

RESEARCH

Open Access



Identification of genetic and non-genetic modifiers of genomic imprinting through screening of imprinted DMR methylation in humans

Francesco Cecere^{1,2†}, Raissa Relator^{3,4†}, Michael Levy³, Ankit Verma¹, Haley McConkey^{3,4}, Bruno Hay Mele⁵, Laura Pignata¹, Carlo Giaccari¹, Emilia D'Angelo¹, Subham Saha¹, Abu Saadat¹, Angela Sparago¹, Claudia Angelini⁶, Flavia Cerrato¹, Bekim Sadikovic^{3,4*†} and Andrea Riccio^{1,2*†}

Abstract

Background Genomic imprinting is required for normal development, and abnormal methylation of differentially methylated regions (iDMRs) controlling the parent of origin-dependent expression of the imprinted genes has been found in congenital disorders affecting growth, metabolism, neurobehavior, and in cancer. In most of these cases the cause of the imprinting abnormalities is unknown. Also, these studies have generally been performed on a limited number of CpGs, and a systematic investigation of iDMR methylation in the general population is lacking.

Results By analysing a vast number of either in-house generated or online available whole-genome methylation array datasets of unaffected individuals, and patients with complex and rare disorders, we determined the most common iDMR methylation profiles in a large population and identified many genetic and non-genetic factors contributing to their variability in blood DNA. We found that methylation variability was not homogeneous within the iDMRs and that the CpGs closer to the ZFP57 binding sites are less susceptible to methylation changes. We demonstrated the methylation polymorphism of three iDMRs and the atypical behaviour of several others, and reported the association of 25 disease- and 47 non-disease-complex traits as well as 15 Mendelian and chromosomal disorders with iDMR methylation changes. The most significantly associated complex traits included ageing, intracytoplasmic sperm injection, African versus European ancestry, female sex, pre- and postnatal exposure to pollutants and blood cell type compositions, while the associated genetic diseases included Down syndrome and

[†]Francesco Cecere and Raissa Relator equally contributed to this study.

[†]Andrea Riccio and Bekim Sadikovic jointly supervised this work.

*Correspondence:

Bekim Sadikovic
bekim.sadikovic@lhsc.on.ca
Andrea Riccio
andrea.riccio@unicampania.it

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

the developmental disorders with molecular defects in the DNA methyltransferases DNMT1 and DNMT3B, H3K36 methyltransferase SETD2, chromatin remodelers SRCAP and SMARCA4 and transcription factor ADNP.

Conclusions These findings identify several genetic and non-genetic factors including new genes associated with genomic imprinting maintenance in humans, which may have a role in the aetiology of the diseases with imprinting abnormalities and have clear implications in molecular diagnostics.

Keywords Genomic imprinting, Epigenetics, DNA methylation, EWAS, Developmental disorder

Background

Genomic imprinting is a regulatory mechanism resulting in the mono-allelic and parent of origin-dependent expression of a few hundred mammalian genes [1]. In its canonical form, imprinting results from differential establishment and maintenance of DNA methylation on the maternally and paternally derived autosomal chromosomes. More than fifty 1–2 kb-long regions that stably maintain differential DNA methylation (iDMRs) in multiple tissues, including blood, have been identified within the imprinted loci through the analysis of reciprocal genome-wide uniparental disomy (UPD) leukocyte samples [2]. The iDMRs include regions acquiring DNA methylation in the germline (gDMR) that are also defined as imprinting control regions (ICRs) in that they control the expression of all the imprinted genes within a cluster through lncRNAs or other mechanisms. Other iDMRs, also known as somatic or secondary DMRs (sDMRs), are generally established on gene promoters during development, and their methylation depends on that of the closest gDMR.

Genomic imprinting is required for normal development, and molecular defects affecting imprinted gene expression and function occur in rare clinical disorders affecting growth, metabolism, endocrine and neuro-behavioral functions (Imprinting Disorders, ImpDis) [3]. Typically, each ImpDis is caused by molecular abnormalities in a single locus, in many cases consisting of either hypo- or hyper-methylation of the locus-specific gDMR [4]. In recent years, it has become evident that a subset of patients with ImpDis display mosaic multi-locus imprinting disturbance (MLID) that can be detected as loss of DNA methylation at one or more gDMRs in addition to that usually associated with the specific disease [5]. Although the clinical relevance of MLID is currently under investigation, this phenomenon appears to be of biological importance because several MLID cases display atypical clinical features of ImpDis, including high recurrence risk and maternal reproductive problems.

In some ImpDis cases, the methylation changes of the gDMRs are secondary to genetic variants [4]. For example, either hypo- or hyper-methylation of a single gDMR can be caused by variants occurring in *cis*, while MLID can involve variants occurring in *trans*. The latter group includes *trans*-acting recessive variants of the

transcription factor gene *ZFP57* in Transient Neonatal Diabetes and maternal-effect variants of the components of the oocyte Sub-Cortical Maternal Complex (SCMC) mainly in the Beckwith-Wiedemann syndrome (BWS) and Silver-Russell syndrome [4, 6]. Also, environmental factors may affect DNA methylation imprinting and an association between assisted reproduction technology (ART) and ImpDis has been reported by several studies [4]. Nevertheless, the mechanisms underlying most iDMR methylation changes in ImpDis are unknown and their aetiology may be multifactorial [4].

The methylation abnormalities in ImpDis are often incomplete, suggesting that epigenetic mosaicism originated in early development [4]. Diagnosis of ImpDis generally involves quantification of methylation levels at a few representative CpGs within a limited number of iDMRs relative to unaffected control baseline [5]. Thus, very little is known about the methylation levels of the iDMRs in the general population. The recent advancement in the use of genome-wide methylation arrays has allowed to accumulate large datasets of whole-genome DNA methylation profiles in control population and patients affected with rare disorders so that the methylation levels of thousands of CpGs can be studied in vast cohorts of individuals and eventually associated with complex or mendelian traits [7–9]. Despite the availability of all this information, the methylation of the imprinted loci has not been systematically studied beyond the few ImpDis-related loci.

To investigate genome-wide maintenance of imprinted methylation, we analyzed the methylation array datasets obtained on blood DNA and studied the distribution of the iDMR methylation levels and their association with many factors in several thousand individuals, including patients affected by developmental disorders. The results describe the normal profile and variability of the iDMR CpG methylation and identify many genetic and non-genetic factors affecting it in humans.

Methods

Variability of iDMR methylation in human blood DNA

To determine the DNA methylation profiles of the iDMRs in the general population, the normalized Beta-values for each probe covered by the 450 K Illumina methylation array across 2,711 samples were downloaded from

GSE55763. After identifying and removing technical replicates and cross-reacting probes [10], the probes covered across all samples were retained, resulting in 398,094 probes and 2,664 samples for analysis. The iDMR coordinates derived from the information uploaded by David Monk's laboratory (<http://www.humanimprints.net/#data>) and originally published in Court et al. 2014 [2]. The data have been obtained through the analysis of methyl-seq data from whole blood samples and have base-pair resolution. The `bed_intersect` function from the `valr` R package (v0.7.0) was used to identify the 710 probes covering the iDMRs. The distribution of methylation levels across 49 iDMRs was visualized using ridge plots, while variability was assessed using standard deviation (SD) and mean methylation thresholds. Using this criterion, the iDMR CpGs were classified in LOWvar, Mvar, SDvar, and MSDvar CpGs based on their variability thresholds.

Definition of the iDMR central and peripheral segments

To determine the central and peripheral positions of the CpGs, a custom R script was used to divide the iDMRs genomic coordinates into three segments with a 1:2:1 proportion and with the central and peripheral parts covering 2/4 of the length each and including 364 and 346 probes, respectively. By using the information about the variability thresholds, a Fisher's exact test was performed to assess the enrichment of LOWvar, Mvar, SDvar, and MSDvar CpGs in central vs. peripheral regions.

Evaluation of the effect of TFs proximity on CpG methylation

ChIP-seq peaks for the transcription factors were downloaded from publicly available datasets [11] (see Data availability). Using a custom R script (see Data availability), the closest TF peak was assigned to each iDMR CpG. "TF_Close" CpGs were defined as those within a specific distance from the ChIP-seq peak (250 bp for ZFP57, 400 bp for CTCF and ZNF445), while "TF_Far" CpGs were defined as those located beyond these thresholds. To evaluate the relationship between TF proximity and CpG variability, Fisher's exact tests were performed to compare the distribution of MSDvar and LOWvar CpGs between TF_Close and TF_Far groups for each transcription factor. Moreover to investigate about the SNPs, the iDMRs central and peripheral coordinates were used as input for `MethylToSNP` and `snpByOverlaps` function from `SNPlocs.Hsapiens.dbSNP144.GRCh37` R package v(0.99.20).

EWAS analysis

To investigate the relationship between iDMR CpG methylation and complex traits, the trait association data were retrieved from the EWAS Atlas (<https://ngdc.cncb.ac.cn/ewas/atlas/index> [9]). Only the associations

confirmed in blood DNA were selected. By intersecting the probeID of the CpGs covering the iDMRs, the information on the traits associated with the methylation of the imprinted loci were retrieved. To assess the impact of blood cell type composition (BCTC) on iDMR methylation, the DNA methylation profiles of 25 blood cell types were downloaded and the iDMR CpGs with a standard deviation (SD) > 0.05 across blood cell types were identified.

iDMR methylation levels of developmental disorders (DD) patients and controls

The iDMR methylation levels of affected individuals with DDs and unaffected controls were determined by employing our Episign Knowledge Database (EKD, see <https://episign.com/>) DNA methylation data derived from peripheral blood. The EKD consists of cohorts for specific conditions as well as unaffected and healthy controls with varying age and gender, and type of array (Illumina 450k and EPIC). To ensure that the differences between methylation were due to genetic variation, controls were matched to cases by sex, age, and array type, if applicable. For each episignature cohort, the minimum number of controls was set to 150, and depending on the number of cases in a cohort, the case-control ratio ranged from 1:1 to 1:30. Beta values for probes in 50 iDMRs were investigated for both groups. Probes with missing values for at least 40% of samples in a group were removed. Missing values for the remaining probes were then imputed using the median value of that probe in the group. Ridge plots were generated to show density histogram of methylation beta values for each group in each signature cohort for every region. Furthermore, statistics was calculated to quantify methylation differences between each case cohort and its control cohort using two-sided z-test, with p-values adjusted to control for false discovery rate using the Bonferroni-Hochberg method, and population mean and standard deviation were computed using the data from matched controls. To identify differentially methylated probes (DMPs), the methylation levels of the case cohorts were compared to a control group using a z-score approach based on pooled variance. For each probe, the pooled variance was computed, which considers both the disorder and control variances. The pooled standard deviation (SD) was obtained by taking the square root. Z-scores were then calculated to quantify how much the disorder's mean methylation differs from the control group, normalized by the pooled SD. To classify probes as differentially methylated, ± 2 standard deviations (SDs) were set as threshold.

