Age-specific DNA methylation alterations in sperm at imprint control regions may contribute to the risk of autism spectrum disorder in offspring

Eugenia Casella[1,*], Jana Depovere[1,*], Chantal Delger[1], Mariia Butynets[1], Philipp Antczak[2,3,4], Thomas Price[5], Randy L. Jirtle[6], Susan K. Murphy[7], Cathrine Hoyo[6], Adelheid Soubry[1]

[1]Department of Human Genetics, Epigenetic Epidemiology Lab, Faculty of Medicine, KU Leuven, University of Leuven, Leuven 3000, Belgium
[2]Department II of Internal Medicine, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne 50937, Germany
[3]Center for Molecular Medicine Cologne (CMMC), Cologne 50931, Germany
[4]Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne and University Hospital Cologne, Cologne 50931, Germany
[5]Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC 27701, USA
[6]Department of Biological Sciences, Center for Human Health and the Environment, North Carolina State University, Raleigh, NC 27633, USA
[7]Department of Obstetrics and Gynecology, Division of Reproductive Sciences, Duke University Medical Center, Durham, NC 27701, USA
[*]Co-first authors

Correspondence to: Adelheid Soubry; email: adelheid.soubry@kuleuven.be

ABSTRACT

Research findings suggest that advanced paternal age is associated with an increased risk of autism spectrum disorder (ASD) in children. The biological process behind this father-to-child inheritance of a disease may be driven by sperm epigenetic marks. This has been suggested earlier, but the identification of epigenomic regions responsible for these age-related responses have not been further elaborated. To identify sperm-specific marks, we conducted an epigenome-wide association study in sperm from 63 men, using the Illumina 450K array. Linear regression modeling was applied to identify differentially methylated CpGs (DMCs) by age; we controlled for body mass index, patient status, and multiple testing. We found 14,622 statistically significant age-related DMCs; most (69%) were inversely correlated. We identified 95 imprinted genes and emphasized 747 age-related DMCs adjacent to an imprint control region (ICR). Altered methylation patterns in ICRs may result in disturbed expression of imprinted genes and are suspected to be at the origin of several diseases in offspring, including neurodevelopmental disorders. Mapping our results to other databases revealed the following set of imprinted genes linked to ASD: *OTX1, PRDM16, PTPRN2, B4GALNT4, KCNQ1, KCNQ1OT1, DLGAP2, PLAGL1, GNAS, GRB10, MAGEL2, CDH24,* and *FBRSL1.* Further research on these genes could help understand the contribution of paternal age on the development of autism. A change in DNA methylation levels in ICRs before conception may contribute to the heterogeneity and complexity of ASD. Measured DNA

methylation effect sizes were subtle, but small epigenetic disturbances in sperm may be important on a population level, especially if men continue delaying fatherhood. Public health would benefit from the development of preventive and educational programs.

INTRODUCTION

Animal models and human epidemiological data indicate that male ageing can interfere with multiple health outcomes, including fertility, embryo growth, and pregnancy success [1–3]. If pregnancy succeeds, older men have an increased risk to father a child with a psychiatric disorder, compared to younger fathers [4]. A worrisome observation is that these findings have been reported within "normal age-ranges" of fatherhood for the first time [5]. Exploring the role of preconceptional age in autism is a time-wise question. Especially because autism diagnosis has risen dramatically over the last decade. Between 2011 and 2022 rates of ASD diagnosis in the US increased by 175% in children between 5 and 8 years old [6]. Within the same period of time, the average age of American men fathering a first child shifted from 30.8 to 32.1 years [7]. A genome-wide DNA methylation study in sperm from mice indicated an age-related decrease in DNA methylation in regions associated with transcriptional regulation. Offspring of older males exhibited reduced exploratory behaviors and showed a transcriptional dysregulation at genes linked to schizophrenia and autism [3]. A link between advanced age of the father and autism in offspring has been demonstrated through multiple studies [8–10]. However, the underlying biological or epigenetic processes are not well understood. A few genome-wide association studies have provided evidence that sperm DNA methylation in human is sensitive to ageing [11–14]. Some were designed to predict male age from a semen sample (e.g. in forensics). Limitations of these studies included the use of distinct populations, such as mixtures of fertility patients, sperm donors and men recruited from the general population. Other reports used limited numbers of young and healthy participants; and populations also included smokers [12]. Smoking is a known confounder in age-related epigenetic studies [14]. Other factors may also affect age-DNA methylation associations, such as body mass index (BMI), diet, and medical conditions [15, 16]. In brief, most published data from age-related epigenetic studies in sperm were not designed -or their analytic approach was not fully optimized- to identify genes involved in inheritance or offspring health.

In this publication we perform a genome-scale DNA methylation study and comprehensive analysis using the Infinium HumanMethylation450 array to identify age-associated alterations in sperm that may possibly be transmitted to offspring. Our male population largely consists of healthy and non-smoking volunteers. Considering current knowledge on imprint control regions (ICRs), whose parent-of-origin CpG methylation patterns are established during gametogenesis and resistant to postfertilization epigenetic reprogramming, we specifically explore these regions for age-related alterations [17]. ICRs regulate imprinted genes and potentially acquired imprint instability from pre-conceptional exposures (including ageing and age-related factors) may be carried on to the next generation, increasing the risk for development of a chronic disorder. We hypothesize that a better understanding of age-related DNA methylation patterns in the male germ line will provide valuable information about father-to-child inheritance of diseases.

RESULTS

Characteristics of study participants and semen samples

Because of variability in human populations, we studied a subgroup of the general population. Non-smoking men with a healthy reproductive profile were recruited at the Duke Fertility Center. Study design and inclusion or exclusion criteria are illustrated in Supplementary Figure 1. Socio-demographic data of our study population is shown in Table 1. In brief, a total of 63 men between 18 to 35 years old was included in our statistical analysis. Men aged less than 25 years old represented 44% ($n = 28$), 27% ($n = 17$) were between 25 and 29 years old, and nearly 29% ($n = 18$) were between 30 and 35 years old. Most men lacked a graduate degree ($n = 46$, 74.1%) and had not fathered a child ($n = 55$, 87.3%). The majority were healthy volunteers from outside the clinic ($n = 48$, 76.2%). The mean age of those that were patients was 24 years, while that of the non-patient subgroup was 30 years; this difference was significant (p-value <0.001). Abnormal clinical sperm parameters were found in 16 men (25.4%); 5 in the patient subgroup and 11 in the non-patient subgroup. About one-third ($n = 20$) was categorized as being overweight or obese. Having a BMI of 25 or more was strongly associated with older age (p-value = 0.001).

Distributions and categorization of DNA methylation outcomes in sperm

Primary analyses of DNA methylation percentages in sperm on the Illumina array, such as global means and

Table 1. Socio-demographic data of TIEGER participants and semen characteristics.

Variables	Categorization	*n*	%
Age (years) **(median: 25.00, mean: 25.48)**	18–24	28	44.4
	25–29	17	27.0
	30–35	18	28.6
BMI	Normal weight (18 ≤ BMI <25)	43	68.2
	Overweight/obese (BMI ≥25)	20	31.8
Patient at fertility clinic	Yes	15	23.8
	No	48	76.2
Highest degree of education	High school graduate/GED	6	9.7
	Some college	18	29.0
	College graduate	22	35.5
	Graduate degree	16	25.8
Biological children	Yes	8	12.7
	No	55	87.3
Total motile count (TMC)	$\leq 39 \times 10^6$ (abnormal)	10	16.1
	$>39 \times 10^6$	52	83.9
Concentration	$<15 \times 10^6$ (abnormal)	3	4.8
	$\geq 15 \times 10^6$	59	95.2
Motility	<40% (abnormal)	11	17.7
	≥40%	51	82.3
Sperm quality	normal	47	74.6
	abnormal	16	25.4

Legend: The study population includes 63 men between 18 and 35 years old. Variables included in our statistical analysis are shown. Sperm quality is categorized as "abnormal" if at least one of the sperm quality metrics (concentration, motility, TMC) is below the WHO recommended threshold. If the sum is not 63, due to missing data, the percentage was calculated on known data (missing data: TMC: 1; concentration: 1; motility: 1).

individual means, were performed on BMIQ normalized β-values. An overview of our analytic approach is illustrated in Figure 1. In total, we evaluated 482,287 CpG sites. The mean DNA methylation at all CpG sites was 47.63% (β-value = 0.4763). We classified each CpG site by its mean β-value in the following subgroups: unmethylated (UM), hemi-methylated (HM) and fully methylated (FM), as described in the methods section. The number of CpG sites in each methylation subgroup is presented in Table 2. Most CpG sites in sperm were either unmethylated (UM) (46.95%) or fully methylated (FM) (43.02%); the remaining 10.03% was categorized as hemi-methylated (HM). Because β-values have severe heteroscedasticity at low and high methylation values, we used M-values in our validation of individual CpG outcomes. Distributions of the average M-values per site, before and after normalization, are shown as the average densities across 63 samples in Supplementary Figure 2. For the purpose of data interpretation, we also show (and discuss) β-value outcomes.

