* genes, we considered genotypes
of individuals from 15 populations sequenced by phase 3
of the 1000 Genomes Project: five of East Asian origin
(i.e., Han Chinese in Beijing—CHB, Japanese in Tokyo—
JPT, Southern Han Chinese—CHS, Kinh in Ho Chi Minh
City—KHV), five of European origin (i.e., Utah Residents
with Northern and Western European ancestry—CEU,
Toscani in Italy—TSI, Finnish in Finland—FIN, British in
England and Scotland—GBR, Iberian population in Spain—
IBS), five of African origin (i.e., Yoruba in Ibadan, Nigeria—
YRI, Luhya in Webuye, Kenya—LWK, Gambian in Western
Divisions in The Gambia—GWD, Mende in Sierra Leone—
MSL, Esan in Nigeria—ESN). Data were filtered for allelic
state (i.e., only biallelic loci were retained), then a pruned
subset of SNPs in approximate linkage equilibrium with each
other was generated (Auton et al. 2015) and monomorphic

loci were excluded. Discriminant analysis of principal compo-
nents (DAPC) was applied allowing the description of homo-
geneous genetic clusters using few synthetic variables. For
the analysed genes, F_{st} values for each SNP were also esti-
mated according to the Wright formula.

Results

In Silico Epigenome-Wide Analyses

A recently developed region-centric approach based on the
identification of BOPs (Bacalini et al. 2015) was applied for
the analysis of publicly available data generated with the
Infinium 450k assay (Heyn et al. 2013) to detect popula-
tion-specific epigenomic signatures. First, we analysed a
panel of 288 samples (96 AFR, 96 ASN and 96 EU). After
quality check and filtering procedures, we retained for the
analysis 440,793 out 485,577 loci. The entire pipeline of
analysis is described in details in the materials and methods
section and an overview is reported in supplementary fig. S1,
Supplementary Material online.

According to the bioinformatics pipeline proposed by
Bacalini *et al*, we identified 77, 217 and 301 BOPs in Class
A that discriminated AFR and EU, AFR and ASN, EU and ASN
respectively (q-values $< 5*10^{-8}$). By considering these loci,
a multidimensional scaling was computed and reported in
fig. 1A. Then, methylation values of the top ranking re-
gions—according to lowest q-values and filtering only signif-
icant BOPs containing at least two adjacent differentially
methylated CpG sites—were reported in fig. 1B.

A total of 15, 43 and 122 BOPs belonging to Class B re-
sulted differentially methylated in the AFR/EU, AFR/ASN, and
EU/ASN comparisons, respectively (q-values $< 10^{-8}$).
Methylation values of probes belonging to Class C and Class
D were compared between the examined groups by ANOVA.
A total of 53, 161 and 1,815 CpG probes mapping in Class C
and 62, 107 and 1,559 mapping in Class D were proved to be
differentially methylated between AFR and EU, AFR and ASN,
EU and ASN, respectively (q-values $< 10^{-8}$). Frequencies of
differentially methylated BOPs/CpGs are reported in supple-
mentary table. 1S, Supplementary Material online.

The list of Class C probes was used to perform a single
CpG-based analysis. As regards the AFR-EU Class C DMRs,
we identified two genes that were differentially methylated
in more than one CpG: the *PRDM16* locus, which codes for a
protein involved in the differentiation of brown adipose tissue,
and *MLPH* that codes for melanophilin, a protein found in
pigment producing cells such as melanocytes. The observed
AFR-ASN Class C DMRs were instead located in genes in-
volved in UDP glucuronosyltransferase. In particular, the
CpGs cg07952421 and cg10632656 are located in correspon-
dence of the binding site of the transcription factor POLR2A
that is responsible for messenger RNA synthesis located up-
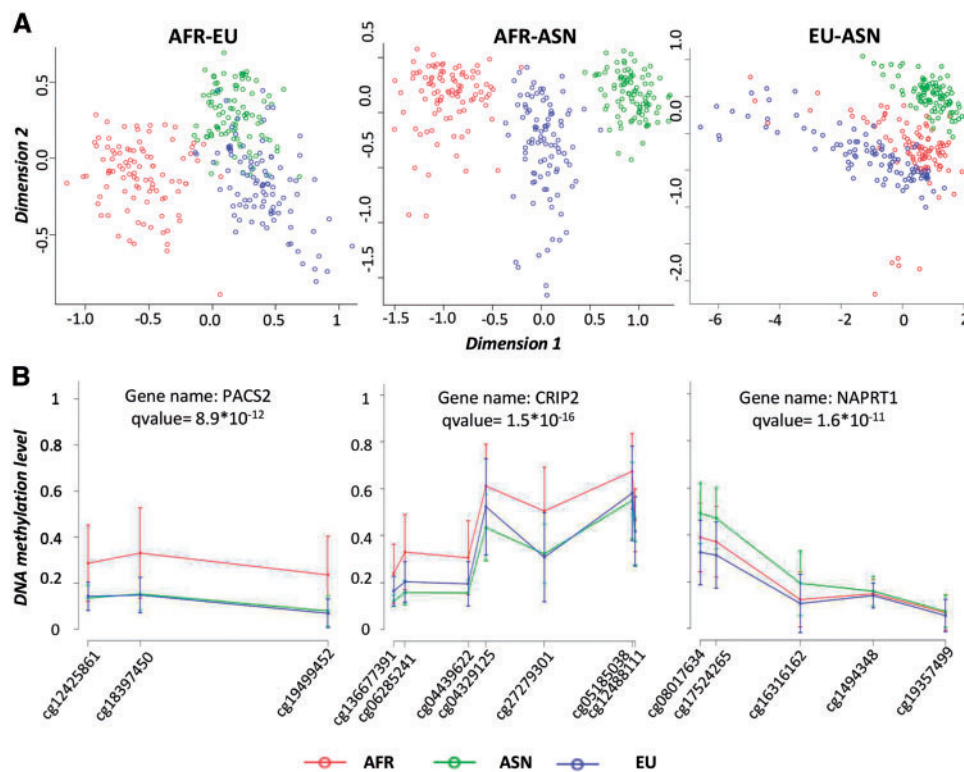stream the *UGT2B17*. Similarly, the ASN-EU comparison

FIG. 1.—Multidimensional scaling (MDS) considering the BOPs in Class A that discriminated AFR and EU, AFR and ASN, EU and ASN, respectively (A). Line plot of methylation values of the first top ranking regions— according to lowest q-values and filtering only significant BOPs containing at least two adjacent differentially methylated CpGs (B).

pointed to Class C DMRs related to genes involved in cellular glucoronidation (Tukey & Strassburg 2000).