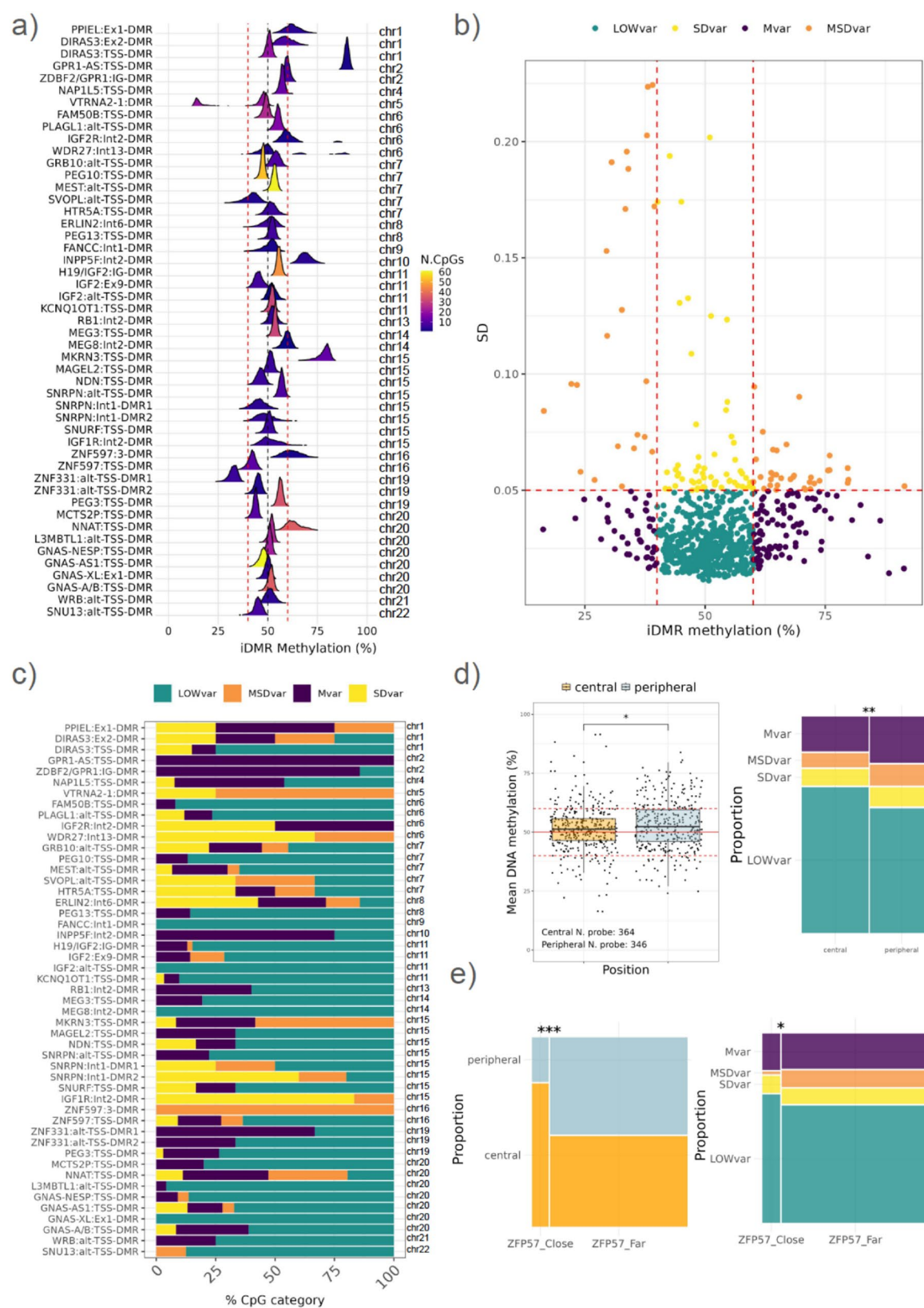


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Methylation profiles of the iDMRs in the general population. **(a)** Ridge plot showing the DNA methylation profile (Bvalues) of 2,664 analyzed control samples across 49 iDMRs. The x-axis represents the DNA methylation value, while the y-axis lists the iDMRs. The density color indicates the number of probes covering each region. **(b)** Scatterplot representing the relationship between the average DNA methylation value (x-axis) and the standard deviation (y-axis) for each iDMR CpG. Dots are color-coded based on variability thresholds defined by the red dashed lines: LOWvar CpGs with $sd < 0.05$ and $40\% < x < 60\%$ methylation (teal), Mvar CpGs with methylation outside the $[40\%, 60\%]$ range (dark purple), SDvar CpGs with $sd > 0.05$ (yellow) and MSDvar CpGs with methylation outside the $[40\%, 60\%]$ range and $sd > 0.05$ (orange). **(c)** Stacked barplot showing the percentage distribution of LOWvar, Mvar, SDvar and MSDvar CpGs in each iDMR. **(d)** iDMR subregion analysis: *Left*: Boxplot comparing the methylation values of the CpGs located in the peripheral versus central segments within the iDMRs. The significance is calculated using the Wilcoxon rank sum test. *Right*: Mosaic plot showing the proportion of SDvar, Mvar, MSDvar and LOWvar categories among the CpGs of the central or peripheral iDMR segments. The significance of the higher LOWvar-central probes is calculated using the Fisher-exact test. **(e)** ZFP57 Peak Proximity: *Left*: Mosaic plot showing the proportion of the CpGs that are ≤ 250 bp (Close) or > 250 bp (Far) from the ZFP57 peak summits in the peripheral and central iDMR subregions. The significance of the higher proportion of the CpGs “Close” to ZFP57 peaks in the central subregions is calculated using the Fisher-exact test. *Right*: Mosaic plot representing the proportion of SDvar, Mvar, MSDvar and LOWvar categories among the ZFP57 “Close” and “Far” CpGs. The significance of the higher proportion of the MSDvar among the CpGs “Far” from the ZFP57 peaks is calculated using the Fisher-exact test

Results

Methylation profiles of the iDMRs in the general population

We determined the methylation level of 49 human iDMRs (Supplementary Table 1) in the peripheral blood leukocyte DNA of a normal population by analyzing the methylation array datasets of 2664 individuals [12, 13]. Consistent with the differential methylation of their maternal and paternal alleles, most iDMRs displayed a mean methylation level close to 50% (Fig. 1a, Supplementary Table 2). Also, the distribution of their methylation levels generally resembled symmetric unimodal curves with a median value of 50% and was not strongly affected by the number of CpGs per iDMR assayed. However, certain loci deviated from this norm. For example, MKRN3:TSS, INPP5F: Int2 and GPR1-AS: TSS had median methylation levels well above 50%, and ZNF331:alt-TSS-DMR1 median methylation levels well below 50%. In addition, VTRNA2-1, WDR27:Int13 and IGF2R: Int2 showed multiple peaks, consistent with the polymorphic methylation described for one of these loci [14, 15]. Other iDMRs, such as NNAT: TSS, DIRAS3:Ex2 and IGF1R: Int2, are more dispersed around the median indicating the presence of higher variability in the population.

Our analysis covered 710 CpGs with 1–61 CpGs for each iDMR in the cohort under study (Fig. 1a, Supplementary Table 3). To determine how the individual CpGs influenced the distribution of iDMR methylation, we calculated the standard deviation (sd) and the mean methylation level of each iDMR CpG under the assumption of normal distribution of their methylation level. We identified 117 CpGs with $sd > 0.05$ (SDvar CpGs) indicating high variability within the control population and 194 CpGs with mean methylation outside the $[40\%, 60\%]$ range (Mvar CpGs) indicating atypical profiles. Of these, 58 CpGs met the criteria of both the SDvar and Mvar CpGs (MSDvar CpGs) (Fig. 1b, Supplementary Table 4). We observed that VTRNA2-1, MKRN3:TSS, IGF1R:Int2, PPIEL: Ex1, WDR27:Int13, IGF2R:Int2 and ZNF597:3' were enriched for both SDvar and

Mvar CpGs, IGF1R:Int2 and SNRPN:Int1 were particularly enriched in SDvar CpGs, while ZDBF2/GPR1:IG, ZNF331:alt-TSS, INP55F:Int2 and GPR1-AS: TSS were particularly enriched in Mvar CpGs, (Fig. 1c and Supplementary Table 5). Apart from these, 457 CpGs (LOWvar CpGs) displayed $sd < 0.05$ and methylation level between 40% and 60%, meeting the characteristics expected from imprinted loci. The LOWvar CpGs represented $> 50\%$ of the CpGs in 29 iDMRs and overall were present in 41/49 loci (Fig. 1c). Indeed, when we reanalyzed the distribution of the iDMR mean methylation in our population after removing the SDvar CpGs, only the VTRNA2-1, WDR27:Int13, IGF1R:Int2 and ZNF597:3' DMRs were completely lost, while the other iDMRs showed a more homogeneous profile (Supplementary Fig. 1a). On the other hand, when the Mvar CpGs were omitted, MEG8:Int2, GPR1-AS:TSS, INPP5F:Int2 and ZDBF2/GPR1 that strongly deviated from 50% methylation were lost, but the other iDMRs showed a mean methylation closer to 50% (Supplementary Fig. 1b). VTRNA2-1, WDR27:Int13 and IGF2R:Int2 represent an exception, in which the methylation level of all their CpGs are polymorphic. Finally, we obtained a methylation profile with a tiny peak close to 50% (Supplementary Fig. 1c) for 41/49 iDMRs, when only the LOWvar CpGs were considered.