Global DNA methylation outcomes in sperm by age

After calculating global DNA methylation at all sites, we searched for a potential correlation between age and DNA methylation. No significant relationship was found. Neither if age was used as a continuous variable in a linear regression model (*p*-value = 0.76) (Supplementary Figure 3), nor through stratifying by age (below and above median age) (*p*-value = 0.97) (Supplementary Figure 4). Next, we verified if age-associations could be found after stratification of our population by patient status or by obesity status. Although an opposite trend could be seen by these strata, these correlations were not significant (Supplementary Figures 5 and 6). We further verified if global DNA methylation differed by age within each CpG subgroup (UM, HM and FM). We found no significant correlations; *p*-values were 0.74 at UM sites, 0.92 at HM sites, and 0.96 at FM sites (Supplementary Figure 7).

Site specific DNA methylation outcomes in sperm by age: directions and magnitudes of change

At each CpG site of the 450K array a potential association between DNA methylation and age was evaluated using a linear regression model, adjusting for BMI, patient status and multiple testing. We found a small fraction of CpG sites (3.03%) that was significantly altered by age; 14,622 CpG sites out of 482,287 CpGs could be identified as differentially methylated sites (DMCs). Among those, most DMCs (69.01%; n = 10,091) were negatively associated with age (Table 2). A Volcano plot illustrates all DMCs that are significantly altered by age (Figure 2A). After subdividing our probes by subgroups, associations were as follows: in the UM subgroup, 91.84% of the DMCs were decreased by age (Table 2 and Figure 2B); in the HM subgroup, 74.73% DMCs were negatively associated with age (Table 2 and Figure 2C); and in the FM subgroup, significant DMCs were nearly equally distributed, with 53.3% (n = 2,719) being positively associated and 46.7% (n = 3,099) being negatively associated with age (Table 2 and Figure 2D). We present a short list of highly significant age-related DMCs (top 30 within each subgroup); with highest magnitude in change (absolute delta M-value higher than 0.1) underlined in Table 3.
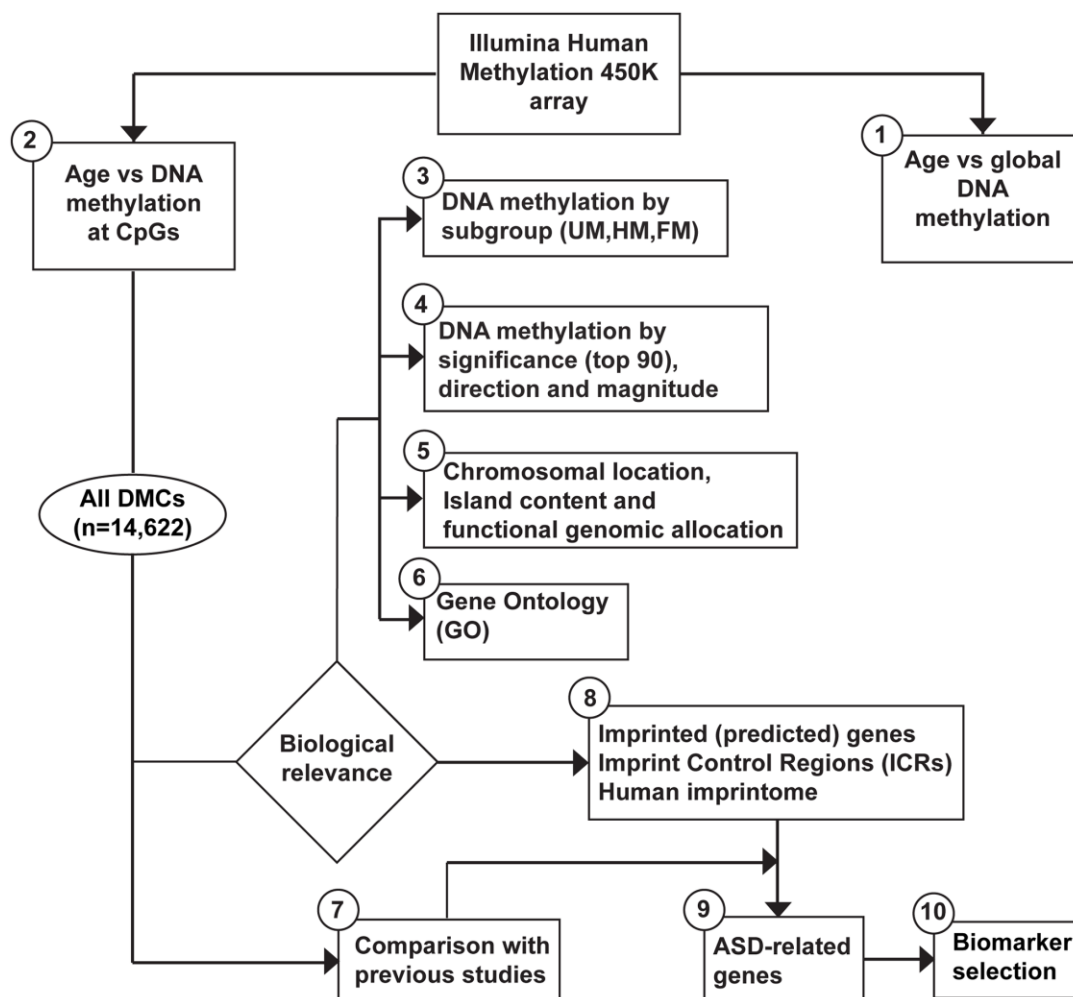

Figure 1. Workflow to assess age-related DNA methylation changes in sperm and biomarker selection for ASD. Data generated from the Illumina HumanMethylation450 BeadChip went through the following analytic procedures. (1) global DNA methylation (mean per subject) by age. (2) DNA methylation at individual CpGs by age. Note, additional tests were performed, as described in the methods section. (3) We subdivided our outcome data as follows: Unmethylated (UM, mean β-value <0.2), Hemi-methylated (HM, 0.20≤ mean β-value ≤0.80), and Fully methylated (FM, mean β-value >0.80), defined by the mean DNA methylation per CpG site (of all subjects). The top 30 of the most significant results (lowest adjusted p-value) were classified by: (4) significance (top 90), direction (positive, negative), and magnitude (absolute Delta-M >0.1). All age-related DMCs (n = 14,622) were analyzed in terms of the following approaches: (5) chromosome location, island content, functional genomic allocation, and (6) gene ontology (GO). We compared our results with: (7) similar published reports, and (8) listed data on imprinting. Finally, (9) a focus was applied on all DMCs using Simons Foundation Autism Reference Initiative database and other publicly available databases on ASD. (10) We selected a set of potential biomarkers for ASD within our set of imprinted genes.

Table 2. Frequencies and distributions of CpG sites in sperm by DNA methylation subgroup and by age.

DNA Methylation subgroup (SG)	β-value	M-value	Analyzed CpG sites		Significant DMCs by age		
			N	*%*	*N*	*%*[1]	*%*[2]
All sites	All	All	482,287	100	14,622 10,091[−]; 4,531[+]	100 69.01[−]; 30.99[+]	3.03 2.09[−]; 0.94[+]
Unmethylated (UM)	<0.2	(⁻inf, −2)	226,432	46.95	2,414 2,217[−]; 197[+]	16.51 91.84[−]; 8.16[+]	1.07 0.98[−]; 0.09[+]
Hemi-methylated (HM)	(0.2–0.8)	(−2, 2)	48,391	10.03	6,390 4,775[−]; 1,615[+]	43.70 74.73[−]; 25.27[+]	13.20 9.87[−]; 3.33[+]
Fully methylated (FM)	>0.8	(2, +inf)	207,464	43.02	5,818 3,099[−]; 2,719[+]	39.79 53.27[−]; 46.73[+]	2.80 1.49[−]; 1.31[+]

According to the mean β-value all CpGs of the 450K are categorized by subgroup (SG): Unmethylated (UM), Hemi-methylated (HM), or Fully methylated (FM) (UM: β-value <0.2; HM: β-value (0.2–0.8); FM: β-value >0.8). Within each subgroup the number of analyzed CpG sites on the array are shown, as well as the number (and %) of significant age-related DMCs (FDR <0.05). (+) number of sites that are increased by age. (−) number of sites that are decreased by age. [1]% relative to the total number of significant DMCs. [2]% relative to the total number of CpGs assessed on the 450K within each subgroup.