Because Class A BOPs map in CpG islands and surrounding regions associated to genes, variation in their methylation status is more likely to have phenotypic consequences. Therefore, candidates to be further characterized with a targeted gene approach were selected from Class A BOPs. According to Class A BOPs, we reported 40 genes (supplementary table 2S, Supplementary Material online) that resulted differentially methylated between populations in both our region-centric approach and in the paper published by Heyn and colleagues (Heyn et al. 2013) that performed a classical single site analysis.

Pathway, Network and Gene Analyses

We performed pathway and network analyses by considering the list of Class A BOPs differentially methylated between the studied populations. Pathway analysis performed on the 77 genes corresponding to the AFR-EU DMRs identified molecular and cellular functions linked to cell morphology, cell death, cellular assembly, and organization (supplementary fig. 2S-A, Supplementary Material online). In particular, results reported in supplementary fig. 2S-B, Supplementary Material online show enrichment of DMRs located in genes involved in gap junction signalling, 14-3-3-mediated signalling, GM-CSF signalling, remodelling of epithelial adherens junction, as well as in the cardiac hyperthrophy regulated by NFAT and in the assembly of RNA polymerase I complex. Network analysis generated a network including 40 genes of the examined list (supplementary fig. 2S-C, Supplementary Material online).

AFR-ASN DMRs are located in the *MAD1L1* gene, one of the accelerated regions (HAR3) typical of the human lineage (Pollard et al. 2006; Hubisz & Pollard 2014). As displayed in supplementary fig. 3S-B, Supplementary Material online, most of the AFR-ASN DMRs mapped in genes involved in cellular immune response pathways, such as those related to antigen presentation, OX40 signalling, which has been proved to be important in T cell priming and cytokine production, Cd42 signalling, and antiviral innate immunity mediated by RIG1-like receptors. Also genes involved in the NF-kB activation by virus were affected by methylation changes. Network analysis identified three main networks that include 25, 23, and 23 genes, respectively (supplementary fig. 3S-A, Supplementary Material online).

Finally, EU-ASN DMRs predominantly mapped on genes involved in glycosaminoglycans biosynthesis (Dermatan Sulfate Biosynthesis, Chondroitin Sulfate Biosynthesis) (supplementary fig. 4S-B, Supplementary Material online), as well as

in genes, such as *CAPN5*, *NOS1*, *PPP3CC*, *PRKCZ*, involved in nNOS signalling and in the superpathway of citrulline metabolism.Pathway analysis also revealed enrichment of genes involved in nutrient sensing mediated by G-protein-coupled receptor (GPCR) of the enteroendocrine cells. Specific DNA methylation signatures were observed in the UVA-Induced MAPK signalling, proline degradation and intrinsic prothrombin activation pathways. Network analysis revealed a number of plausible networks (supplementary fig. 4S-A, Supplementary Material online), four of which (network 1, 3, 4, 5) shared many genes and pathways, although network 2 included genes involved mainly in dermatological characteristics. Network 1 is reported in supplementary fig. 4S-C, Supplementary Material online.

Comparative Analysis on Pathways Including Populations DMRs

To identify the most informative pathways among those described for each population, we applied a comparative and integrative analysis. In fig. 2, the pathways identified according to each pairwise comparison (EU-ASN, AFR-ASN, and AFR-EU) are reported. The calculated significance (*P*-values) indicates the probability of association of molecules identified in each pair of comparisons to the canonical pathway by random chance alone. We observed that some pathways are significant in all the comparisons and that a subset of them may be considered as a signature related to local adaptations, such as in the case of NF-kB activation by viruses (supplementary fig. 5S, Supplementary Material online), GPCR-mediated nutrient sensing in enteroendocrine cells (supplementary fig. 6S, Supplementary Material online), UVA-induced MAPK signalling (supplementary fig. 7S, Supplementary Material online).

Comparative analysis identified 11 pathways typical of the ASN populations, such as dermatan sulfate biosynthesis (and the late stage), chondroitin sulfate biosynthesis (and the late stage), PTEN signalling, OX40 signalling, and the antigen presentation. Only two pathways turned out to be typical of AFR populations, that is, the activation of IRF by the cytosolic pattern recognition receptor and the role of RIG1-like receptors in antiviral innate immunity. Finally, no pathways typical of EU populations emerged, although the possibility that this result is due to the reduced number of considered genes ($N = 77$, AFR-EU) cannot be excluded.

Candidate Gene DNA Methylation Variability between Macro-Geographic Groups

The three most significant CpGs (according to Class A BOPs *P*-values and differences in DNA methylation levels) observed in the LCL dataset, in the whole blood dataset and in the study of Heyn et al. (2013) were selected for further analyses. DNA methylation levels of 16 CpG sites nearby cg21248554 (*KRTCAP3* gene, GRCh37/hg19 chr2:27,665,151), 13 nearby cg16658412 (*MAD1L1* gene, GRCh37/hg19 chr

7:1,883,420) and six nearby cg15465743 (*BRSK2* gene, GRCh37/hg19 chr11:1,413,145) were thus measured in a new sample group composed of 17 Chinese, 13 Philippines, 16 Moroccans, 14 Nigerians and 30 Italians recruited in Italy (see materials and methods section for details). The observed DNA methylation patterns were compared across populations for each genomic region and the average values were calculated for individuals of Asian, European, and African ancestry (fig. 3).When considering the region located in the *KRTCAP3* gene, DNA methylation values of the CpGs sites nearby cg21248554 showed a concomitant hypomethylation in individuals of African origins, but not in people of European or Asian ancestry (fig. 3A). Individuals of African origins presented lower values of DNA methylation than subjects of European and Asian ancestry also at the region located in the *MAD1L1* gene, as well as for CpGs sites nearby cg16658412 (fig. 3B). DNA methylation levels of CpGs located in the *BRSK2* gene were lower in individuals of Asian origin than in Europeans and Africans (fig. 3C). A detailed description of significant sites for each pair of population comparisons is reported in fig. 3.