To understand how CpG methylation was distributed within the iDMRs, we divided these regions into three segments: two peripheral and a central one. We observed that the methylation level of the central segments was closer to 50% with respect to the peripheral ones. Also, the peripheral segments had a higher number of Mvar and SDvar CpGs than the central ones, which instead were enriched in LOWvar CpGs (Fig. 1d, Supplementary Table 6), suggesting that the core of the iDMRs is more stable than their boundaries. This was particularly evident for some iDMRs (Supplementary Fig. 2a). To study the dependence of this trend on transcription factor binding, we intersected the coordinates of the central and peripheral iDMR segments with the binding sites (ChIPseq peaks) of ZFP57, which is a key protein in iDMR methylation maintenance [16, 17]. We found that

the CpGs that are closer (≤ 250 bp) to the summit of the ZFP57 peaks were more enriched in the central segments compared to the peripheral ones, while the CpGs that are more distant (> 250 bp) from the ZFP57 peaks were similarly distributed between peripheral and central segments (Fig. 1e, Supplementary Fig. 2b, Supplementary Table 7). Moreover, the CpGs that are more distant from the ZFP57 peaks were more represented by MSDvar than the CpGs that are closer to these binding sites (Fig. 1e, Supplementary Table 7). Similar to ZFP57, we found that the CpGs that are closer (≤ 400 bp) to the ChIPseq peaks of ZNF445, another transcription factor implicated in imprinting maintenance [18], were more enriched in the central region of the iDMRs. However, in this case, we did not find any difference in the methylation variability of the CpGs that are closer or further to ZNF445 binding sites (Supplementary Fig. 2c). Finally, we investigated the role of CTCF, a transcription factor that have implications in chromatin architecture and transcriptional regulation of genes flanking its binding sites [19]. Contrary to ZFP57 and ZNF445, we found that the CpGs that are closer (≤ 400 bp) to the CTCF ChIPseq peaks were more enriched in the peripheral than in the central iDMR regions (Supplementary Fig. 2d) and these CpGs were more enriched in LOWvar CpGs with respect to the CpGs that are further from the CTCF binding sites, suggesting that CTCF binding is important to preserve the function of the iDMRs and their methylation stability, particularly at their boundaries.

To investigate if the higher methylation variability of the iDMR periphery was associated with single nucleotide polymorphisms (SNPs) possibly acting as DNA methylation quantitative trait loci (mQTL), we quantified the number of known SNPs (from dbSNP) within central and peripheral iDMR regions. We found comparable numbers of SNPs in the two groups (central: 3,702 SNPs; peripheral: 3,472 SNPs; Supplementary Fig. 2e), suggesting that SNP density and mQTL effects are not the primary causes of the regional differences in methylation variability of the iDMRs.

In summary, two-thirds of the iDMR CpGs displayed a methylation level close to 50% with relatively low variability in blood DNA of a control population. However, a few iDMRs displayed highly variable CpG methylation and consistent deviation from 50% methylation level, including three polymorphic loci. Although variably methylated CpGs were present in the whole length of the iDMRs, these were less frequent in the central segments and in the vicinity of the ZFP57 and CTCF binding sites.

Association of blood iDMR methylation with complex traits

Once assessed the variation of the iDMRs methylation level, we asked what factor affected the methylation of these regions within the general population. For this

purpose, we retrieved the information deposited in the EWAS Data Hub and Atlas for the CpGs of the iDMRs. These data included high-quality associations of about 700 complex traits with CpG methylation in 3000 cohorts and 200 tissues/cells [9]. For further analysis, we considered only the traits significantly associated with variation of iDMR methylation level in blood DNA (Supplementary Table 8). Also, we distinguished the traits into two categories: (i) non-disease traits including environmental factors and phenotypic-behavioral traits; (ii) disease traits, including neoplastic and non-neoplastic diseases. We found 107 iDMR CpGs associated with 47 non-disease traits and 111 iDMR CpGs associated with 25 disease traits (Fig. 2a; Table 1 and Supplementary Table 9). Overall, NNATTSS, VTRNA2-1, GNAS-AS1:TSS, MEST:alt-TSS and H19/IGF2:IG exhibited the highest number of CpGs associated with these traits (Fig. 2b).

DIRAS3:Ex2, ERLIN2:Int6 and PEG13:TSS were the iDMRs more strongly associated with the non-disease traits (Fig. 2c). In particular, the methylation of 8/8 DIRAS3:Ex2 CpGs were positively correlated (increased on average) with ageing and 7/8 of them also negatively correlated (decreased on average) with African ancestry, meaning that methylation of this locus was higher in Africans than in Europeans, East Asians and Aboriginal Australians (Fig. 2d and Supplementary Fig. 3). Ageing also affected fewer CpGs in several other iDMRs. In addition, 4/7 ERLIN2:Int6 CpGs were positively correlated with female sex (Fig. 2d). Furthermore, 30 CpGs distributed in 18 iDMRs (particularly FAM50B:TSS, GNAS-A/B:TSS-DMR and MEST:alt-TSS) were associated with air pollution (Supplementary Fig. 3a). However, this effect was complex, depending on the nature of the pollutant examined (Supplementary Table 9). Furthermore, intracytoplasmic sperm injection (ICSI), a method of assisted reproduction technology, was associated with methylation changes of several iDMRs, particularly with 3 NAP1L5:TSS CpGs (Fig. 2d). Finally, maternal exposure to several potentially dangerous substances were associated with iDMR methylation. Particularly relevant is the negative correlation of maternal smoking with 3 MKRN3:TSS CpGs and the negative correlation of several GNAS-AS1:TSS CpGs with prenatal exposure to Per- and Polyfluorinated Alkylated Substances (PFAS) (Supplementary Fig. 3a, Supplementary Table 9).

Apart from MEG8:Int2 and IGF2:alt-TSS with single CpGs covered, WDR27:Int13 (3/3), VTRNA2-1 (17/20), INPP5F:Int2 (3/4) and NNAT:TSS (22/36) were the iDMRs with the highest percentage of CpGs associated with disease-traits (Fig. 2e, Supplementary Table 11). In particular, the methylation of many iDMRs but VTRNA2-1, which was positively correlated, were negatively correlated with Down syndrome (DS, Fig. 2f). Several iDMRs, including VTRNA2-1, were also correlated