In the UM subgroup, 20 DMCs were highly susceptible to ageing with a high DNA methylation impact at the following genes: *HOXA11; HOXA11AS, LRRC25, LHX1, LHX2, LHX5, ZIC4, GRASP, C16orf13, LOC283999, ZNF385A, LOXL3; DOK1, MAGEL2, GNA13, IRF9* and *OSR2*. In the HM subgroup we identified eight DMCs, annotated to the following genes: *NOTCH4, RSPO1, LOXL3, HOXA5, LOC283999, HBA1* and *ESAM*. In the FM subgroup we found four DMCs, corresponding to *LOC283999, FAM184A*, as well as some unidentified (NA) genes. For instance, the gene *LOC283999* showed an opposite direction in DNA methylation change by age at the FM subgroup; this may have biological consequences. Older men may have less sperm cells that are fully methylated at the corresponding site (cg09445803); a ten-year increase in age was associated with a 2.7% decrease in DNA methylation, or a drop from 82.56% to 79.86% (Table 3). As indicated in our methods section, we controlled for BMI, patient status and multiple testing. In our sensitivity analyses (excluding outliers or patients) and in our extended models (including additional co-factors) our results remained consistent with the main analysis (data not shown).

Genomic allocations, island content and other functional characteristics of age-associated DMCs or annotated genes

A Miami plot was created of all significant age-related DMCs across the chromosomes, separated by hyper-methylated sites and hypomethylated sites (Figure 3). A remarkable observation is a general hypomethylation at sites annotated to imprinted genes. Next, we did a parallel analysis of all 14,622 DMCs in terms of their CpG content and neighborhood content (island, shore,

shelf, open sea). CpG islands and open sea areas are the most frequent targets on the microarray, accounting for 30.91% and 36.28% of all analyzed CpG sites, respectively (Supplementary Table 1, Figure 4A). We detected limited age-related DMCs at islands; accounting for nearly 5% (*n* = 722) of the DMCs (Figure 4B), and only 0.48% of all analyzed sites (which is in large contrast to the 30.9% in the complete array) (Supplementary Tables 1 and 2). A comparison of increase (or hypermethylation) *versus* decrease (hypomethylation) in numbers and in ratios - by age, by island and by neighborhood content - are shown in Supplementary Table 2. Only a small difference was seen between the prevalence of negative and positive associations by age (55.7% *versus* 44.3 %, respectively); ratio is 1.26 if all DMCs are included, but it is 2.38 if only UM sites are considered (Supplementary Table 2). Still, an overall ratio of 1.26 at CpG islands is modest if compared to a high ratio of 24.8 at shores (hence, 24.8 times more negative associations than positive associations are measured at shores). An opposite but small ratio (0.40) is seen at open seas (or 2.5 times more positive associations than negative associations).

We further examined variations in DNA methylation in relation to CpG island content by our three subgroups of DNA methylation (Supplementary Table 2 and Figure 4C–4E). Island associated DMCs were predominantly at unmethylated sites (UM), at which 13% of the age-related DMCs could be detected. Calculated ratios indicate that the direction of age-related changes is mostly negative in UM and positive in FM. Out of the 325 UM DMCs at islands, age had a positive effect on a small subset of 96 sites. And, out of the 117 FM DMCs at islands, 36 sites showed a

decrease in DNA methylation. These sites -opposite to the characteristics of these subgroups (being low or high in DNA methylation, respectively)- may be interesting in the context of potential imprinting by parent of origin (discussed below). Most FM sites at shores and at shelves showed a decrease in methylation by age (approximately 95%, 1,611 sites out of 1,700 at shore sites; 81%, 776 out of 961 at shelf sites, respectively). Next, we explored functional genomic allocations (e.g., in subgroups such as promoter, body, 3′UTR, and intergenic) (Supplementary Table 3 and Figure 5). Overall, distributions of age-associated DMCs by functional allocations did not differ substantially from what was expected from the 450K array; although a higher representation of promotors within our DMCs was seen at the UM subgroup (28%) *versus* the FM subgroup (13%). Gene Ontology (GO) term enrichment analysis using all DMCs showed 11 significant GO terms related to male age (Supplementary Figure 8). Specifically, the following terms were captured with the highest enrichment: cell morphogenesis (biological process), and plasma membrane protein complex (cellular component).
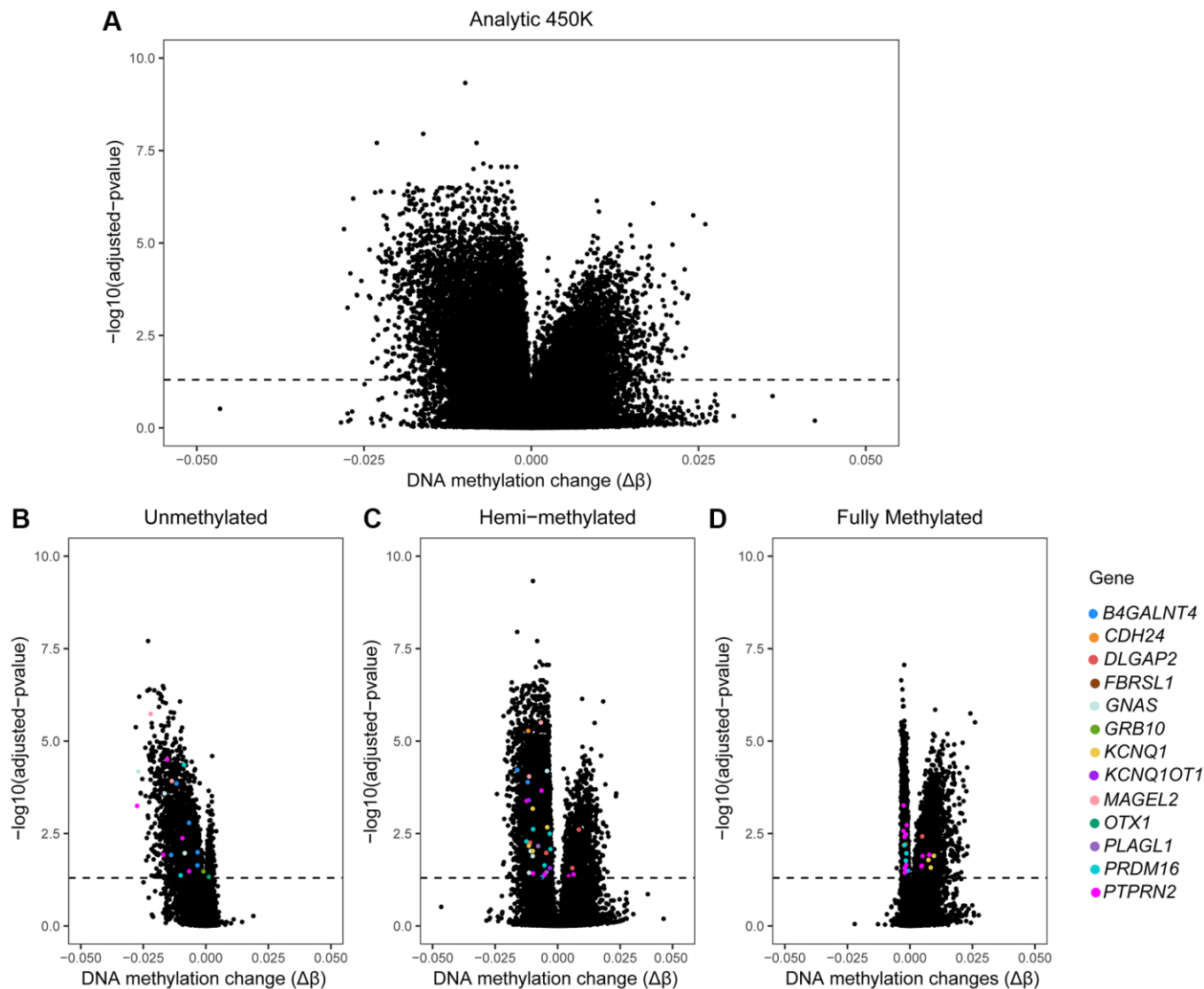

Figure 2. Volcano plots of 450K data by age in sperm. Volcano plots of our age-association study on the 450K array. X-axis: Delta β-values (Δβ) representing age-associated changes in DNA methylation (after linear regression, adjusted for BMI and patient status). Y-axis: logarithmic transformation of the adjusted *p*-value (BH-method). Dashed line: indicates where the adjusted *p*-value is 0.05; dots above the dashed line represent CpG sites where sperm DNA methylation is significantly associated with male age. (**A**) All sites of the 450K are included (*n* = 482,287) and dots above the dashed line are age-associated DMCs (*n* = 14,622). (**B**–**D**) Volcano plots by subgroup of DNA methylation. Colored dots: significant age-related DMCs mapped to genes selected by their potential role in inheritance of ASD from father to child; these include: *OTX1, PRDM16, PTPRN2, B4GALNT4, KCNQ1, KCNQ1OT1, DLGAP2, PLAGL1, GNAS, GRB10*, *MAGEL2, CDH24* and *FBRSL1*. Each colored dot represents one DMC; multiple dots with same color are allocated to the same gene.