Candidate Gene DNA Methylation Variability within Macro-Geographic Groups

To investigate differences within each area (i.e., at the microgeographic level), we analysed the diversity of individuals who belong to different African populations (i.e., Moroccans and Nigerians) and to different Asian populations (i.e., Chinese and Philippines).Pairwise *t*-test between individuals from Morocco and Nigeria, as well as between individuals from China and Philippines were performed by considering DNA methylation levels of the candidate regions located in the *KRTCAP3*, *MAD1L1* and *BRSK2* genes. In particular, *KRTCAP3* and *MAD1L1* showed similar DNA methylation values within macro-geographic groups, with *t*-tests not revealing statistically significant differences. On the contrary, the *BRSK2* gene showed differences between individuals of Moroccan and Nigerian ancestries (pairwise *t*-test: CpG 6 *P*-value = 5.2×10^{-3}; CpG 9 *P*-value = 2.6×10^{-3}; CpG 10 *P*-value = 3.1×10^{-3}; CpG 11 *P*-value = 2.8×10^{-3}). In fig. 4, mean DNA methylation levels of all populations were reported according to their microgeographic category. .

Population Genetic Structure at the Identified Candidate Genes

Because the possible genetic influence on DNA methylation levels, we investigated patterns of population genetic structure at the three identified candidate genes by taking advantage from the whole genome data produced for 15 worldwide human groups by the 1000 Genomes Project (Auton et al. 2015). Indeed, elucidating these patterns could be crucial to disentangle the role of genetics in driving the differences observed in DNA methylation levels. Genomic localizations and
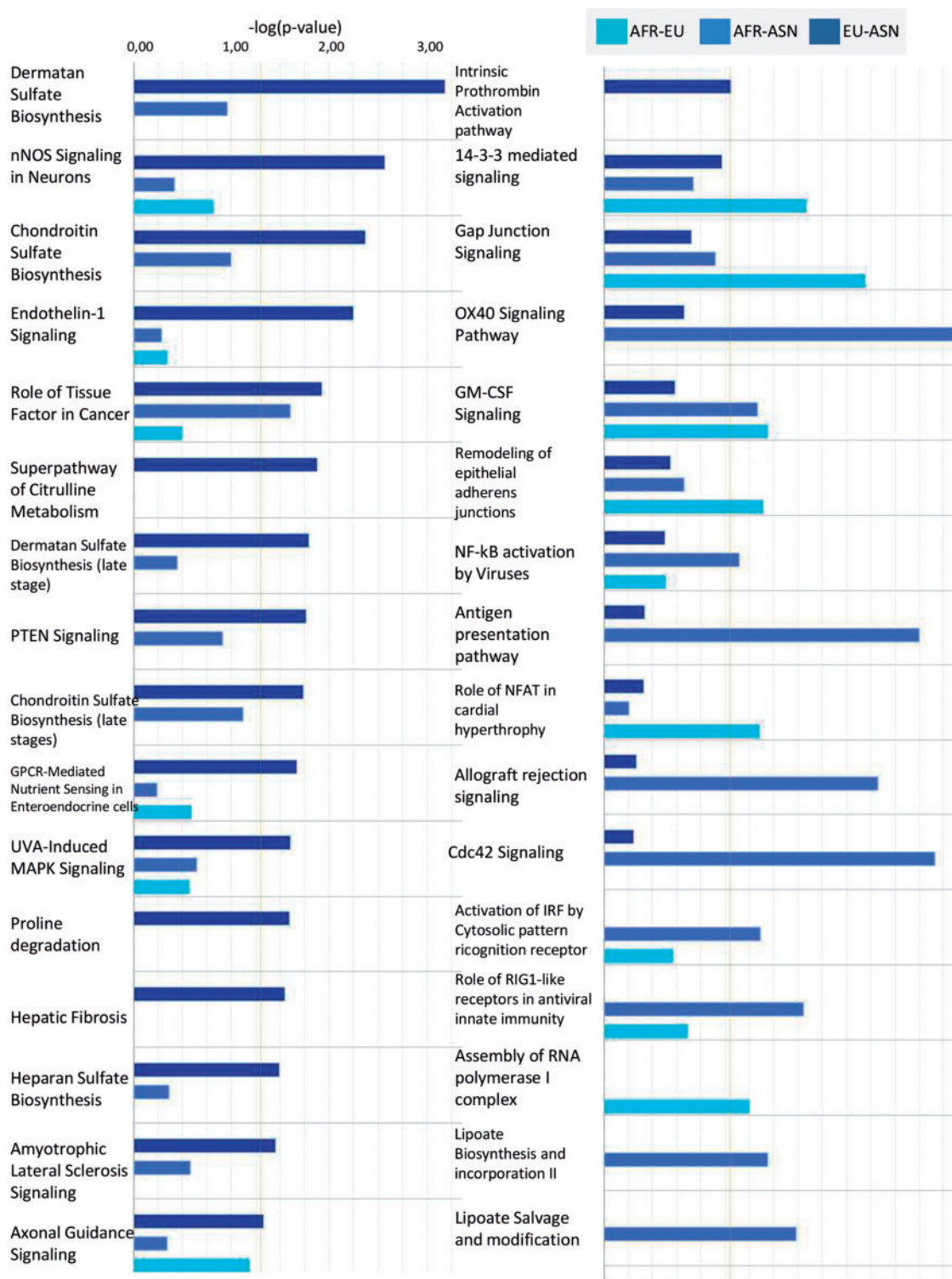
FIG. 2.—Comparative pathway analysis considering the three pairwise comparisons between AFR-EU (light blue), AFR-ASN (medium blue) and EU-ASN (dark blue). The bar indicate -log(P-value) for each comparison. P-value indicates the probability of association of molecules identified in each pair of comparisons to the canonical pathway by random chance alone.
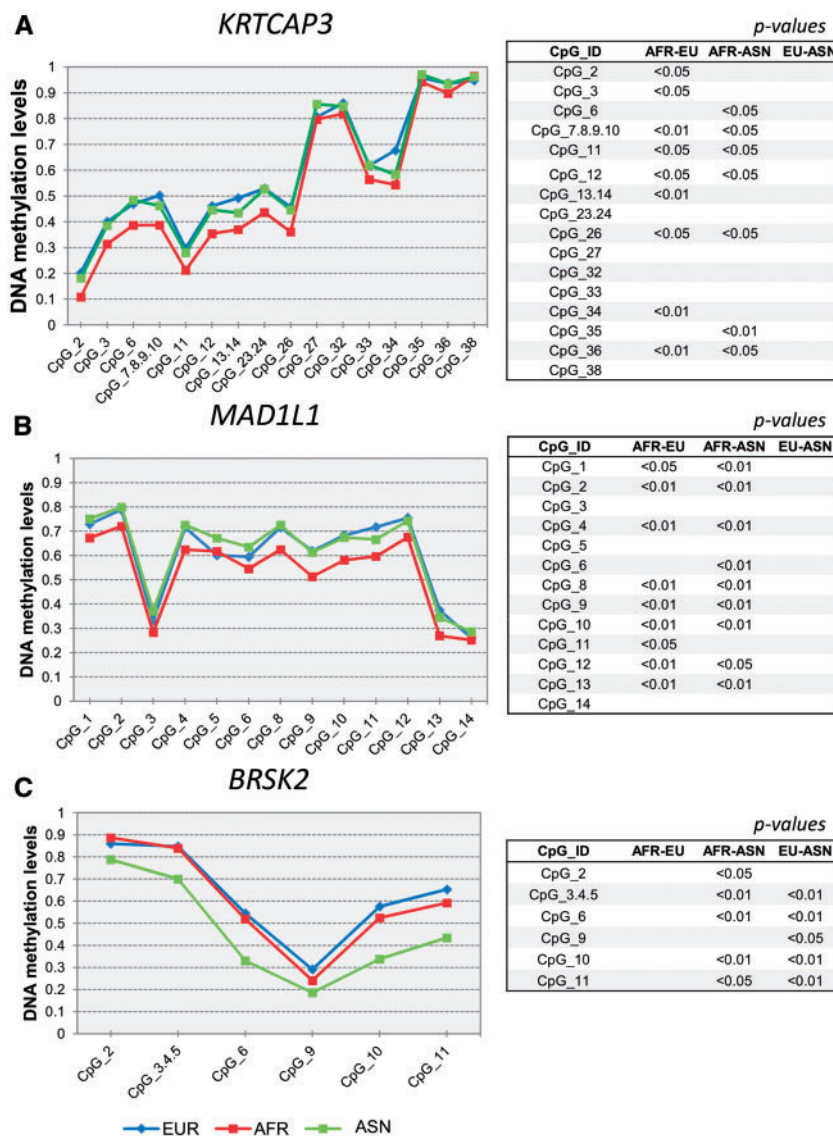
Fig. 3.—Line plots of average DNA methylation levels at *BRSK2*, *KRTCAP3* and *MAD1L1* calculated for each macro-geographic group of populations, DNA methylation of African individuals were reported in red, of Europeans in blue and of individuals of Asian origin in green. For each CpGs sites of each region, *P*-values (pairwise *t*-test) for AFR-EU, AFR-ASN and EU-ANS were reported.