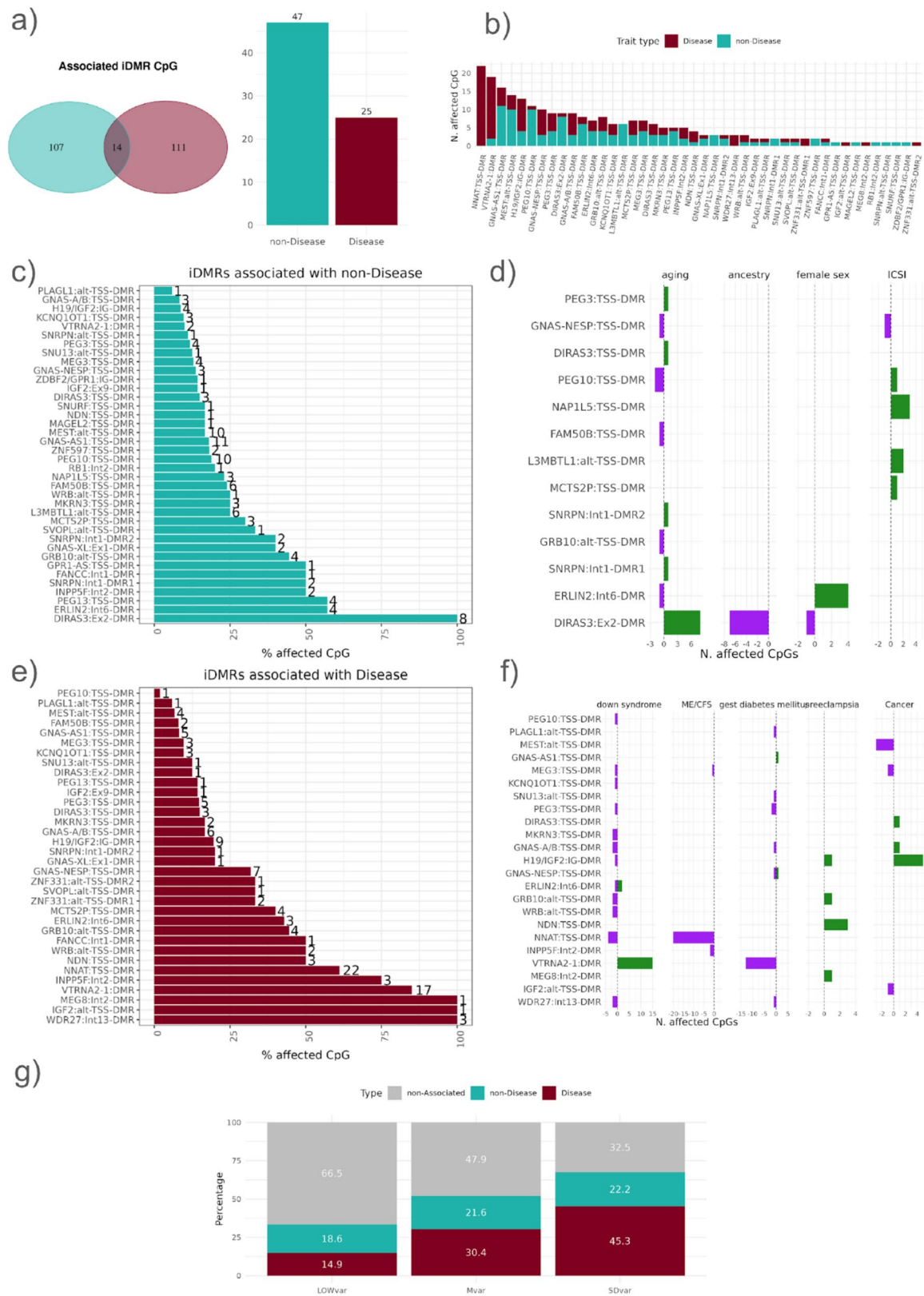


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Association of the complex traits listed in the Supplementary Table S8 with iDMR methylation. **(a)** Associated traits with iDMR CpGs: *Left*: Venn diagram displaying the CpGs associated with complex non-disease and disease traits. *Right*: Barplot showing the number of unique non-disease (cyan) and disease (red) traits associated with iDMR methylation. **(b)** Barplot showing the number of affected CpGs for each iDMR associated with non-disease (cyan) or disease (red) traits. **(c)** Barplot representing the percentage of CpGs of each iDMR associated with non-disease traits. **(d)** Barplot displaying the number of CpGs associated with specific non-disease traits. Negative correlations are in purple, positive correlations in forestgreen. **(e)** Barplot representing the percentage of CpGs of each iDMR associated with disease traits. **(f)** Barplot displaying the number of CpGs associated with specific disease traits displayed as in (d). **(g)** Stacked Barplot showing the percentage of disease (red) and non-disease (cyan) trait associations with LOWvar, Mvar and SDvar CpGs

mostly negatively with gestational diabetes mellitus (Fig. 2f). Furthermore, we observed a negative correlation between myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) and 20 NNAT: TSS CpGs (Fig. 2f), as well as a positive correlation between preeclampsia and 3 NDN:TSS CpGs (Fig. 2f). Also, methylation of several iDMRs, including MCTS2P:TSS, GNAS-NESP:TSS, GNAS-AS1:TSS and PEG3:TSS was correlated mostly positively with the inflammatory disorder Behcet's disease. Blood DNA methylation of several DMRs were associated with cancer (Fig. 2f). In particular, we found 5 H19/IGF2:IG CpGs positively correlated with lung cancer, 3 MEST:alt-TSS CpGs negatively correlated with breast cancer, as well as the positive correlation of DIRAS3:TSS and GNAS-AB:TSS CpGs and the negative correlation of IGF2:alt-TSS and MEG3:TSS CpGs with further various cancer types (Supplementary Fig. 3b).

The availability of vast data on iDMR methylation variability, gave us the possibility to explore the presence of combined effects on multiple loci. For this purpose, we performed a correlation study among all the covered iDMRs CpGs. Selecting a correlation value >0.60 , we found several iDMRs with concordant methylation variability, such as MEST:alt-TSS and GNAS-A/B:TSS, MEST:alt-TSS and MEG3:TSS, KCNQ1OT1:TSS and GNAS-A/B:TSS, and MEG3:TSS and GNAS-A/B:TSS (Supplementary Fig. 5 and Supplementary Table 11).

In order to investigate the relationship between the effect of the complex traits and the variability of iDMR methylation in the normal population, we intersected the information related to our control cohort and the association data (Supplementary Table 10). We found that more than 67% of the SDvar and 52% of the Mvar CpGs but only 33.5% of the LOWvar CpGs were associated with a trait (Fig. 2g). The prevalence of the Mvar and SDvar among the affected CpGs was more substantial for the disease-associated traits, but was evident for both trait categories. In summary, we found that iDMR methylation changes were associated with several environmental and complex traits and more concentrated on the SDvar and Mvar CpGs. These associations may explain part of the methylation variability observed in the general population.

Blood cell type-specific methylation of the iDMRs

Blood is a tissue composed of many different cell types. A potential factor contributing to the variability of iDMR

methylation levels in the human population could be ascribed to differences in blood cell-type composition (BCTC). To address this issue, we looked at the methylation level of the iDMR CpGs in the datasets of different blood cell types in the EWAS datahub (Supplementary Table 11). We found 63 iDMR CpGs whose methylation level had $sd > 0.05$ across all the blood cell types (Fig. 3a). The BCTC effect was more substantial on NNAT:TSS and DIRAS3:Ex2 with 22/36 and 7/20 affected CpGs, respectively (Fig. 3b). In particular, the methylation level of NNAT:TSS was higher in T lymphocytes and primarily in CD127 regulatory and the CD45 memory T cells with respect to the other blood cell types, while that of DIRAS3:Ex2 was higher in monocytes and neutrophils (Fig. 3c). 2/4 of the WRB:alt-TSS CpGs and 2/3 of the SVOPL:alt-TSS CpGs were also affected by BCTC. In the other iDMRs, the number of affected CpGs did not exceed 20%. Overall, 20/51 SDvar and 30/118 Mvar but only 25/432 LOWvar CpGs were affected by BCTC, with the highest number of SDvar and Mvar CpGs in NNAT:TSS and DIRAS3:Ex2 (Fig. 3d). To investigate the effect of BCTC on their methylation profiles, we removed the affected CpGs from the NNAT:TSS and DIRAS3:Ex2 DMRs and plotted their methylation distribution in the control population. We observed that by using this filter the mean of the NNAT:TSS and DIRAS3:Ex2 distributions was closer to 50% (Fig. 3e). In summary, although BCTC affected the methylation level of a few CpGs in most iDMRs, NNAT:TSS and DIRAS3:Ex2 included a higher number of CpGs with different methylation in blood cell types and this may contribute to the high variability and deviation from the 50% level observed for these loci in the general population.

iDMR methylation in developmental disorders (DDs)

Through the interrogation of a vast number of DNA methylation array datasets, methylation changes have been recently identified in the blood DNA of individuals affected by mendelian and chromosomal DDs, and diagnostic epigenatures have been identified in many cohorts [7, 20]. To identify genes that possibly affect genomic imprinting in humans, a screen for modifiers of iDMR methylation was conducted in samples of patients affected by 56 DDs that display mutations in the proteins of the epigenetic machinery and exhibit genome-wide DNA methylation epigenatures within our EpiSign Knowledge Database (EKD) (Supplementary Table 12, see

Table 1 iDMRs most significantly associated with complex traits

iDMR	Trait	Frac- tion of affected CpGs
DIRAS3:Ex2-DMR	aging	8/8
NAP1L5:TSS-DMR	intracytoplasmic sperm injection	3/3
NNAT:TSS-DMR	myalgic encephalomyelitis/chronic fatigue syndrome	20/22
DIRAS3:Ex2-DMR	African versus European ancestry	7/8
VTRNA2-1:DMR	Down syndrome	15/18
NDN: TSS-DMR	preeclampsia	3/4
VTRNA2-1:DMR	gestational diabetes mellitus	13/18
ERLIN2:Int6-DMR	female sex	4/6
MKRN3:TSS-DMR	prenatal exposure to smoking	3/5
FAM50B:TSS-DMR	air pollution	4/7
GNAS-AS1:TSS-DMR	prenatal PFAS exposure	8/16
ERLIN2:Int6-DMR	Down syndrome	3/6
MCTS2P:TSS-DMR	Behcet's disease	3/6
GNAS-NESP:TSS-DMR	Behcet's disease	4/9
H19/IGF2:IG-DMR	cancer	5/13
GNAS-A/B:TSS-DMR	air pollution	3/8
PEG3:TSS-DMR	Behcet's disease	3/9
MEST:alt-TSS-DMR	air pollution	4/14
MEST:alt-TSS-DMR	cancer	3/14
GNAS-AS1:TSS-DMR	Behcet's disease	3/16
NNAT:TSS-DMR	Down syndrome	4/22

also <https://episign.com/>). Control samples ($N=250$) of healthy individuals and individuals negative for an existing EpiSign disorder were matched to DD cases using age, sex, and array type. It was observed that the matched controls exhibit the same trends seen in the general population, with most peaks near 50% methylation (Supplementary Fig. 4, Supplementary Table 13). Comparison of methylation status between EpiSign cohort cases and controls showed statistically significant (i.e., at least 10% mean difference and adjusted p -value <0.05) difference in mean methylation levels for at least one iDMR in 15 cohorts (Table 1; Fig. 4a, and Supplementary Table 14).

In particular, between two and nine iDMRs were significantly deregulated in Immunodeficiency, centromeric instability, facial anomalies syndrome 1 (ICF1), Immunodeficiency, centromeric instability, facial anomalies syndrome 2,3,4 (ICF2,3,4), Luscan–Lumish syndrome (LLS), DS and Floating-Harbour syndrome (FLHS) cohorts. Methylation of one iDMR was significantly altered in the remaining ten disorders, and for many of these most of the other iDMRs followed a similar trend (Table 2, Supplementary Tables 14 and Supplementary Figs. 5–S60). In addition, the methylation change was homogeneous along the iDMRs (Supplementary Fig. 7). Both gDMRs and sDMRs were affected, but some loci (e.g., NNAT:TSS, IGF1R:Int2, PPIEL:Ex1 and the SNRPN sDMRs) were more frequently deregulated than others (Supplementary

Table 14). The mean iDMR methylation differences between DD patients and controls ranged from 10 to 20% (Fig. 4a and Supplementary Table 14), with an average of 13%, and the profiles of the patients and controls were partially overlapping (Fig. 4b). Although statistical significance was calculated for the entire cohort, we estimate that the number of patients with methylation level different from controls exceeds 50% in all the disorders and iDMRs (Supplementary Fig. 5). Most of the cohort cases presented hypomethylated patterns compared to matched controls, except for the Cerebellar ataxia, deafness, and narcolepsy, autosomal dominant syndrome (ADCADN), Coffin–Siris syndrome 4 (CSS4_c.2650), Hunter–McAlpine craniosynostosis syndrome (HMA) and Williams–Beuren deletion syndrome (Williams) cohorts, which were hypermethylated in their respective deregulated DMRs (Fig. 4b).