Table 3. Alterations at the top 90 most significant DMCs by age in sperm.

SG	Probe ID	Delta M	Mean M	*p*-value	Mean β	Delta β	Gene	Chr	Chr location
<0.2	cg16038003	−0.139	−2.434	2e-08	0.169	−0.023	HOXA11; HOXA11AS	7	chr7:27227520-27229043
<0.2	cg07834682	−0.099	−2.274	3.2e-07	0.179	−0.017	SIX2	2	chr2:45231211-45231482
<0.2	cg11207353	−0.104	−2.834	3.8e-07	0.131	−0.017	LRRC25	19	NA
<0.2	cg01298678	−0.132	−2.994	4e-07	0.123	−0.023	LHX2	9	chr9:126773246-126780953
<0.2	cg06490225*	−0.128	−2.693	4.2e-07	0.146	−0.021	LHX1*	17	chr17:35291899-35300875
<0.2	cg03861097	−0.138	−3.850	4.3e-07	0.073	−0.023	NA	21	chr21:34395128-34400245
<0.2	cg08178072	−0.111	−2.149	5e-07	0.193	−0.019	NA	2	chr2:45164561-45166567
<0.2	cg00896370	−0.106	−2.689	5.4e-07	0.143	−0.018	ZIC4	3	chr3:147115764-147116421
<0.2	cg01447831*	−0.158	−3.373	6.3e-07	0.102	−0.027	NA*	7	chr7:155246390-155251955
<0.2	cg22566906	−0.103	−2.292	6.3e-07	0.177	−0.018	GRASP	12	chr12:52400467-52401696
<0.2	cg25156915	−0.129	−4.268	7.2e-07	0.055	−0.018	NA	4	chr4:174427891-174428192
<0.2	cg16757281*	−0.109	−4.821	8.5e-07	0.038	−0.010	C16orf13*	16	chr16:682634-687106
<0.2	cg04856022	−0.100	−2.603	8.7e-07	0.148	−0.017	PPT2	6	chr6:32121829-32122529
<0.2	cg16872071	−0.094	−2.500	9.4e-07	0.158	−0.016	RALGDS	9	chr9:135995969-135996954
<0.2	cg24764168ˢ	−0.143	−2.199	1.2e-06	0.196	−0.019	NA	6	chr6:1381743-1385211
<0.2	cg14699734	−0.085	−2.491	1.7e-06	0.157	−0.014	NA	19	chr19:5803734-5806023
<0.2	cg02675859	−0.102	−3.796	1.7e-06	0.072	−0.013	LOC283999	17	chr17:76228110-76228380
<0.2	cg10073842	−0.127	−2.573	1.8e-06	0.156	−0.022	MAGEL2	15	NA
<0.2	cg21619325	−0.095	−2.725	2.4e-06	0.138	−0.015	OSR1	2	chr2:19560963-19561650
<0.2	cg00260116	−0.079	−2.472	2.4e-06	0.158	−0.013	C7orf52	7	chr7:100815484-100816995
<0.2	cg08788246	−0.097	−3.373	3.1e-06	0.094	−0.015	ZNF74	22	chr22:20759743-20760923
<0.2	cg08940570	−0.125	−2.852	3.3e-06	0.133	−0.022	LOXL3; DOK1	2	chr2:74781494-74782685
<0.2	cg17465423	−0.105	−2.530	3.7e-06	0.157	−0.018	ZNF385A	12	chr12:54784900-54785238
<0.2	cg09600715	−0.086	−2.188	3.8e-06	0.186	−0.015	NA	7	chr7:27227520-27229043
<0.2	cg19556208	−0.110	−3.334	4e-06	0.098	−0.017	IRF9	14	NA
<0.2	cg18911395	−0.193	−3.374	4.2e-06	0.112	−0.028	LHX5	12	chr12:113908887-113910681
<0.2	cg17413194	−0.125	−2.593	4.2e-06	0.154	−0.022	GNA13	17	chr17:63051893-63053355
<0.2	cg15287594	−0.089	−3.758	4.3e-06	0.073	−0.010	FITM2	20	chr20:42939450-42940043
<0.2	cg03157027	−0.107	−3.804	5e-06	0.073	−0.017	OSR2	8	chr8:99960497-99961438
<0.2	cg04479580	−0.100	−2.315	5.2e-06	0.176	−0.017	NA	12	chr12:6438272-6438931
(0.2–0.8)	cg02991082	−0.086	0.009	4.7e-10	0.502	−0.010	TSC22D3	X	chrX:106959378-106959914
(0.2–0.8)	cg02494945	−0.098	−1.762	1.1e-08	0.234	−0.016	TBX4	17	chr17:59531723-59535254
(0.2–0.8)	cg26854298	−0.100	0.250	2e-08	0.542	−0.008	NA	4	chr4:174443365-174443948
(0.2–0.8)	cg11671335	−0.104	0.819	7.1e-08	0.634	−0.007	NOTCH4	6	chr6:32163292-32164383
(0.2–0.8)	cg03767475	−0.090	1.785	8.7e-08	0.770	−0.004	COMP	19	chr19:18899037-18902284
(0.2–0.8)	cg01862311ˢ	−0.099	1.156	8.7e-08	0.685	−0.006	NA	8	chr8:26305853-26306825
(0.2–0.8)	cg15032957	−0.061	1.985	8.7e-08	0.796	−0.004	TLE2	19	chr19:3035638-3035872
(0.2–0.8)	cg24049629	−0.086	0.304	9.9e-08	0.551	−0.009	RASSF1	3	chr3:50377803-50378540
(0.2–0.8)	cg04508739	−0.088	1.142	2.3e-07	0.683	−0.006	NA	14	chr14:57264638-57265561