numbers of SNPs considered in the analyses are listed in [supplementary table 3S, Supplementary Material](#) online.

As regards the *BRSK2* gene, DAPC suggested the presence of three distinct population clusters, one including African groups, one composed by individuals of European ancestry, and one entailing East Asians (fig.5A). We then selected the SNPs with the highest F_{st} values between populations of European ancestry and population of nonEuropean ancestry, in the attempt to disentangle the causes of the differences in DNA methylation patterns described in fig. 4. When F_{st} values were computed for all possible EU/nonEU comparisons, only two SNPs (rs61869028 and rs34805614) turned out to significantly differentiate these population groups (AFR-EU

$F_{st} = 0.23$, *P*-value < 0.01 and 0.07, respectively; ASN-EU $F_{st} = 0.33$ and 0.26, respectively, *P*-values < 0.01).

DAPC identified only two main population clusters when variation at the *KRTCAP3* gene was considered. It is to note that this subdivision is the result of the analysis of a modest number of SNPs in *KRTCAP3*, although the other two genes included a considerably higher number of genetic variants. The first group corresponded to African populations, although the second one included both Europeans and Asians (fig. 5B). This pattern is in agreement with the differences identified via epigenetics analysis, assigning individuals of African origin apart from European and East Asian ones. We calculated F_{st} values to search for the most informative variants responsible
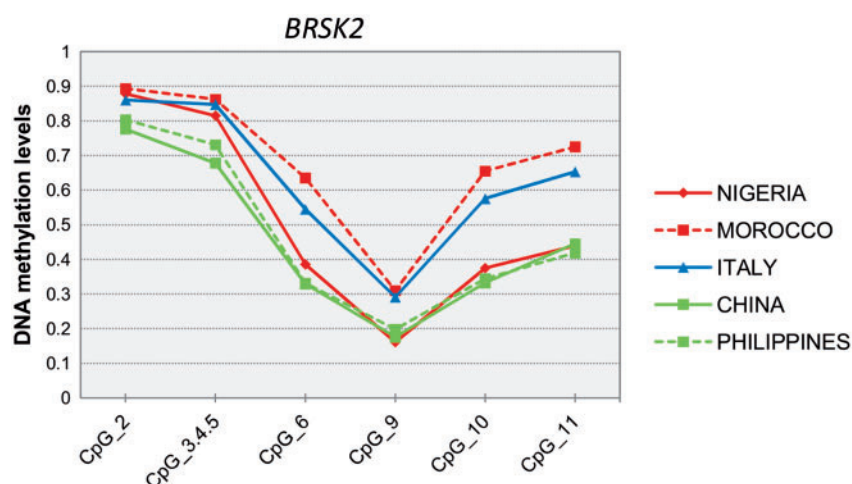
FIG. 4.—Average DNA methylation for the six CpGs analysed in the *BRSK2* gene. Red lines indicate individuals of African origin (from Nigeria and Morocco), blue line indicate Europeans (Italian individuals) and green lines indicate individuals of Asians origin (from China and Philippines).

for genetic differentiation at this locus between AFR and nonAfrican individuals. The SNPs that significantly differentiated these populations were rs780108, rs1260327, rs780105, rs780110, rs1728911, (F_{st} AFR-ASN 0.61, 0.60, 0.59, 0.53, 0.51 and *Fst* AFR-EU 0.28, 0.25, 0.26, 0.23, 0.21, respectively; *P*-value < 0.01).

For the *MAD1L1* gene (fig. 5C), DAPC identified three main clusters: the first included African populations, the second is made up of individuals of European ancestry, and the third included only East Asians. In particular, LD1 enabled to differentiate individuals of African ancestry from nonAfricans, as already suggested by DNA methylation patterns. The SNPs driving such differentiation were rs73288707, rs55934553, rs6968659, and rs3839699 (F_{st} AFR-ASN 0.29, 0.23, 0.36, 0.49 and F_{st} AFR-EU 0.28, 0.23, 0.35, 0.36, respectively; *P*-values < 0.01).

Discussion

This study explored population-specific DNA methylation changes that occur at both macro- and microgeographic scales. We applied a new bioinformatics pipeline to analyze publicly available epigenome-wide data, we selected the three so identified most significant CpGs and we analysed DNA methylation levels at these candidate genes in independent populations, in order to investigate the dynamics that create epigenetic diversity between and within different human groups.