We then looked at the methylation of the individual CpGs of the defective iDMRs in the DD cohorts and found 76 CpGs significantly deregulated in at least one disorder compared to the control cohort (Supplementary Table 15). According to their variability in the general population, we found 18 LOWvar, 14 SDvar, 24 Mvar and 20 MSDvar CpGs deregulated in the DD patients (Fig. 4c). In summary, these results strongly suggest that the epigenetic regulators that are mutated in the DDs control the maintenance of iDMR methylation in blood DNA, with a stronger effect on the CpGs with moderate to high variability in the general population.

Unfortunately, it is difficult to study the consequence of iDMR methylation changes on imprinting expression in blood cells, because most imprinted genes are poorly expressed in these cell types [21]. Nevertheless, by retrieving an available microarray dataset [22], we found that the expression of several imprinted genes was altered in the brain of DS patients (Supplementary Fig. 8).

Discussion

Defective iDMR methylation caused by unknown molecular mechanisms has been reported in ImpDis and cancer [4, 23]. However, the methodologies used to investigate the iDMRs in these studies were diverse and generally targeted a few CpGs. Although some evidence suggested polymorphic imprinting and interaction with environmental factors for some loci, the small sample size generally used in these studies limited information on iDMR methylation in the general population. On the other hand, high-throughput studies based on whole-genome methylation arrays have accumulated in the last few years and used to determine the association of environmental factors and complex and mendelian traits with DNA methylation changes. We employed this vast dataset to establish the iDMR methylation profiles in the general population and identified genetic and non-genetic

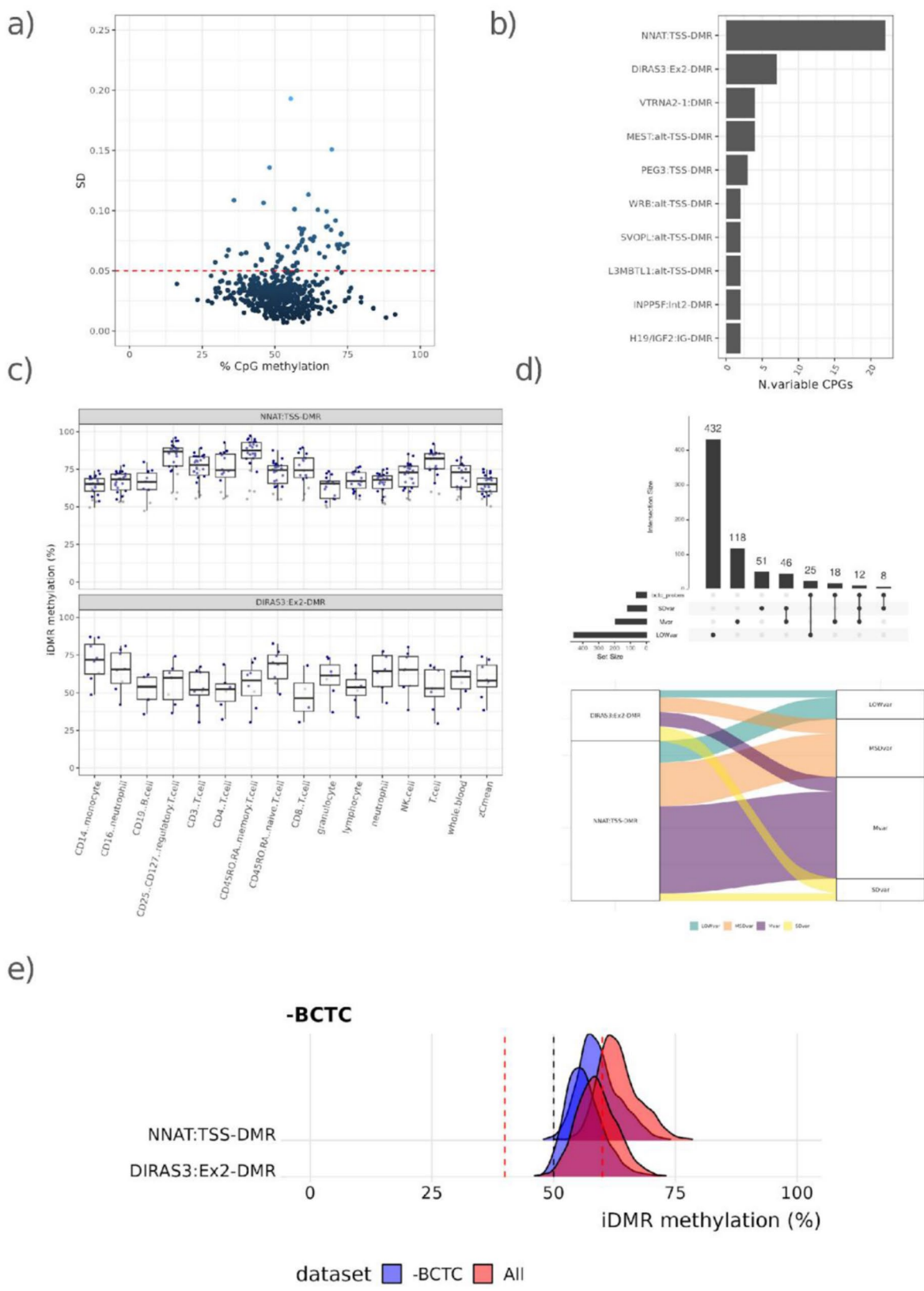


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Effect of blood cell-type composition (BCTC) on iDMR methylation. **(a)** Scatterplot displaying the average values and standard deviation of iDMR methylation across different blood cell types. The color scale of the dots indicates the SD value. **(b)** Barplot displaying the number of iDMR CpGs affected by BCTC. **(c)** Boxplot displaying the DNA methylation levels (β values) of the NNAT:TSS and DIRAS3:Ex2 DMRs in different blood cell types. The color scale of the dots (from dark grey to blue) indicates the strength of the BCTC effect (SD). **(d)** Distribution of the BCTC-affected CpGs among the SDvar, Mvar, MSDvar and LOWvar categories. *Top*: Upset plot showing the intersection between SDvar, Mvar, MSDvar and LOWvar and the BCTC-affected CpGs. *Bottom*: Alluvial plot showing the proportions of SDvar, Mvar, MSDvar and LOWvar within the NNAT:TSS and DIRAS3:Ex2 iDMRs. **(e)** Ridge plot showing the DNA methylation profiles (β values) of the DIRAS3:Ex2 and NNAT:TSS DMRs in the 2,664-individuals control cohort before and after the BCTC probes filtering

factors that contribute to the variability of these patterns in human blood DNA. We demonstrated that the methylation of about one-third of the iDMR CpGs deviates from the expected 50% level, is polymorphic or highly variable and identified several environmental and genetic factors contributing to this variability in the general population or in cohorts of individuals affected by developmental disorders.

Due to their gamete-of-origin-specific methylation, the iDMRs are expected to show a 50% methylation level in somatic cells. Although most of these regions displayed profiles compatible with these expectations, this was not the case for many loci when examined in a large population. In particular, the gDMRs ZNF331:alt-TSS1, INPP5F:Int2 and GPR1-AS:TSS and the sDMR MKRN3:TSS showed a mean methylation level different from 50%, indicating either *de novo* or loss of methylation in blood cells during development. Other iDMRs, such as VTRNA2-1, WDR27:Int13 and IGF2R:Int2 DMRs, displayed polymorphic methylation in our control cohort. These results are consistent with the reported polymorphic methylation imprinting of VTRNA2-1 and polymorphic expression imprinting of IGF2R:Int2 [14, 15, 24, 25]. Other gDMRs, including NNAT:TSS, DIRAS3:Ex2, ERLIN2:Int6, PPIEL:Ex1, SVOPL:alt-TSS and ZNF597:3', showed very high variability within our population. Also, methylation variability was not uniform within the iDMRs. The higher stability of the CpGs located in the vicinity of the ZFP57 binding sites that are more abundant in the center of the iDMRs and the CpGs near the CTCF binding sites that are more frequent in the iDMR periphery indicate that transcription factor binding contributes to the maintenance of methylation imprinting in the general population. On the other hand, the higher variability of the iDMR borders is reminiscent of the methylation changes of the CpG island shores occurring in cancer and tissue differentiation, possibly because of paucity of DNA methylation-blocking histone modifications [26, 27].

Because blood cell types have diverse methylation level of many CpGs, changes in BCTC may influence the methylation profiles determined in blood DNA of an individual [28]. However, the effect of BCTC on iDMR methylation has not been investigated so far. We demonstrated that BCTC has little effect on most imprinted loci. However, NNAT:TSS and DIRAS3:Ex2 are exceptions, in that the methylation level of the majority of their

tested NNAT:TSS CpGs and several DIRAS3:Ex2 CpGs differed among blood cell types. Although their CpGs were poorly covered, the WRB:alt-TSS and SVOPL:alt-TSS DMRs were also affected by BCTC. This may contribute to their high sd and deviation from 50% methylation level of the NNAT:TSS and DIRAS3:Ex2 DMRs observed in the general population and in some specific cohorts. For example, it is possible that the NNAT:TSS methylation changes we observed in DS are partly due to the frequent BCTC alterations these patients have [29]. Similarly, BCTC changes may underly at least in part the NNAT:TSS CpG hypomethylation detected in individuals affected by ME/CFS, as well as that of FLHS, LLS, RMNS and TBRS. Thus, it would be advisable to exclude these BCTC-affected CpGs when assaying methylation imprinting in diagnostic tests.