(0.2–0.8)	cg24114899[s]	−0.081	0.972	2.3e-07	0.659	−0.007	*MGC12982; FOXD2*	1	chr1:47902793-47905518
(0.2–0.8)	cg13054119	−0.115	−1.459	2.6e-07	0.276	−0.018	*RSPO1*	1	chr1:38099677-38100864
(0.2–0.8)	cg02021127	−0.100	1.404	2.6e-07	0.719	−0.005	*PIP5KL1*	9	chr9:130692839-130693331
(0.2–0.8)	cg04044203	−0.109	0.059	2.7e-07	0.510	−0.010	*ESAM*	11	chr11:124632063-124633239
(0.2–0.8)	cg10836509	−0.083	−0.834	3.1e-07	0.363	−0.012	NA	1	chr1:226297287-226298586
(0.2–0.8)	cg23674882	−0.117	−0.602	3.2e-07	0.403	−0.012	*LOXL3*	2	chr2:74781494-74782685
(0.2–0.8)	cg05269451[s]	−0.103	−1.157	3.2e-07	0.317	−0.014	NA	2	chr2:172963623-172964135
(0.2–0.8)	cg03107888	−0.084	−0.677	3.2e-07	0.388	−0.012	*HOXA13*	7	chr7:27238690-27240311
(0.2–0.8)	cg06097213[s]	−0.080	−1.002	3.2e-07	0.337	−0.013	*LOC285830*	6	chr6:29716468-29717158
(0.2–0.8)	cg20437892	−0.091	−0.046	3.2e-07	0.492	−0.011	NA	2	NA
(0.2–0.8)	cg00215611	−0.088	0.063	3.2e-07	0.511	−0.009	NA	2	chr2:219860927-219861242
(0.2–0.8)	cg19701577	−0.131	−0.824	3.2e-07	0.371	−0.012	*HOXA5*	7	chr7:27182613-27185562
(0.2–0.8)	cg10919107[cr]	−0.094	0.201	3.2e-07	0.534	−0.009	*PMEPA1*	20	chr20:56227252-56227687
(0.2–0.8)	cg04730768	−0.077	0.710	3.2e-07	0.618	−0.007	*CHAT*	10	chr10:50817095-50817309
(0.2–0.8)	cg23327992	−0.079	−0.472	3.2e-07	0.421	−0.012	*IRF9*	14	NA
(0.2–0.8)	cg20774552	−0.057	0.707	3.3e-07	0.619	−0.006	NA	11	chr11:46410921-46414687
(0.2–0.8)	cg20005923	−0.102	−0.365	3.6e-07	0.440	−0.011	*LOC283999*	17	chr17:76228110-76228380
(0.2–0.8)	cg02825052	−0.076	1.115	3.8e-07	0.681	−0.006	*PICALM*	11	NA
(0.2–0.8)	cg01704105	−0.114	−1.250	3.8e-07	0.305	−0.016	*HBA1*	16	chr16:226173-227254
(0.2–0.8)	cg00176496	−0.075	0.579	3.8e-07	0.597	−0.008	*PITPNM3*	17	NA
(0.2–0.8)	cg01344797[cr]	−0.067	0.977	4e-07	0.661	−0.005	*PSD*	10	chr10:104168558-104169153
>0.8	cg02882504[s]	−0.085	2.748	8.7e-08	0.866	−0.002	NA	16	chr16:3238805-3239492
>0.8	cg02680566	−0.083	2.100	2.3e-07	0.806	−0.004	***OSR1***	2	chr2:19560963-19561650
>0.8	cg08604523[s]	−0.089	2.171	4e-07	0.812	−0.003	***MDK***	11	chr11:46406904-46407441
>0.8	cg01943692	−0.077	2.409	7.7e-07	0.837	−0.003	NA	1	chr1:26686516-26687281
>0.8	cg00019678	−0.065	2.564	1.1e-06	0.852	−0.003	***C12orf34***	12	chr12:110151327-110152758
>0.8	cg09445803	−0.101	2.325	1.2e-06	0.826	−0.003	***LOC283999***	17	chr17:76228110-76228380
>0.8	cg12001078	0.064	2.708	1.4e-06	0.864	0.010	*DAPP1*	4	NA
>0.8	cg18672716	0.144	3.521	1.8e-06	0.907	0.024	NA	2	NA
>0.8	cg24321297	−0.046	2.627	2.8e-06	0.859	−0.002	***ST6GALNAC2***	17	chr17:74580974-74582396
>0.8	cg24888609[s]	0.152	3.321	3.1e-06	0.894	0.026	*FAM184A*	6	NA
>0.8	cg21483922	−0.067	2.220	3.1e-06	0.820	−0.003	***SAMD14***	17	chr17:48206663-48207601
>0.8	cg05525368	−0.062	2.950	3.8e-06	0.883	−0.002	***CDC20***	1	chr1:43824134-43825059
>0.8	cg11433866[s]	−0.070	2.122	3.8e-06	0.809	−0.003	NA	12	chr12:115124729-115125152
>0.8	cg01405985	−0.107	2.441	4.2e-06	0.835	−0.002	NA	19	chr19:36347044-36348101
>0.8	cg24073074	−0.069	2.175	4.9e-06	0.815	−0.003	***WIPI1***	17	NA
>0.8	cg04078221	−0.064	2.042	5.2e-06	0.801	−0.003	NA	1	chr1:47899125-47899398
>0.8	cg18231690	−0.075	2.092	5.3e-06	0.805	−0.004	***ZNF592***	15	chr15:85291079-85291697
>0.8	cg21690489	−0.063	2.621	5.6e-06	0.857	−0.003	***CREB3L1***	11	NA
>0.8	cg08692104	−0.055	3.365	6.3e-06	0.910	−0.001	***ZNF823***	19	chr19:11849359-11849796
>0.8	cg11277190[cr]	0.093	2.905	6.4e-06	0.876	0.015	*FAM184A*	6	NA
>0.8	cg21574186	0.079	3.519	6.4e-06	0.916	0.009	*UNC5C*	4	NA

>0.8	cg00471142	−0.059	3.040	6.5e-06	0.889	−0.002	*LRFN1*	19	chr19:39804621-39805954
>0.8	cg01109643	−0.083	2.152	6.6e-06	0.810	−0.003	NA	13	chr13:113548643-113549127
>0.8	cg24842142	−0.069	2.175	6.9e-06	0.815	−0.003	***RNF39***	6	chr6:30042918-30043500
>0.8	cg14101976	0.081	4.098	7.3e-06	0.942	0.010	*UNC5C*	4	NA
>0.8	cg12391323ˢ	−0.046	2.435	7.5e-06	0.842	−0.003	NA	5	chr5:139027443-139030219
>0.8	cg05878073	−0.067	3.194	7.6e-06	0.899	−0.002	***ABCD4***	14	chr14:74769366-74769815
>0.8	cg11338156	−0.078	2.657	7.9e-06	0.858	−0.003	NA	10	NA
>0.8	cg13244241	−0.082	3.833	8.2e-06	0.931	−0.001	NA	10	chr10:102489343-102491011
>0.8	cg00730441	−0.061	2.187	8.4e-06	0.817	−0.003	***TBX2***	17	chr17:59485573-59485780

Probes related to DMCs by age were first sorted by subgroups (UM: β-value <0.2; HM: β-value (0.2–0.8); FM: β-value >0.8), then by adjusted *p*-values (Benjamini-Hochberg method; FDR <0.05). Only the top 30 most significant results by each subgroup (SG) are presented. Delta-M and Delta-β: estimated changes by age in M-value and in β-value, respectively. Annotated gene names and chromosome locations are shown according to the UCSC Genome Browser. *Probe related to CpG Islands; ᶜʳcross-reactive probe; ˢSNP; (bold) direction is opposite (e.g., hypomethylated in FM subgroup); (underlined) highest magnitude in DNA methylation changes (Delta-M is higher than 0.1).
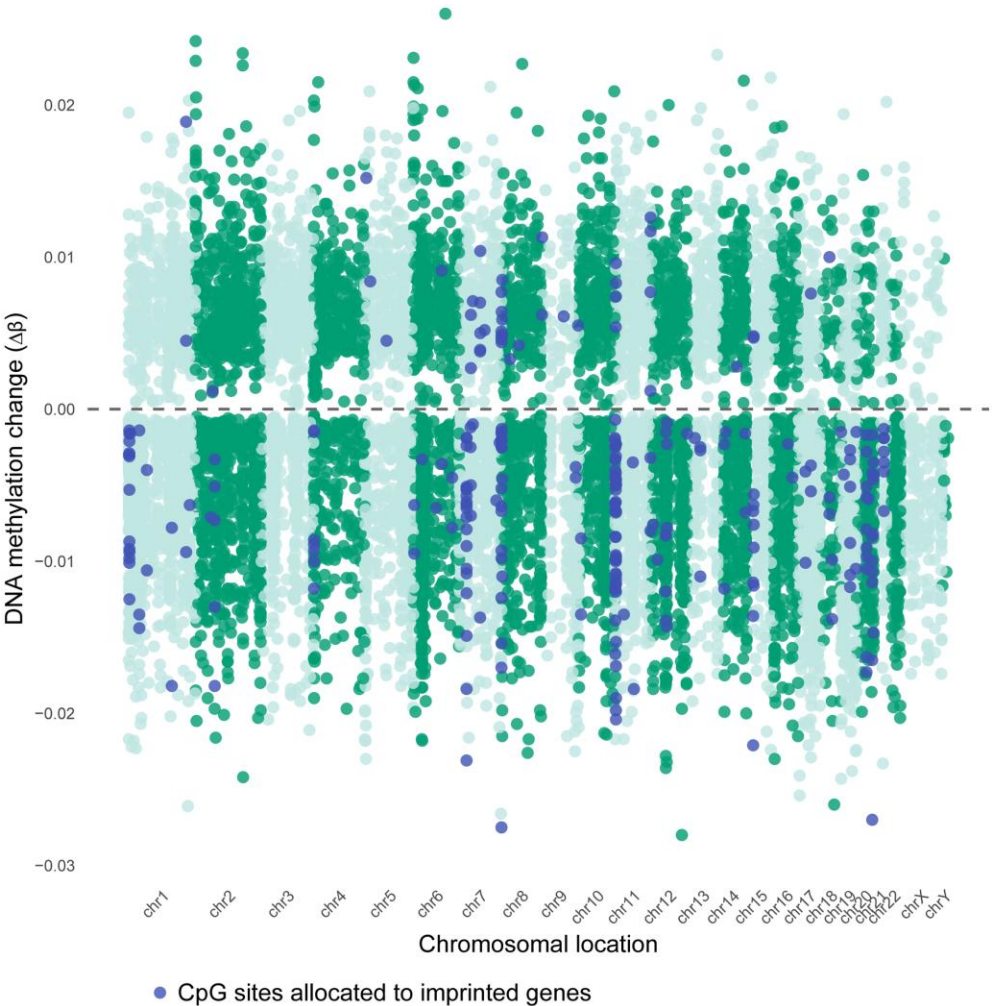

Figure 3. Chromosomal distribution of age-associated DMCs in sperm. Miami plot of age-associated differentially methylated CpG sites (DMCs), shown by direction of change (Delta β >0 relates to an increase in number of sperm cells that are methylated, Delta β <0 means a decrease in number of sperm cells that are methylated). Dashed line: no change (Δβ = 0). Delta β-values are calculated from M-values, as shown in the methods section ($\Delta\beta = \dfrac{2^{(M_0 + \text{delta M})}}{1 + 2^{(M_0 + \text{delta M})}} - \dfrac{2^{M_0}}{1 + 2^{M_0}}$). Significant DMCs correspond to FDR <0.05. Blue dots are sites allocated to imprinted genes, listed in the Geneimprint database.