Genome-Wide Analysis

We performed a pathway analysis on the most variable genes in terms of methylation levels, to get new hints on the environmental stimuli that have the potential to create epigenetic variability in humans. Accordingly, we reasoned that an

oscillation of DNA methylation levels of single CpGs could be less informative than variations at many CpGs located in functional elements, such as islands, shores or shelves. With this approach based on block of probes, we identified 40 loci that resulted differentially methylated between populations and in accordance to results presented by Heyn and colleagues (Heyn et al. 2013) based on a single probe analysis. These two different statistical approaches revealed a double validated subset of genes that more likely could be crucial in generating phenotypic variability in different human groups.

The most interesting results come from comparative analysis that suggested population-specific variations at pathways linked to diet, UVA exposure and pathogens load. It is important to note that, although DNA methylation is a tissue specific mechanism, some studies have actually addressed the correspondence between methylation levels of certain loci in immune cells and methylation levels in other tissues (Ai et al. 2012; Li et al. 2012; Marsit & Christensen 2013).

In all comparisons, we identified the pathway of GPCR-mediated nutrient sensing in enteroendocrine cells, the first level where stimuli from the gut lumen are detected. Some authors supposed that taste-signalling molecules of the gastrointestinal mucosa might participate in the functional detection of nutrients and harmful substances in the lumen, thus preparing the gut to absorb them or to initiate a protective response (Sternini et al. 2008). Moreover, these cells participate in the circuit that starts from gastrointestinal tract until the activation of gut-brain axis (Sternini et al. 2008).

The UVA-induced MAPK signalling is another pathway affected by methylation changes across all comparisons, suggesting that a further level of adaptation to UVA exposure in different human populations could be mediated by DNA methylation changes. This is a fundamental finding because many genetic studies showed that pigmentation genes have
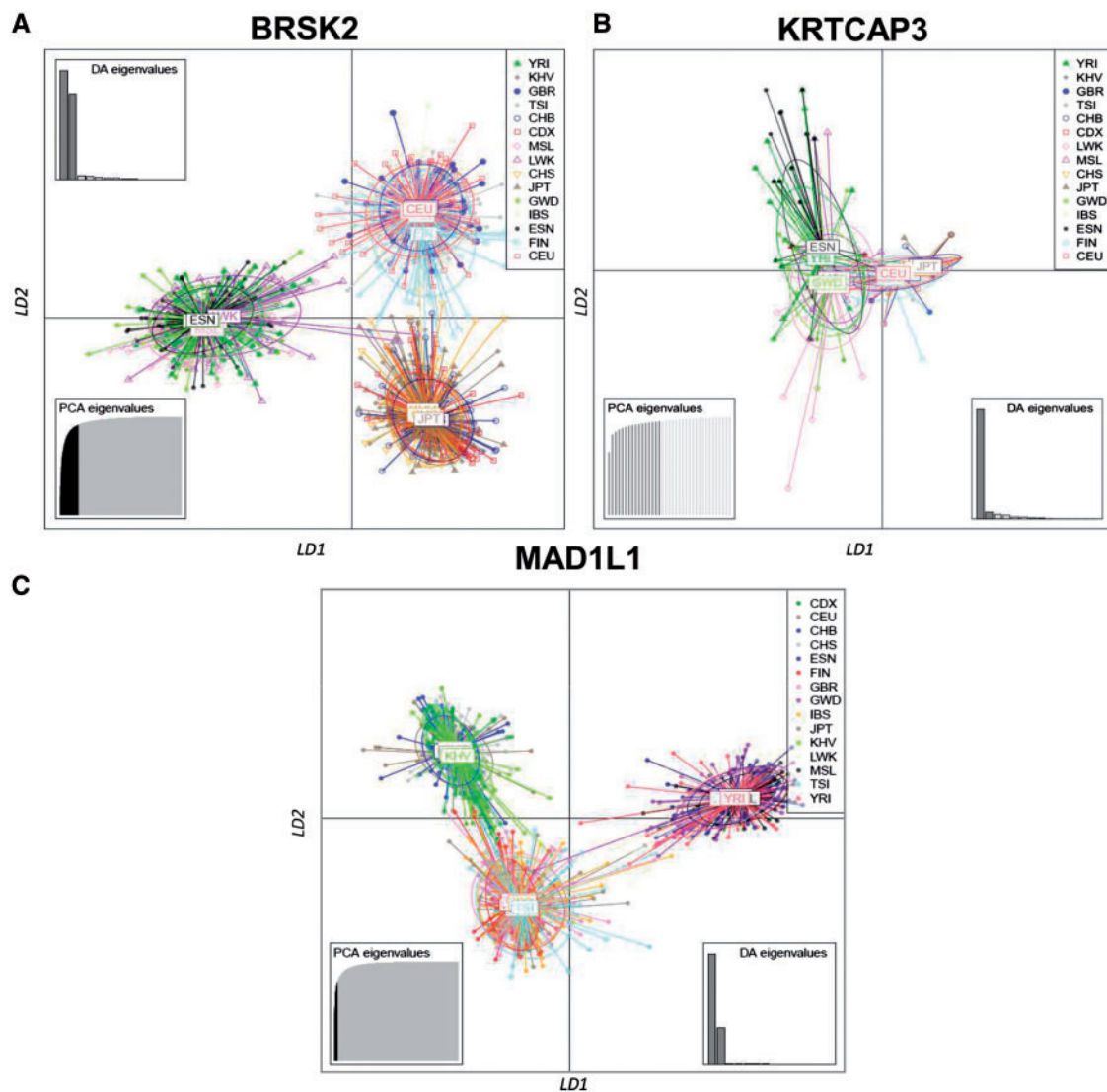
FIG. 5.—DAPC scatterplots related to the *BRSK2* (*A*), *KRTCAP3* (*B*) and *MAD1L1* (*C*) genes and computed by considering 15 populations of African, Asian and European origins. The plots represent individuals as dots and populations as inertia ellipses. Eigenvalues and principal components considered in the analysis are displayed in left and right squares.

undergone positive selection (Sturm & Duffy 2012). It has been suggested that dark skin pigmentation could have been selected as a protective trait against the deleterious effects of solar UV exposure in tropical climate areas. On the other side, lighter skin types may have been selected in the northern hemisphere due to low UVR exposure that allows the optimal level of vitamin D3 synthesis (Jablonski & Chaplin 2010; Arciero et al. 2015). However, this hypothesis is highly debated (Robins 2009) and some authors suggested that this is inconsistent with the fact that no pigmentation gene variants have been found to be associated with serum vitamin D levels in GWAS (Wang et al. 2010) or with blood folate concentrations (Tanaka et al. 2009). A possible additional interpretation of these observation could be consistent with the model proposed by Klironomos and colleagues (Klironomos et al. 2013), in which early adaptation via epigenetic mutations causes a reduction of selective constraints on genetic variables. The strength of selection operating on the genetic variants could be weakened because epigenetic variability in the same pathway can generate an additional level of phenotypic variation (Schmitz et al. 2011; Klironomos et al. 2013). These considerations are in line with recently published data that showed how combination of DNA methylation variation and genetic polymorphisms could improve the evolution of adaptive phenotypes (Fagny et al. 2015).