Genomic DNA methylation is deeply influenced by genetic and environmental factors [9, 30]. iDMR methylation establishment and maintenance are primarily controlled by the maternal-effect genes encoding the member of the SCMC complex and the transcription factors ZFP57 and ZFP445 during gametogenesis and preimplantation embryo development [4, 5]. However, if and how further genetic and environmental factors interacting with the DNA methylation machinery affect imprinting methylation is poorly defined. The availability of many whole-genome array datasets in which association with environmental conditions and complex traits was tested, prompted us to investigate to what extent iDMR methylation was affected by these relatively common factors. We found that many disease and non-disease-associated traits affect iDMR CpG methylation. Consistent with the dynamics of DNA methylation during human lifetime [31], ageing was the non-disease factor associated with more iDMRs. Although only one or two CpGs were significantly affected in most loci, this result demonstrates the importance of age-matched controls when measuring iDMR methylation in humans. Differently from the other iDMRs, many CpGs of DIRAS3:Ex2 were associated with ageing. These same CpGs were influenced by European versus African ancestry and BCTC. While the methylation differences between populations may be attributed to DNA sequence variants in the DIRAS3 locus [32], the association with both ageing and BCTC suggests that methylation of this iDMR may be influenced by natural changes in the relative proportion of blood cell types in elder individuals [33]. Another interesting association is

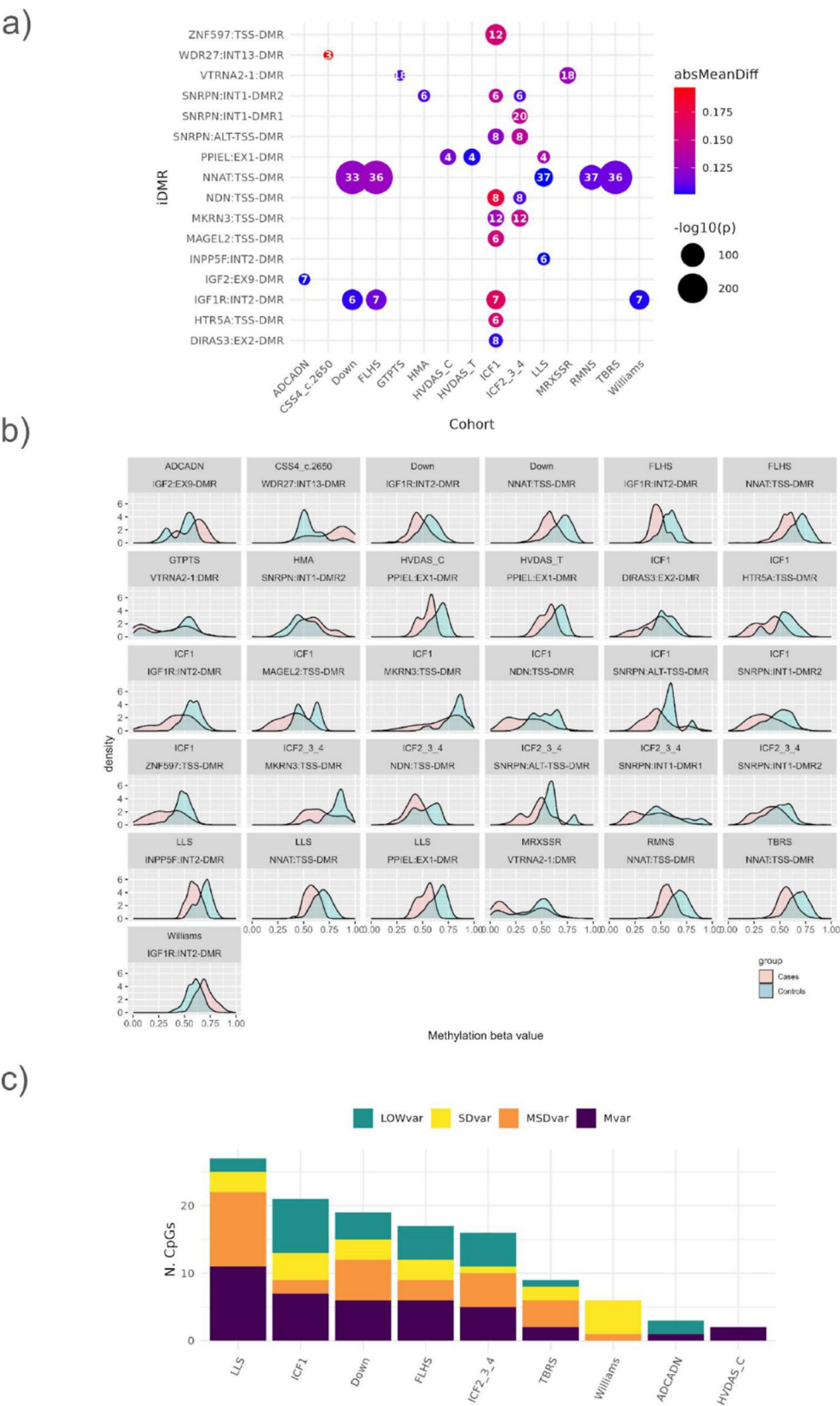


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 iDMR methylation changes in DDs. **(a)** Bubble plot showing the differentially methylated iDMRs in the DD cohorts. The disorders are indicated on the x-axis and the iDMRs on the y-axis. The size of each bubble indicates the statistical significance ($-\log_{10}$ p-value), the color represents the absolute mean methylation difference (absMeanDiff) and the figure within the bubble the number of affected CpGs. **(b)** Density plots comparing the methylation distributions of differentially methylated iDMRs between control individuals (blue) and DD patients (pink). **(c)** Distribution of the CpGs with different methylation level in DDs vs. controls among the SDvar, Mvar, MSDvar and LOWvar categories

between *ERLIN2:Int6* and the female sex. This DMR that is located within the imprinted Endoplasmic Reticulum Lipid Raft-Associated Protein 2 gene that is responsible for the *Spastic Paraplegia 18 A* [34] is more methylated in girls compared to boys.

Several studies have reported an increased risk of developing ImpDis and particularly BWS in children conceived through Assisted Reproduction Technology, suggesting a disturbance in imprinting maintenance caused by the procedures used to enhance reproductive success [35]. Consistent with these observations, we report an effect of intra-cytoplasmic sperm injection on several iDMRs. Also, air pollution had an effect on the methylation of several iDMRs. More specific associations were evident with prenatal exposure to smoking and the pollutants PFHxS and PFOA. In particular, maternal smoking was strongly associated with hypomethylation of *MKRN3:TSS*, a locus involved in Central Precocious Puberty [36], while PFHxS and PFOA were associated with hypomethylation of *GNAS-AS1:TSS*, a molecular defect of Pseudohypoparathyroidism [37].

Trisomy 21 (DS) has a widespread impact on DNA methylation, possibly caused by an imbalance of epigenetic regulator genes located in Chromosome 21 [38]. However, its consequence on genomic imprinting has not been investigated so far. We demonstrated hypomethylation of many iDMR CpGs in the blood DNA of DS patients. An elevated number of deregulated DMRs was also found in children exposed to gestational diabetes mellitus who are known to be at high risk of metabolic diseases [39]. The affected *PLAGL1*, *GNAS* and *PEG3* imprinted genes may have a role in the aetiology of these metabolic disturbances. Also, the expression of several imprinted genes was altered in the post-mortem brain of DS patients. The observed association of DS and gestational diabetes with *VTRNA2-1* methylation should be taken with caution, because of the polymorphic imprinting of this locus. Indeed, *VTRNA2-1* results can be explained by genetic association or bias in the frequency of allelic variants between patients and controls. Disturbed Placental Imprinting has been reported in preeclampsia [40]. We found methylation changes of several iDMRs in blood leukocytes of women with early-onset preeclampsia, which may be explained by epigenetic modifications predisposing to this disease originated in early development. Finally, we demonstrated the association of iDMR methylation with several types of cancer. Although deregulated imprinting is a frequent

finding in cancer cells, its alteration in blood leukocyte DNA has not been systematically investigated [41, 42]. The iDMR methylation changes we detected, particularly those of *H19-IGF2:IG*, *IGF2:alt-TSS*, *MEST:alt-TSS* and *DIRAS3:TSS*, may be originated in early development and represent markers for early tumour diagnosis.

Screening for DNA methylation modifiers by employing the array datasets of patients with developmental disorders harbouring mutations in proteins of the epigenetic machinery was successfully used in previous studies [43]. Using a similar approach, we identified several genes associated with iDMR methylation changes. First, we confirmed and expanded the results previously obtained in *ICF1* and *ICF2-4* [44], which indicate the important role of *DNMT3B* and associated factors in imprinting maintenance. Among the new genes associated with iDMR methylation changes, we identified the one coding for the H3K36 methyltransferase *SETD2* that in the mouse is required to establish imprinted methylation in oocytes [45]. Further genes affecting iDMR methylation were those coding for the DNA methyltransferases *DNMT1*, the chromatin remodelers *SRAP* and *SMARCA4* and the transcription factor *ADNP*. Finally, the association of iDMR methylation changes with the HMA syndrome and the Williams syndrome that are characterized by the deletion of chromosome 17q23.1-q24.2 and 7q11.23, respectively, indicates that further epigenetic regulator genes controlling methylation imprinting maintenance are located in these regions. Concerning the associations between DDs and *NNAT:TSS* and *VTRNA2-1* methylation, the considerations reported above for these two loci should be taken into accounts. Interestingly, *ICF1* and *ICF2-4* patients had many hypomethylated iDMRs that do not show any BCTC-effect in the general population. Furthermore, *DIRAS3:Ex2* but not *NNAT:TSS* were affected and *DIRAS3:Ex2* hypomethylation was not restricted to a few CpGs but homogeneously spread in all the iDMR (Supplementary Fig. 5). Thus, the methylation changes observed in ICF patients are likely not caused by their memory B-cell deficiency, but rather secondary to defective DNA methylation machinery. Findings of disrupted DNA methylation profiles in iDMRs are consistent with and part of the broader, genome-wide disruption of DNA methylation profiles in rare neurodevelopmental disorders. Stability and reproducibility of these DNA methylation profiles across individuals affected by these disorders allow for development of highly accurate episignatures