Identifying age-related DMCs linked to imprinted genes and candidate ICRs

We screened our findings on age-related DMCs for their potential role in imprinting (Figure 1). First, imprinted genes were defined using Geneimprint, with a list of 227 imprinted genes in human; including 104 predicted and 123 confirmed imprinted genes [18]. Out of the 14,622 age-related DMCs, we found 271 DMCs (1.85%) that could be mapped to 95 (predicted) imprinted genes [17]; these genes are listed by subgroup in Figure 6. Because our 450K array includes 215 (potentially) imprinted genes, this means that 44% of these imprinted genes theoretically present on the array are sensitive to influences from ageing. Out of the 95 age-related imprinted genes, 49 are maternally expressed (or paternally imprinted), 42 are paternally expressed (or maternally imprinted), and imprinting is isoform dependent in four of the identified genes (Supplementary Table 4). Second, we extended our approach using a database published by Jima et al. [17], which includes a genomic map of 1,488 putative Imprint Control Regions (ICRs) or so-called Imprintome. Of our 14,622 DMCs, 747 DMCs could be linked to 380 candidate ICRs (Supplementary Table 5). More precisely, out of these 747 ICR-associated DMCs, 152 DMCs were allocated to 94 ICRs with paternal origin of methylation and 74 genes were identified; 193 DMCs were linked to 107 maternally methylated ICRs and 86 genes were found; and, the remaining 402 DMCs were associated to regions with unknown parent-of-origin methylation; corresponding to 179 ICRs or 158 genes (Supplementary Table 5). After a stringent selection of age-related genes based on a combined use of our findings above by Geneimprint list ($n = 95$) and by the Imprintome ($n = 318$), we selected 22 imprinted genes with previously reported ICRs (Figure 6). This set of genes corresponds to 94 age-associated DMCs; 79 DMCs are hypomethylated and 15 DMCs are hyper-methylated (Table 4 and Figure 6). For instance, the highest magnitude at hypomethylated DMCs was measured at the *PTPRN2* gene, being −28% per 10 years of ageing (cg18285788; Delta β = −0.028, $p = $ 5.64e-04); and, the highest magnitude at hyper-
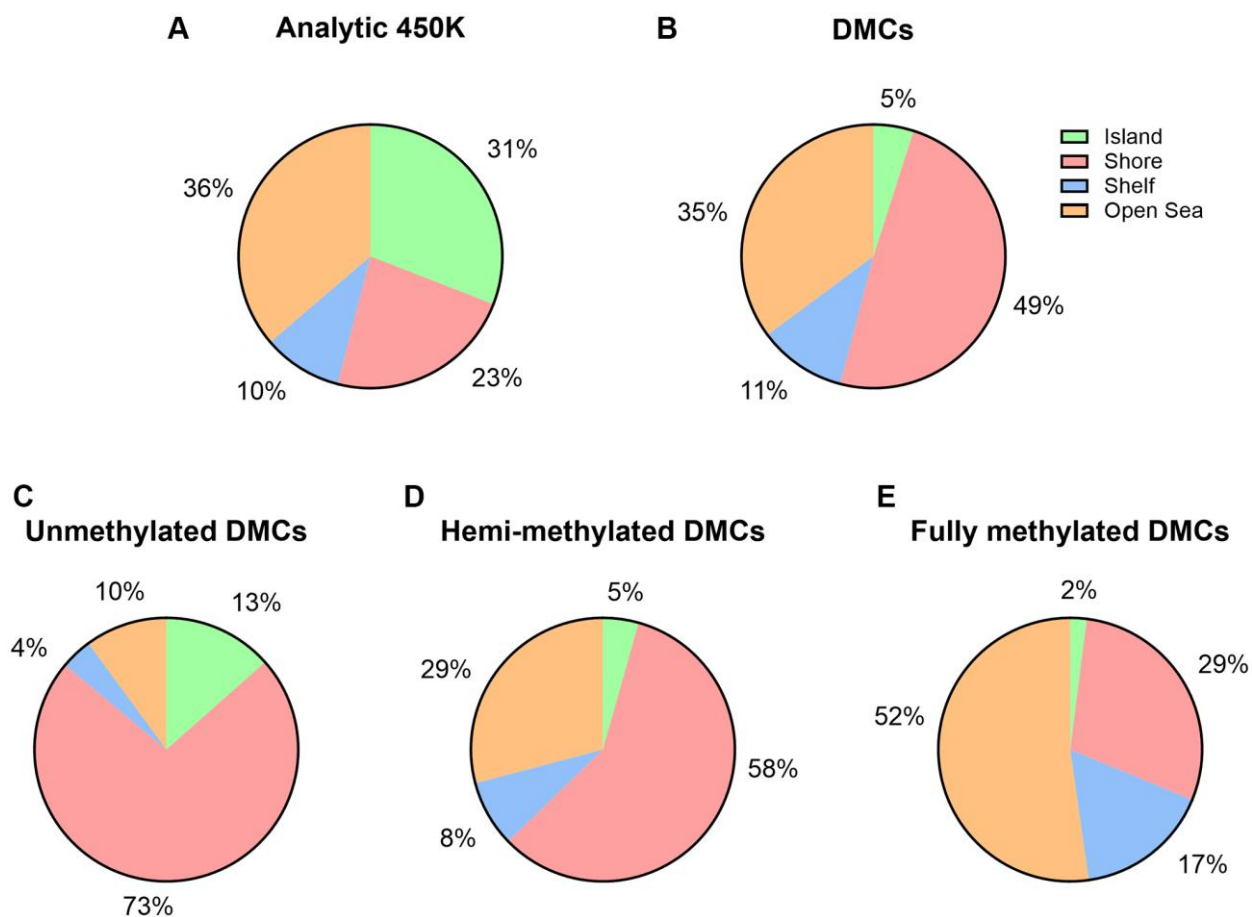

Figure 4. Island content of CpGs within 450K array and age-associated DMCs. (**A**) Percentages of CpGs by island content are displayed with respect to the 450K ($n = 482,287$). (**B**) Percentages of CpGs by island content are displayed with respect to the number of significant age-related DMCs ($n = 14,622$). (**C**–**E**) By DNA methylation subgroup (Abbreviations: UM: unmethylated, mean β-value <0.2; HM: hemi-methylated, 0.20 ≤ mean β-value ≤0.80; FM: fully methylated, mean β-value>0.80).

methylated DMCs was found at the *KCNQ1* gene, being +10% per 10 years of ageing (cg17416793; Delta β = +0.010, p = 0.0127). In our search to define a set of age-related biomarkers of inheritance we could classify these 22 imprinted genes into the following categories: seven genes are paternally imprinted (or maternally expressed) (*H19, KCNQ1DN, PTPRN2, KCNQ1, SVOPL, B4GALNT4, FBRSL1*; the last two are yet predicted imprinted genes); twelve genes are known to be maternally imprinted (or paternally expressed) (*MAGEL2, DIRAS3, FAM50B, DLGAP2, PLAGL1, ZIM2, GLI3, DLK1, CDH24, RB1, SNRPN, KCNQ1OT1*); and, in 3 genes imprinting is isoform dependent (*GNAS, BLCAP, GRB10*). Notably, we repeated our analyses after exclusion of reported SNPs and cross-reactive probes [19, 20]. This test did not significantly change our results, with two exceptions: *ZIM2* and *C6orf145* (indicated in Figure 6 and Tables 3 and 4).