Pathway and network analyses showed that methylation levels of genes involved in response to pathogens change in many of the performed comparisons. When AFR and ASN

groups were compared, many DMRs resulted located in genes involved in OX40 signalling pathway, Cdc42 signalling, antigen presentation pathway and role of RIG1-like receptors in antiviral innate immunity, which turned out to be the most significant pathways. The ASN-EU comparison revealed instead significant pathways involved in the Dermatan Sulfate Biosynthesis (DS) and Chondroitin Sulfate Biosynthesis (CS). Despite the high number of functions of these pathways (e.g., central nervous system development, wound repair, infection, growth factor signalling, morphogenesis, and cell division, etc.), it is interesting to note that an increasing number of pathogens, including viruses, parasites, and bacteria, have been shown to use cell-surface DS for their attachment to host cells (Menozzi et al. 2002; Trowbridge & Gallo 2002). These results are in agreement with those obtained by Heyn and colleagues (Heyn et al. 2013) that identified many pop-CpGs in genes involved in immune response, indicating that their methylation patterns have been likely shaped by local pathogen landscapes. This issue was recently discussed also in Bierne et al. (2012). As pathogens exert a very strong influence on human genetic/epigenetic variation, some authors suggested that local pathogens diversity constitutes one of the main selective pressures having acted through human evolution (Fumagalli et al. 2011). We can hypothesize that epigenetic variation can provide a further level of variability important to counteract pathogens impact. Moreover, a recent study published by Marr and colleagues (Marr et al. 2014) elucidated the possible role of DNA methylation in pathogens response by demonstrating that intracellular parasites can alter the host cell DNA methylation patterns, altering gene expression that could lead to disease conditions. An example is represented by the uropathogenic *Escherichia Coli* that, in uroepithelial cells, can induce down-regulation of *CDKN2A* (a cell cycle inhibitor), thus potentially increasing cell proliferation and pathogens persistence (Tolg et al. 2011). However, the exact mechanism by which DNA methylation contributes to pathogens response remains unclear and we do not have enough data to understand whether DNA methylation variability represents a way to adapt to pathogens infection (in order to maintain the survival of the organism) or a way for the pathogens to maintain their survival. We also identified NF-kB activation by virus as a significant pathway including genes whose level of DNA methylation change in all the performed comparisons. It is noteworthy that this pathway is activated by different virus, such as EBV, HIV1, HSV, HBV, and CMV (Santoro 2003; Schroder et al. 2004), which increased their diffusion in the last century shaping DNA methylation profiles and thus producing population specific patterns.

Candidate Gene Analysis

A candidate gene approach was applied to the three most robust signals emerged from genome-wide analyses: *BRSK2*, a Serine/threonine-protein kinase that is abundantly expressed in pancreatic islets and that plays a key role in polarization of neurons and axonogenesis, cell cycle progress, as well as in insulin secretion (Chen et al. 2012; Wang at al. 2012; Friemel et al. 2014); *KRTCAP3*, a keratinocyte-associated protein (Andres et al. 2009; George et al. 2011; Han et al. 2013; Khan et al. 2014); *MAD1L1*, a component of the mitotic spindle-assembly checkpoint associated with chromosomal stability, whose methylation levels have been demonstrated to vary according to treatment with phytoestrogens, which are naturally present in a high number of edible plants, such as lupin, fava beans, soybeans, kudze, and psoralea (Karsli-Ceppioglu et al. 2015).

In details, hypomethylation at *KRTCAP3* and *MAD1L1* genes was observed in African individuals, although hypomethylation at *BRSK2* characterized East Asian subjects. According to this scenario, we demonstrated that differences between populations affect a larger gene region and not only one CpG site, further revealing that DNA methylation differences between populations are extended also to the CpGs flanking those implemented in the Illumina 450k BeadChip.

Notably, in our replication study we used whole blood samples for methylation analysis. Although we cannot completely exclude influence of heterogeneous cell composition of whole blood on observed methylation patterns, we noted that estimated methylation levels are similar to those reported by Heyn and colleagues (Heyn et al. 2013) and measured in LCLs.

DNA methylation differences at the *KRTCAP3*, *MAD1L1* and *BRSK2* genes were observed in individuals living in Bologna (Italy), as well as in the study published by Heyn and colleagues that includes individuals living in America (Heyn et al. 2013). This indicates that ancestry is the principal factor that influences DNA methylation profiles at these loci. The different environments (American and Italian city) seems not to have a crucial role in shaping DNA methylation profiles of these regions. Although more data on gene expression and protein analysis are needed to draw further conclusions, these findings strongly suggest that the examined genomic regions have high probability to have exerted an important role in the recent evolutionary history of human populations.

Moreover, the examined populations allowed us to perform comparisons aimed at investigating epigenetic variability within macro-geographic areas (i.e., Africa, Asia). Noteworthy, *BRSK2* methylation showed differences between individuals from Morocco and from Nigeria (pairwise *t*-test: CpG 6 P-value$=5.2*10^{-3}$; CpG 9 P-value$=2.6*10^{-3}$; CpG 10 P-value$=3.1*10^{-3}$; CpG 11 P-value$=2.8*10^{-3}$), pointing to high methylation diversity between populations belonging to the same continent, even if characterized by appreciably different demographic histories. A possible hypothesis for the observed pattern includes the European influence (in terms of both cultural and genetic admixture) evident on people from Morocco (and vice versa) (Henn et al. 2012; Botigué et al. 2013). To our knowledge, this level of variability

has never been addressed before because information regarding the countries of origin are missing in most studies.

It is noteworthy that this peculiar methylation pattern has been observed only for the *BRSK2* region and not for *MAD1L1* and *KRTCAP3* genes. This could indicate that microgeographic dynamics differently shaped some genomic regions according to their biological function. This variability among African populations is in line with the recent findings of Fagny and colleagues (Fagny et al. 2015) that investigate the effect of changes in subsistence strategies and ecological habitats on methylation profiles by considering Central African populations.