Table 2 Developmental disorders with affected iDMRs

Disorder	Gene(s) involved	Category	Patients number	Affected iDMRs
Cerebellar ataxia, deafness, and narcolepsy, autosomal dominant (ADCADN)	<i>DNMT1</i>	DNA methyltransferase	5	IGF2:Ex9
Coffin–Siris syndrome 4 (CSS4_c.2650)	<i>SMARCA4</i>	Chromatin remodeler	3	WDR27:Int13
Down Syndrome (DS)	<i>HMGN1</i>	Chromatin remodeler	40	IGF1R:Int2 NNAT:TSS
	<i>MIS18A</i>	Chromatin remodeler		
	<i>AIRE</i>	Transcription factor		
	<i>USP16</i>	Histone ubiquitination		
	<i>PRMT2</i>	Histone methyltransferase		
	<i>DNMT3L</i>	DNA methyltransferase		
	<i>RBM11</i>	Alternative splicing		
	<i>CHAF1B</i>	Chromatin remodeler		
	<i>BRWD1</i>	Histone reader		
Floating-Harbour syndrome (FLHS)	<i>SRCAP</i>	Chromatin remodeler	21	IGF1R:Int2 NNAT:TSS
Genitopatellar syndrome (GTPTS)	<i>KAT6B</i>	Histone acetyltransferase	4	VTRNA2-1
Helsmoortel–Van der Aa syndrome (ADNP syndrome (Central) (HVDAS_C)	<i>ADNP</i>	Transcription factor	14	PPIEL:Ex1
Helsmoortel–Van der Aa syndrome (ADNP syndrome (Terminal) (HVDAS_T)	<i>ADNP</i>	Transcription factor	21	PPIEL:Ex1
Hunter–McAlpine craniiosynostosis syndrome (HMA)	<i>NPM1</i>	Histone chaperone	8	SNRPN:Int1-DMR2
	<i>UIMC</i>	Histone reader		
	<i>NSD1</i>	Histone methyltransferase		
Immunodeficiency, centromeric instability, facial anomalies syndrome 1 (ICF1)	<i>DNMT3B</i>	DNA methyltransferase	8	DIRAS3:Ex2
				HTR5A:TSS
				IGF1R:Int2
				MAGEL2:TSS
				MKRN3:TSS NDN:TSS
				SNRPN:alt-TSS
				SNRPN:Int1-DMR2
Immunodeficiency, centromeric instability, facial anomalies syndrome 2,3,4 (ICF2-3-4)	<i>ZBTB24</i> , <i>CDCA7</i> , <i>HELLS</i>	Transcription factor and chromatin remodelers	7	ZNF597:TSS
				MKRN3:TSS NDN:TSS
				SNRPN:alt-TSS
				SNRPN:Int1-DMR1
				SNRPN:Int1-DMR2
Intellectual developmental disorder, X-linked, syndromic, Snyder–Robinson type (MRXSSR)	<i>SMS</i>	Enzyme	17	VTRNA2-1
Luscan–Lumish syndrome (LLS)	<i>SETD2</i>	Histone methyltransferase	4	INPP5F:Int2
				NNAT:TSS PPIEL:Ex1
Rahman syndrome (RMNS)	<i>HIST1H1E</i>	Linker histone	9	NNAT:TSS
Tatton–Brown–Rahman syndrome (TBRS)	<i>DNMT3A</i>	DNA methyltransferase	30	NNAT:TSS
Williams–Beuren deletion syndrome (Williams)	<i>BAZ1B</i>	Histone phosphorylation	22	IGF1R:Int2
	<i>GTF2I</i>	Transcription factor		

that can be used as diagnostic biomarkers in these disorders [8].

Several of the environmental and genetic conditions investigated had concordant effects on multiple iDMRs, suggesting common underlying molecular mechanisms and supporting the hypotheses of Imprinted Gene Network [4, 46]. It is important to note that all the genetic and environmental factors analyzed in this study were associated with iDMR methylation changes between 10% and 20% and are not comparable to those observed in the individuals affected by ImpDis, which were excluded in our screening. However, it is possible that some of the factors causing subtle changes in iDMR methylation in the general population or in individuals carrying

mutations in epigenetic modifiers contribute to the imprinting disorder cases with multi-factorial etiology.

A limitation of this study is the lack of reliable data on imprinted gene expression in blood cells. Because most of the imprinted genes are poorly expressed, it is difficult to draw firm conclusions on the impact of iDMR methylation variability in blood cell types.

Conclusions

The faithful maintenance of genomic imprinting throughout cell division and differentiation is necessary for human health. However, the genetic elements and factors that more stably maintain their epigenetic characteristics and the molecular mechanisms controlling this process

are poorly defined. By looking at the methylation level of the imprinted loci in a large human cohort and the association of the imprinting methylation changes with many environmental and clinical variables, we have determined the CpGs of the imprinting control regions that are more resistant and those that change more frequently their methylation and demonstrated many genetic and non-genetic factors affecting their epigenetic variability. We hope this research will improve diagnostic tests and shed some light on the aetiology of human diseases with disturbances of genomic imprinting.

Abbreviations

ADCADN	Cerebellar ataxia, deafness, and narcolepsy, autosomal dominant syndrome
ART	assisted reproduction technology
BCTC	blood cell-type composition
BWS	Beckwith Wiedemann syndrome
ChIPseq	Chromatin-Immunoprecipitation sequencing
CSS4_c.2650	Coffin-Siris syndrome 4
DDs	developmental disorders
DS	Down syndrome
EKD	EpiSign Knowledge Database
EWAS	Epigenome-wide Association Study
FLHS	Floating-Harbour syndrome
gDMR	germline DMR
HMA	Hunter-McAlpine craniostosis syndrome
ICF1	Immunodeficiency, centromeric instability, facial anomalies syndrome 1
ICF2,3,4	Immunodeficiency, centromeric instability facial anomalies syndrome 2,3,4
ICRs	imprinting control regions
iDMRs	imprinted differentially methylated regions
ImpDis	Imprinting Disorders
LLS	Luscan-Lumish syndrome
LOWvar	CpGs displayed $sd < 0.05$ and mean methylation level between 40% and 60%
ME/CFS	myalgic encephalomyelitis/chronic fatigue syndrome
MLID	multi-locus imprinting disturbance
mQTL	DNA methylation quantitative trait loci
MSDvar CpGs	CpGs meeting the criteria of both the SDvar and Mvar CpGs
Mvar CpGs	CpGs with mean methylation outside the [40%, 60%] range
PfHxS	perfluorohexane sulfonate
PFOA	perfluorooctanoate
SCMC	Sub-Cortical Maternal Complex
Sd	standard deviation
sDMRs	somatic or secondary DMRs
SDvar	CpGs CpGs with $sd > 0.05$
UPD	uniparental disomy
Williams	Williams-Beuren deletion syndrome

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13072-025-00612-7>.

Supplementary Material 1

Supplementary Material 2

Author contributions

Fr.C., R.R., B.S. and A.R. designed experiments and interpreted results. Fr.C., R.R., M.L., H.M.C., A.V., L.P., C.G., E.D'A., S.S., Ab.S., A.S., F.C., B.H.M. and C.A. and performed experiments. Fr.C., R.R., B.S. and A.R. wrote the manuscript with input from all authors. All authors reviewed the manuscript.

Funding

This work was supported by grants from the Associazione Italiana Ricerca sul Cancro (AIRC; IG 2020 ID 24405), Fondazione Telethon (GMR23T1062), Italian Ministry of University and Research PRIN 2022B2N2BY awarded to AR, Progetto "National Centre for HPC, Big Data and Quantum Computing" within Piano Nazionale di Ripresa e Resilienza spoke 8 15858/2023 awarded to C.A., and Government of Canada through Genome Canada and the Ontario Genomics Institute (OGI-188) awarded to B.S.

Data availability

The DNA methylation data and metadata of the 2,711 individuals used as control population, the human ZFP57, ZNF445 ChIPseq data and the expression microarray dataset in post-mortem DS brains were obtained from the GEO public repository under the accession numbers: GSE55763, GSE115387, GSE200964 and GSE5390, respectively. The CTCF data were downloaded from the methylation array manifest <https://github.com/zhoulab/InfinitumAnnotationV1/raw/main/Anno/EPIC/EPIC.hg19.CTCF.tsv.gz>. The data on association between CpG methylation and complex traits were retrieved and are available from the EWAS Atlas (<https://ngdc.cncb.ac.cn/ewas/atlas/index>). Part of the DNA methylation data and metadata was obtained from the GEO public repository and are available under the following accession numbers: GSE55763 (control population) and GSE200964 (ZFP57 peaks summit). Individual genomic, epigenomic, or any other personally identifiable data that have not previously been made publicly available for samples in the EKD are prohibited from deposition in publicly accessible databases due to institutional and ethical restrictions. All further data generated in this study, including iDMR methylation level in the general population, list of iDMR CpGs associated with complex traits, iDMR CpG methylation levels in blood cell types and iDMR methylation levels in DDs are available in the Supplementary Tables. The custom scripts are available from the GitHub repository: <https://github.com/ngsFC/EWAS>.

Declarations

Ethics approval and consent to participate

All of the participants provided informed consent prior to sample collection. All of the samples and records were de-identified before any experimental or analytical procedures. The research was conducted in accordance with all relevant ethical regulations. All experimental methods comply with the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

B.S. is a shareholder in EpiSign Inc.

Author details

¹Department of Environmental, Biological and Pharmaceutical Sciences and Technologies (DiSTABiF), Università degli Studi della Campania "Luigi Vanvitelli", Caserta, Italy

²Institute of Genetics and Biophysics (IGB) "Adriano Buzzati-Traverso", Consiglio Nazionale delle Ricerche (CNR), Naples, Italy

³Verspeeten Clinical Genome Centre, London Health Sciences Centre, London, ON, Canada

⁴Department of Pathology and Laboratory Medicine, Western University, London, ON, Canada

⁵Department of Biology, University of Naples "Federico II", Napoli 80126, Italy

⁶Istituto per le Applicazioni del Calcolo (IAC) "Mauro Picone", Consiglio Nazionale delle Ricerche (CNR), Napoli 80131, Italy

Received: 9 April 2025 / Accepted: 15 July 2025

Published online: 24 July 2025

References

- Barlow DP, Bartolomei MS. Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol*. 2014;6(2):a018382.