Comparison of the current findings on age-related DMCs with results from similar studies

In a recent review by the research group of Haaf, 2,355 genes have been reported in sperm with age affected DMRs [11]. Considering the current study results and those found by four other observational studies, we retrieved 2,098 age-associated genes in sperm. Although similar techniques (Illumina platforms) were used, study populations and study conditions differed [9–12]. A Venn diagram illustrates the numbers of overlapping age-affected genes identified by each report (Figure 7). We counted 929 genes that were found by at least two studies, 83 genes that were found by at least three studies, seven genes that were found by at least four studies, and one gene that was found by all five studies. (Supplementary Table 6). In brief, six genes found by four studies are *SLC22A18AS, C7orf50, UTS2R, BEGAIN, GRIN1,* and *PCDH15*; and, a single gene identified by all is *DLGAP2*. Apart from *C7orf50*, all genes recurrently found in age-association studies over the last couple of years have been linked to the development of autism (discussed below) [21–26]. The role of *UTS2R (*or *GPR14)* in autism is unclear. It plays a role in various brain functions, but it has been considered as a candidate gene for autism [27, 28]. Overall, our findings are in line with Bernhardt et al.'s conclusion that age-induced DNA methylation alterations in sperm may contribute to the development of neurodevelopmental disorders in offspring [11].
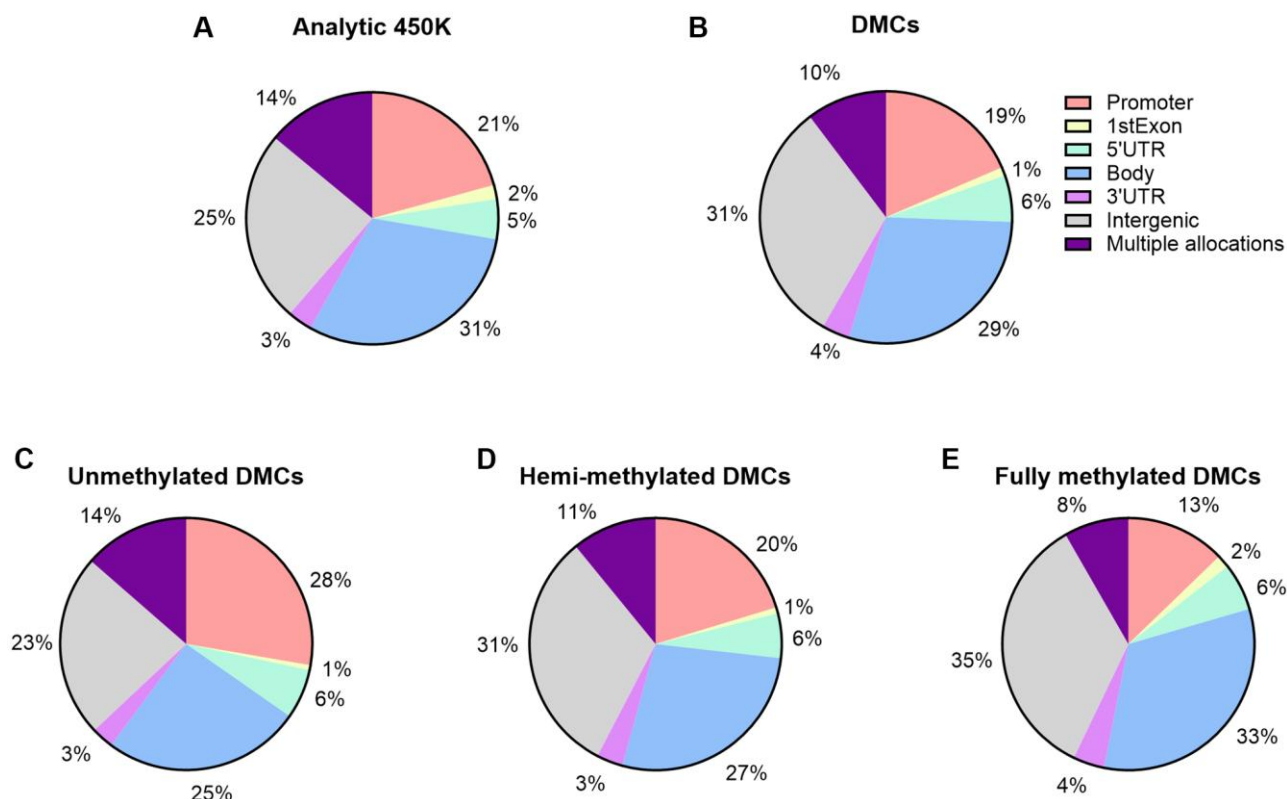

Figure 5. Functional genomic distribution of CpGs within 450K array and age-associated DMCs. (**A**) Percentages of CpGs by functional genomic region are displayed; by promoter, 1stExon, 5'UTR, 3'UTR, body, and intergenic regions. CpG sites allocated to multiple regions are reported as a separate category. (**A**) Percentages with respect to the 450K (n = 482,287). (**B**) Percentages are displayed with respect to the number of significant age-related DMCs (n = 14,622). (**C–E**) By DNA methylation subgroup (Abbreviations: UM: unmethylated, mean β-value <0.2; HM: hemi-methylated, 0.20 ≤ mean β-value ≤0.80; FM: fully methylated, mean β-value >0.80).

Imprinted Genes
227⁻ DMCs, 44⁺ DMCs (95 genes)

	Unmethylated (UM) 2⁺ DMCs , 76⁻ DMCs (45 genes)	Hemi-methylated (HM) 9⁺ DMCs , 109⁻ DMCs (60 genes)	Fully methylated (FM) 33⁺ DMCs , 42⁻ DMCs (45 genes)
M (49 genes)	*HOXA11* PR, *KCNQ1DN* PR, **PTPRN2** *, *OSBPL5* *, **SLC26A10**, **B4GALNT4** PR,*, *HSPA6*, *PHLDA2* PR, *HOXC4* PR, **MAGI2**, *COL9A3* PR, **ANO1**, *RBP5* PR, *CHST8* PR, *WDR8*, *RAB1B* PR, *HOXA5*, **LMX1B**, *C10orf93*, **NTM**, **OTX1** PR,*↑, **ATP10A** PR	*HOXA5, HOXA11, SLC22A18AS, VAX2, HOXC4* PR, **SLC26A10**, *CDKN1C* PR, **B4GALNT4** PR, *HOXA3* PR, *TP73*, **ATP10A** PR, **PTPRN2** *, *CHST8* PR, *IFITM1* PR, **KCNQ1** PR, **UBE3A**, *OSBPL5* PR, **NTM**, **LMX1B**, *BRUNOL4*, *PTPN14*, *PPAP2C* cr, **HOXB3** PR, *SVOPL*, *COL9A3* PR, *LILRB4*, *ACD* PR, *BTNL2* PR, *TCEB3C* PR	*HOXC4* ↓, **SLC26A10** PR↓, *IFITM1* PR↓, **MAGI2**, *CHST8* ↓, *HOXA11* PR↓, **PTPRN2** *↓, **NTM**, **ADAMTS16**, *BRUNOL4*, **FBRSL1** ↓, *H19*, *ATP5EP2* PR↓, **KCNQ1**, **HOXB3**, *SMOC1, PPP1R9A, PTPN14, C10orf91* PR↓, **UBE3A**, *C10orf93* ↓, **B4GALNT4** ↓, *FASTK* ↓, *SALL1* ↓, *FLJ46321* cr
P (42 genes)	**MAGEL2** PR, *GATA3*, **PRDM16** *, *EVX1*, *BMP8B, GDAP1L1* PR, *DIRAS3, SIM2*, *GNASAS, SPON2* PR, **OBSCN** *, *IGF2AS* PR, *IGF2* PR, *PYY2, HECW1* PR↑, *RB1* *, *MRAP2* PR, *FAM50B* PR, **FOXG1** PR, *DNMT1*	*DIRAS3*, **MAGEL2** PR, **CDH24**, *IGF2AS* PR, *IGF2* PR, *WT1*, *GNASAS* PR, *PYY2*, *BMP8B, EVX1* PR, *AIM1*, *SIM2*, **KCNQ1OT1** PR, *SPON2* PR, **PRDM16**, **DLGAP2** *, *ST8SIA1, C15orf2* PR, **PLAGL1** PR, *HECW1*, **GLIS3**, *SNRPN, TMEM52* PR, *GATA3, DLK1* PR, *LIN28B*, *CCDC85A* PR, **C6orf145** PR,cr	*RPL22* PR↓, *ZIM2* s↓, *SPON2* ↓, *GLI3* ↓, *GNASAS* ↓, **DLGAP2**, *SIM2* PR↓, **GLIS3**, **PRDM16** *↓, **CDH24** ↓, *NKAIN3*, *C9orf116, DLK1* ↓, *RB1* ↓, *HECW1*, **RNU5D**, *BMP8B* ↓, *PURG* ↓
I (4 genes)	*GNAS* PR,*, *BLCAP*, **GRB10** PR,*	*BLCAP* PR, **DDC**, *GNAS* PR	*BLCAP* PR↓, **GNAS** ↓