In the study by Heyn and colleagues (Heyn et al. 2013), a high proportion of genetic polymorphisms associated with methylation profiles were identified, suggesting a close link between genetic variability and DNA methylation values, at least for one third of the CpGs that were differentially methylated across populations. However, a limit of that study and of the present one was that no deep genomic characterization (i.e., by sequencing) of these individuals was available. In the attempt to partially overcome this issue, we have taken advantage of whole genome sequence data provided by the 1000 Genomes Project to explore full genetic variation at the three identified candidate genes in 15 human populations belonging to the same continents of origin of the examined subjects. Elucidating these patterns could be crucial to disentangle the role of genetics in driving the differences observed in DNA methylation levels, although the individuals considered in the present analysis belong to two different cohorts.

Appreciable population genetic structure was observed for the *KRTCAP3* and *MAD1L1* genes, potentially accounting also for the identified DNA methylation differences. However, Heyn and colleagues (Heyn et al. 2013) did not identify meQTL for the CpGs located in *MAD1L1* and *KRTCAP3*. This could be due to the fact that their data were generated by a microarray approach that captures only a subset of genetic variation actually present in the genome (i.e., generally common SNPs initially discovered in populations of European origins). Our results suggest that the potential role of the genetic context nearby the pop-CpGs on chromatin structure, and therefore on DNA methylation profiles, cannot be ruled out. Instead, population genetic structure at *BRSK2* seems to be different. In the cohort of samples living in Italy, individuals from Nigeria and Asia shared similar values of DNA methylation, although three main clusters made up of Asian, African and European individuals could be identified from a genetic perspective, suggesting that DNA methylation profiles of *BRSK2* cannot be explained by the genetic structure of the region alone. A panel of genetic variants with high F_{st} values that might contribute to the observed epigenetic variability was thus identified for the three genes. Further studies aimed at considering levels of both genetic and epigenetic variability, together with gene expression data, are needed to elucidate these complex dynamics.

In conclusion, we identified functional pathways whose methylation levels vary across different human groups and we pinpointed nutrients, UVA exposure and pathogens load as some of the main environmental stimuli that are able to shape DNA methylation profiles of human populations. Moreover, the analysis of three candidate genes has shown a further level of epigenetic variability, also within a single continent (i.e., Africa), suggesting that different demographic and adaptive histories, as well as peculiar ecological niches, actually influence individuals DNA methylation profiles together with their genomic background. The extent of the transgenerational inheritance of the observed diversity is largely unknown. Recent studies showed how certain epigenetic variants could be transmitted across generations (Lane et al. 2003; Seisenberger et al. 2012) but to date, further efforts must be devoted to the study of cases of epigenetic inheritance and to characterize the underlying mechanisms. However, it is well known that environmental changes could modulate DNA methylation profiles in a population specific way with the possibility of an indirect effect on the population genetic background (e.g., mutation rates, local transposition or recombination rates) (Richards 2008).

Supplementary Material

Supplementary figures S1–S7 and tables S1–S3 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

Literature Cited

Ai S, Shen L, Guo J, Feng X, Tang B. 2012. DNA methylation as a biomarker for neuropsychiatric diseases. Int J Neurosci. 122:165–176. doi: 10.3109/00207454.2011.637654.

Alisch RS, et al. 2012. Age-associated DNA methylation in pediatric populations. Genome Res. 22:623–632. doi: 10.1101/gr.125187.111.

Andres AM, et al. 2009. Targets of Balancing Selection in the Human. Genome. Mol Biol Evol. 26:2755–2764. doi: 10.1093/molbev/msp190.

Arciero E, et al. 2015. Genes regulated by Vitamin D in bone cells are positively selected in East Asians. Plos One 10(12):e0146072. doi: 10.1371/journal.pone.0146072.

Auton A, et al. 2015. A global reference for human genetic variation. Nature 526:68–74. doi: 10.1038/nature15393.

Bacalini MG, et al. 2015. A meta-analysis on age-associated change in blood DNA methylation: results from an original analysis pipeline for Infinium 450k data. Aging 7(2):97–109.

Bierne H, Hamon M, Cossart P. 2012. Epigenetics and bacterial infections. Cold Spring Harb Perspect Med. 2:a010272. doi: 10.1101/cshperspect.a010272.

Botigué LR, et al. 2013. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. Proc Natl Acad Sci U S A. 110:11791–11796. doi: 10.1073/pnas.1306223110.

Calvano SE, et al. 2005. A network-based analysis of systemic inflammation in humans. Nature 437:1032–1037. doi: 10.1038/nature03985.

Chen X-Y, et al. 2012. Brain-selective kinase 2 (BRSK2) phosphorylation on PCTAIRE1 negatively regulates glucose-stimulated insulin secretion in pancreatic β-cells. J Biol Chem. 287:30368–30375. doi: 10.1074/jbc.M112.375618.

Consortium T, 1000 GP. 2010. A map of human genome variation from population-scale sequencing. Nature 467:1061–1073. doi: 10.1038/nature09534.

Du P, Kibbe WA, Lin SM. 2008. lumi: a pipeline for processing Illumina microarray. Bioinforma Oxf Engl. 24:1547–1548. doi: 10.1093/bioinformatics/btn224.

Fagny M, et al. 2015. The epigenomic landscape of African rainforest hunter-gatherers and farmers. Nat Commun. 6:10047. doi: 10.1038/ncomms10047.

Feinberg AP, Irizarry RA. 2009. Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. Proc Natl Acad Sci U S A. 200906183. 107 Suppl 1:1757–1764. doi: 10.1073/pnas.0906183107.

Fraser HB, Lam LL, Neumann SM, Kobor MS. 2012. Population-specificity of human DNA methylation. Genome Biol. 13(2):R8.

Friemel C, et al. 2014. Array-based DNA methylation profiling in male infertility reveals allele-specific DNA methylation in PIWIL1 and PIWIL2. Fertil Steril. 101:1097–1103.e1. doi: 10.1016/j.fertnstert.2013.12.054.

Fumagalli M, et al. 2011. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. PLoS Genet. 7:e1002355. doi: 10.1371/journal.pgen.1002355.

George RD, et al. 2011. Trans genomic capture and sequencing of primate exomes reveals new targets of positive selection. Genome Res. 21:1686–1694. doi: 10.1101/gr.121327.111.

Giuliani C, et al. 2015. The epigenetic side of human adaptation: hypotheses, evidences and theories. Ann Hum Biol. 42:1–9. doi: 10.3109/03014460.2014.961960.

Gluckman PD, Hanson MA, Spencer HG. 2005. Predictive adaptive responses and human evolution. Trends Ecol Evol. 20:527–533. doi: 10.1016/j.tree.2005.08.001.