2. Court F, Tayama C, Romanelli V, Martin-Trujillo A, Iglesias-Platas I, Okamura K, et al. Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res.* 2014;24(4):554–69.
3. Eggermann T, Monk D, de Nancrales GP, Kagami M, Giabicani E, Riccio A, et al. Imprinting disorders. *Nat Rev Dis Primer.* 2023;9(1):33.
4. Monk D, Mackay DJG, Eggermann T, Maher ER, Riccio A. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat Rev Genet.* 2019;20(4):235–48.
5. Mackay DJG, Gazdagh G, Monk D, Brioude F, Giabicani E, Krzyzewska IM et al. Multi-locus imprinting disturbance (MLID): interim joint statement for clinical and molecular diagnosis. *Clin Epigenetics* [Internet]. 2024 Dec [cited 2025 Mar 17];16(1). Available from: <https://clinicalepigeneticsjournal.biomedcentral.com/articles/https://doi.org/10.1186/s13148-024-01713-y>
6. Eggermann T, Yapici E, Bliet J, Pereda A, Begemann M, Russo S et al. Trans-acting genetic variants causing multilocus imprinting disturbance (MLID): common mechanisms and consequences. *Clin Epigenetics* [Internet]. 2022 Mar 16 [cited 2025 Mar 17];14(1). Available from: <https://clinicalepigeneticsjournal.biomedcentral.com/articles/https://doi.org/10.1186/s13148-022-01259-x>
7. Rooney K, Sadikovic B. DNA methylation epigenotypes in neurodevelopmental disorders associated with large structural copy number variants: clinical implications. *Int J Mol Sci.* 2022;23(14):7862.
8. Levy MA, McConkey H, Kerkhof J, Barat-Houari M, Bargiacchi S, Biamino E, et al. Novel diagnostic DNA methylation epigenotypes expand and refine the epigenetic landscapes of Mendelian disorders. *Hum Genet Genomics Adv.* 2022;3(1):100075.
9. Xiong Z, Yang F, Li M, Ma Y, Zhao W, Wang G, et al. EWAS open platform: integrated data, knowledge and toolkit for epigenome-wide association study. *Nucleic Acids Res.* 2022;50(D1):D1004–9.
10. Chen Y, an, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8(2):203–9.
11. de Tribolet-Hardy J, Thorball CW, Forey R, Planet E, Duc J, Coudray A, et al. Genetic features and performance of human KRAB-zinc finger proteins. *Genome Res.* 2023;33(8):1409–23.
12. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol.* 2015;16(1):37.
13. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature.* 2017;541(7635):81–6.
14. Marttila S, Tamminen H, Rajic S, Mishra PP, Lehtimäki T, Raitakari O, et al. Methylation status of VTRNA2-1/nc886 is stable across populations, monozygotic twin pairs and in majority of tissues. *Epigenomics.* 2022;14(18):1105–24.
15. Romanelli V, Nakabayashi K, Vizoso M, Moran S, Iglesias-Platas I, Sugahara N, et al. Variable maternal methylation overlapping the nc886/vRNA2-1 locus is locked between hypermethylated repeats and is frequently altered in cancer. *Epigenetics.* 2014;9(5):783–90.
16. Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell.* 2011;44(3):361–72.
17. Imbeault M, Helleboid PY, Trono D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature.* 2017;543(7646):550–4.
18. Takahashi N, Coluccio A, Thorball CW, Planet E, Shi H, Offner S, et al. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev.* 2019;33(1–2):49–54.
19. Wang J, Zhuang J, Iyer S, Lin X, Whitfield TW, Greven MC, et al. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* 2012;22(9):1798–812.
20. Kerkhof J, Rastin C, Levy MA, Relator R, McConkey H, Demain L, et al. Diagnostic utility and reporting recommendations for clinical DNA methylation epigenotype testing in genetically undiagnosed rare diseases. *Genet Med Off J Am Coll Med Genet.* 2024;26(5):101075.
21. De Cock L, D'haenens E, Vantomme L, Backers L, Beyens A, Claes KB, et al. Cracking rare disorders: a novel minimally invasive RNA-seq protocol. *NPJ Genomic Med.* 2025;10(1):45.
22. Lockstone HE, Harris LW, Swatton JE, Wayland MT, Holland AJ, Bahn S. Gene expression profiling in the adult down syndrome brain. *Genomics.* 2007;90(6):647–60.
23. Jelinic P, Shaw P. Loss of imprinting and cancer. *J Pathol.* 2007;211(3):261–8.
24. Monk D, Arnaud P, Apostolidou S, Hills FA, Kelsey G, Stanier P, et al. Limited evolutionary conservation of imprinting in the human placenta. *Proc Natl Acad Sci U S A.* 2006;103(17):6623–8.
25. Xu Y, Goodyer CG, Deal C, Polychronakos C. Functional polymorphism in the parental imprinting of the human IGF2R gene. *Biochem Biophys Res Commun.* 1993;197(2):747–54.
26. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet.* 2009;41(2):178–86.
27. Skvortsova K, Masle-Farquhar E, Luu PL, Song JZ, Qu W, Zotenko E, et al. DNA hypermethylation encroachment at CpG island borders in Cancer is predisposed by H3K4 monomethylation patterns. *Cancer Cell.* 2019;35(2):297–e3148.
28. Koestler DC, Christensen BC, Karagas MR, Marsit CJ, Langevin SM, Kelsey KT, et al. Blood-based profiles of DNA methylation predict the underlying distribution of cell types. *Epigenetics.* 2013;8(8):816–26.
29. Zhang Z, Stolrow HG, Christensen BC, Salas LA. Down syndrome altered cell composition in blood, brain, and buccal swab samples profiled by DNA-Methylation-Based cell-Type Deconvolution. *Cells.* 2023;12(8):1168.
30. Aref-Eshghi E, Kerkhof J, Pedro VP, Groupe DI, France, Barat-Houari M, Ruiz-Pallares N, et al. Evaluation of DNA methylation epigenotypes for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. *Am J Hum Genet.* 2020;106(3):356–70.
31. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. *Aging Cell.* 2015;14(6):924–32.
32. Husquin LT, Rotival M, Fagny M, Quach H, Zidane N, McEwen LM, et al. Exploring the genetic basis of human population differences in DNA methylation and their causal impact on immune gene regulation. *Genome Biol.* 2018;19(1):222.
33. Marttila S, Rajic S, Ciantar J, Mak JKL, Junttila IS, Kummola L, et al. Biological aging of different blood cell types. *GeroScience.* 2025;47(1):1075–92.
34. Hedera P. Hereditary Spastic Paraplegia Overview. In: Adam MP, Feldman J, Mirzaz GM, Pagon RA, Wallace SE, Amemiya A, editors. *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle; 1993 [cited 2025 Mar 17]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1509/>
35. Ye M, Reyes Palomares A, Iwarsson E, Oberg AS, Rodriguez-Wallberg KA. Imprinting disorders in children conceived with assisted reproductive technology in Sweden. *Fertil Steril.* 2024;122(4):706–14.
36. Canton APM, Macedo DB, Abreu AP, Latronico AC. Genetics and epigenetics of human pubertal timing: the contribution of genes associated with central precocious puberty. *J Endocr Soc.* 2025;9(2):bvae228.
37. Mantovani G, Bastepe M, Monk D, de Sanctis L, Thiele S, Usardi A, et al. Diagnosis and management of pseudohypoparathyroidism and related disorders: first international consensus statement. *Nat Rev Endocrinol.* 2018;14(8):476–500.
38. Muskens IS, Li S, Jackson T, Elliot N, Hansen HM, Myint SS, et al. The genome-wide impact of trisomy 21 on DNA methylation and its implications for hematopoiesis. *Nat Commun.* 2021;12(1):821.
39. Weng X, Liu F, Zhang H, Kan M, Wang T, Dong M, et al. Genome-wide DNA methylation profiling in infants born to gestational diabetes mellitus. *Diabetes Res Clin Pract.* 2018;142:10–8.
40. Wang X, Liu Y, Wu Y, Lin C, Yang S, Yang Y, et al. Methylation alterations of imprinted genes in different placental diseases. *Clin Epigenetics.* 2024;16(1):132.
41. Goovaerts T, Steyaert S, Vandenbussche CA, Galle J, Thas O, Van Criekinge W, et al. A comprehensive overview of genomic imprinting in breast and its deregulation in cancer. *Nat Commun.* 2018;9(1):4120.
42. Cecere F, Pignata L, Hay Mele B, Saadat A, D'Angelo E, Palumbo O, et al. Co-Occurrence of Beckwith-Wiedemann syndrome and Early-Onset colorectal Cancer. *Cancers.* 2023;15(7):1944.
43. Martin-Herranz DE, Aref-Eshghi E, Bonder MJ, Stubbs TM, Choufani S, Weksberg R, et al. Screening for genes that accelerate the epigenetic aging clock in humans reveals a role for the H3K36 methyltransferase NSD1. *Genome Biol.* 2019;20(1):146.
44. Velasco G, Grillo G, Touleimat N, Ferry L, Ivkovic I, Ribierre F, et al. Comparative methylome analysis of ICF patients identifies heterochromatin loci that require ZBTB24, CDCA7 and HELLS for their methylated state. *Hum Mol Genet.* 2018;27(14):2409–24.
45. Xu Q, Xiang Y, Wang Q, Wang L, Brind'Amour J, Bogutz AB, et al. SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat Genet.* 2019;51(5):844–56.

46. Al Adhami H, Evano B, Le Digarcher A, Gueydan C, Dubois E, Parrinello H, et al. A systems-level approach to parental genomic imprinting: the imprinted gene network includes extracellular matrix genes and regulates cell cycle exit and differentiation. *Genome Res.* 2015;25(3):353–67.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.