Han M, Fan L, Qin Z, Lavingia B, Stastny P. 2013. Alleles of keratin 1 in families and populations. Hum Immunol. 74:1453–1458. doi: 10.1016/j.humimm.2013.05.003.

Henn BM, et al. 2012. Genomic ancestry of North Africans Supports Back-to-Africa Migrations Schierup. PLoS Genet. 8:e1002397. doi: 10.1371/journal.pgen.1002397.

Heyn H, et al. 2013. DNA methylation contributes to natural human variation. Genome Res. 23:1363–1372. doi: 10.1101/gr.154187.112.

Hubisz MJ, Pollard KS. 2014. Exploring the genesis and functions of Human Accelerated Regions sheds light on their role in human evolution. Curr Opin Genet Dev. 29:15–21. doi: 10.1016/j.gde.2014.07.005.

Jablonka E, Lamb MJ. 2005. Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life. Cambridge, MA: MIT Press.

Jablonski NG, Chaplin G. 2010. Human skin pigmentation as an adaptation to UV radiation. Proc Natl Acad Sci U S A. 107:8962–8968. doi: 10.1073/pnas.0914628107.

Karsli-Ceppioglu S, et al. 2015. Genome-wide DNA methylation modified by soy phytoestrogens: role for epigenetic therapeutics in prostate cancer? OMICS J Integr Biol. 19:209–219. doi: 10.1089/omi.2014.0142.

Khan I, et al. 2014. Mammalian keratin associated proteins (KRTAPs) subgenomes: disentangling hair diversity and adaptation to terrestrial and aquatic environments. BMC Genomics 15:779. doi: 10.1186/1471-2164-15-779.

Klironomos FD, Berg J, Collins S. 2013. How epigenetic mutations can affect genetic evolution: model and mechanism. BioEssays News Rev Mol Cell Dev Biol. 35:571–578. doi: 10.1002/bies.201200169.

Lane N, et al. 2003. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 35:88–93. doi: 10.1002/gene.10168.

Lemire M, et al. 2015. Long-range epigenetic regulation is conferred by genetic variation located at thousands of independent loci. Nat Commun. 6:6326. doi: 10.1038/ncomms7326.

Li L, et al. 2012. DNA methylation in peripheral blood: a potential biomarker for cancer molecular epidemiology. J Epidemiol Jpn Epidemiol Assoc. 22:384–394.

Marr AK, et al. 2014. Leishmania donovani infection causes distinct epigenetic DNA methylation changes in host macrophages. PLoS Pathog. 10:e1004419. doi: 10.1371/journal.ppat.1004419.

Marsit C, Christensen B. 2013. Blood-derived DNA methylation markers of cancer risk. Adv Exp Med Biol. 754:233–252. doi: 10.1007/978-1-4419-9967-2_12.

Menozzi FD, et al. 2002. Enhanced bacterial virulence through exploitation of host glycosaminoglycans. Mol Microbiol. 43:1379–1386.

Pollard KS, et al. 2006. Forces Shaping the Fastest Evolving Regions in the Human Genome. PLoS Genet. 2:e168. doi: 10.1371/journal.pgen.0020168.

Richards EJ., 2008 Population epigenetics. Curr Opin Genet Dev.. 18:221–226. doi: 10.1016/j.gde.2008.01.014.

Robins AH. 2009. The evolution of light skin color: role of vitamin D disputed. Am J Phys Anthropol. 139:447–450. doi: 10.1002/ajpa.21077.

Santoro MG. 2003. NEW EMBO MEMBER'S REVIEW: NF-kappaB and virus infection: who controls whom. Embo J. 22:2552–2560. doi: 10.1093/emboj/cdg267.

Schmitz RJ, et al. 2011. Transgenerational epigenetic instability is a source of novel methylation variants. Science 334:369–373. doi: 10.1126/science.1212959.

Schroder K, Hertzog PJ, Ravasi T, Hume DA. 2004. Interferon-γ: an overview of signals, mechanisms and functions. J Leukoc Biol. 75:163–189. doi: 10.1189/jlb.0603252.

Seisenberger S, et al. 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell. 48:849–862. doi: 10.1016/j.molcel.2012.11.001.

Shea N, Pen I, Uller T. 2011 Three epigenetic information channels and their different roles in evolution. J Evol Biol. 24:1178–1187. doi: 10.1111/j.1420-9101.2011.02235.x.

Sternini C, Anselmi L, Rozengurt E. 2008. Enteroendocrine cells: a site of 'taste' in gastrointestinal chemosensing. Curr Opin Endocrinol Diabetes Obes. 15:73–78. doi: 10.1097/MED.0b013e3282f43a73.

Sturm RA, Duffy DL. 2012. Human pigmentation genes under environmental selection. Genome Biol. 13:248. doi: 10.1186/gb-2012-13-9-248.

Szyf M. 2015. Nongenetic inheritance and transgenerational epigenetics. Trends Mol Med. 21:134–144. doi: 10.1016/j.molmed.2014.12.004.

Tanaka T, et al. 2009. Genome-wide association study of vitamin B6, vitamin B12, Folate, and homocysteine blood concentrations. Am J Hum Genet. 84:477–482. doi: 10.1016/j.ajhg.2009.02.011.

Thompson RF, Suzuki M, Lau KW, Greally JM. 2009. A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. Bioinformatics 25:2164–2170. doi: 10.1093/bioinformatics/btp382.

Tolg C, et al. 2011. Uropathogenic E. coli infection provokes epigenetic downregulation of CDKN2A (p16INK4A) in uroepithelial cells. Lab Investig J Tech Methods Pathol. 91:825–836. doi: 10.1038/labinvest.2010.197.

Trowbridge JM, Gallo RL. 2002. Dermatan sulfate: new functions from an old glycosaminoglycan. Glycobiology 12:117R–125R.

Tukey RH, Strassburg CP. 2000. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. Annu Rev Pharmacol Toxicol. 40:581–616. doi: 10.1146/annurev.pharmtox.40.1.581.

Wang TJ, et al. 2010. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. The Lancet 376:180–188. doi: 10.1016/S0140-6736(10)60588-0.

Wang Y, et al. 2012. BRSK2 is regulated by ER stress in protein level and involved in ER stress-induced apoptosis. Biochem Biophys Res Commun. 423:813–818. doi: 10.1016/j.bbrc.2012.06.046.

Wessely F, Emes RD. 2012. Identification of DNA methylation biomarkers from Infinium arrays. Front Genet. 3:(161)1–8.

Associate editor: Sabyasachi Das