Imprinted genes and ICRs
79⁻ DMCs, 15⁺ DMCs (22 genes)

	Unmethylated (UM) 27⁻ DMCs (10 genes)	Hemi-methylated (HM) 5⁺ DMC, 32⁻ DMCs (14 genes)	Fully methylated (FM) 10⁺ DMCs, 20⁻ DMCs (13 genes)
M (7 genes)	*KCNQ1DN* PR, **PTPRN2** *, **B4GALNT4** PR,*	**B4GALNT4** PR, **PTPRN2** *, **KCNQ1** PR, *SVOPL*	**PTPRN2** *↓, **FBRSL1** ↓, *H19*, **KCNQ1**, **B4GALNT4** ↓
P (12 genes)	**MAGEL2** PR, *DIRAS3, RB1* *, *FAM50B* PR	*DIRAS3*, **MAGEL2** PR, **CDH24**, **KCNQ1OT1** PR, **DLGAP2** *, **PLAGL1** PR, *SNRPN, DLK1* PR	*ZIM2* s↓, *GLI3* ↓, **DLGAP2**, **CDH24** ↓, *DLK1* ↓, *RB1* ↓
I (3 genes)	*GNAS* PR,*, *BLCAP*, **GRB10** PR,*	*BLCAP* PR, **GNAS** PR	*BLCAP* PR↓, **GNAS** ↓

ASD-related genes
(27 genes)

M	*ANO1* *LMX1B* *NTM* *OTX1* PR,*↑ *ATP10A* PR *SLC26A10* PR↓ *UBE3A* *HOXB3* PR *MAGI2* *ADAMTS16*
P	*PRDM16* *↓ *OBSCN* * *FOXG1* PR *GLIS3* *RNU5D*
I	*DDC*

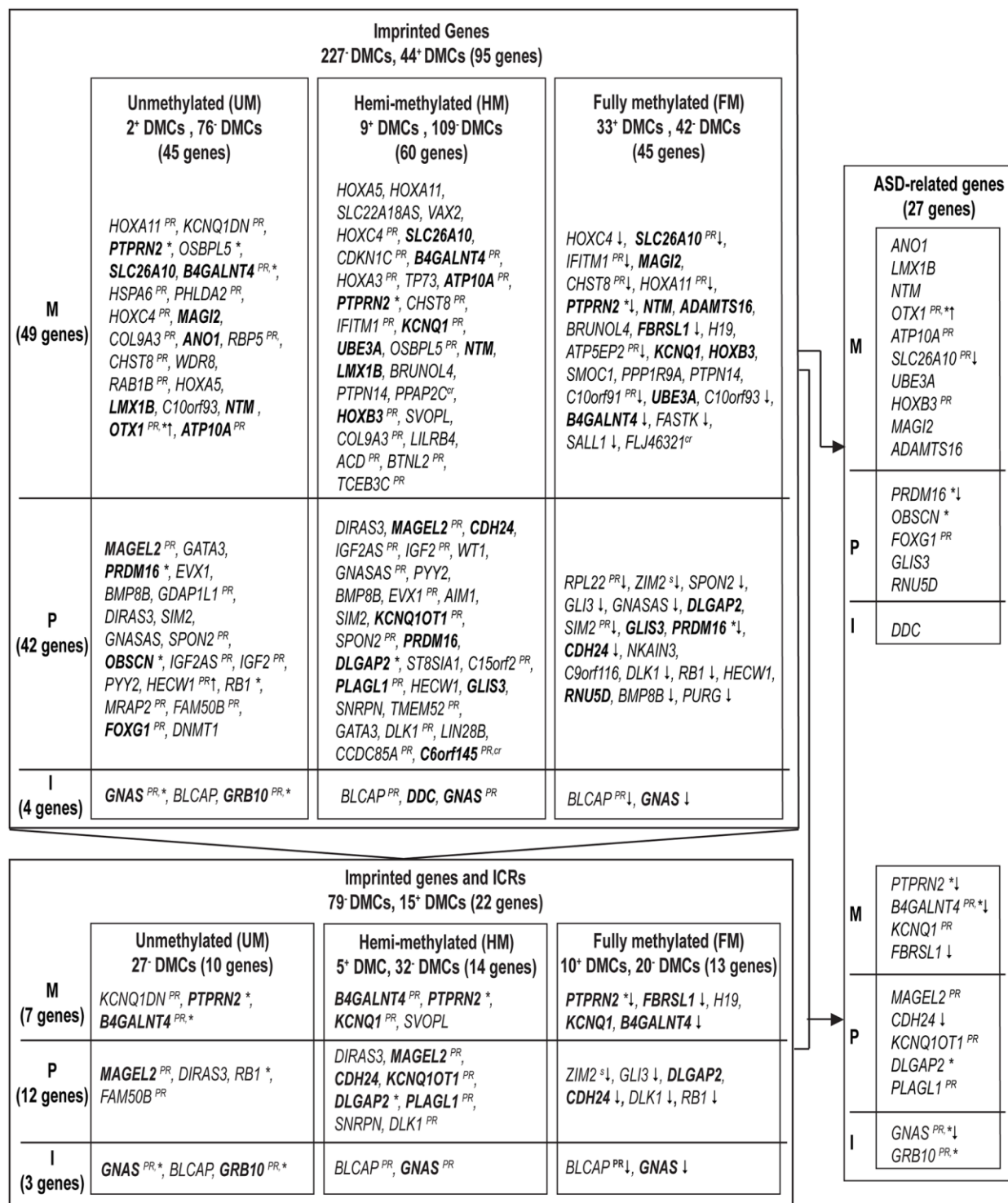|	
M	*PTPRN2* *↓ *B4GALNT4* PR,*↓ *KCNQ1* PR *FBRSL1* ↓
P	*MAGEL2* PR *CDH24* ↓ *KCNQ1OT1* PR *DLGAP2* * *PLAGL1* PR
I	*GNAS* PR,*↓ *GRB10* PR,*

Figure 6. Imprinted genes affected by age in sperm and our approach for biomarker selection to predict autism spectrum disorders in offspring. A summary of (predicted) imprinted genes linked to significant age-associated DMCs is ordered by expressed allele, and significance (results with smallest *p*-values are on top). Upper frame: 95 (predicted) imprinted genes linked to 271 DMCs within DNA methylation subgroups (Abbreviations: UM: unmethylated, mean β-value <0.2; HM: hemi-methylated, 0.20≤ mean β-value ≤0.80; FM: fully methylated, mean β-value >0.80). Lower frame: out of the 95 (predicted) imprinted genes (upper frame) 22 genes have been mapped to ICRs. In bold, genes involved in ASD. Right frame: summary of 27 ASD-related imprinted genes (due to probe cross-reactivity, the gene *C6orf145* with a single DMC was not withheld). The following superscripts are used if at least one DMC had the following characteristic: *, at CpG Islands; PR, at promotor region; ↓ or ↑, an opposite direction in DNA methylation change was measured (increase or decrease in DNA methylation, at UM and FM methylated DMCs, respectively); s, at SNPs; cr, at a cross-reactive probe. M, maternally expressed genes; P, paternally expressed genes; I, isoform dependent.

Table 4. Age-related DMCs and 22 allocated genes in sperm linked to ICRs and imprinting.

SG	EA	Probe ID	Delta β	Mean β	*p*-value	Chr	Gene
<0.2	I	cg21809160	−0.027	0.102	6.65e-05	20	*GNAS*
<0.2	I	cg24058407	−0.017	0.190	2.61e-04	20	*GNAS*
<0.2	I	cg11399589	−0.017	0.139	0.0021	20	*BLCAP*
<0.2	I	cg17696847	−0.008	0.079	0.0107	20	GNAS
<0.2	I	cg27006764	−0.001	0.035	0.0334	7	*GRB10*
<0.2	I	cg13591710	−0.016	0.168	0.0453	20	*BLCAP*
<0.2	I	cg11948874	−0.010	0.104	0.0482	20	*BLCAP*
<0.2	M	cg07439128	−0.017	0.067	2.35e-05	11	*KCNQ1DN*
<0.2	M	cg04937416	−0.015	0.091	3.17e-05	7	*PTPRN2*
<0.2	M	cg10798664	−0.012	0.179	1.40e-04	11	*B4GALNT4**
<0.2	M	cg18285788	−0.028	0.054	5.64e-04	7	*PTPRN2*
<0.2	M	cg26461944	−0.007	0.139	0.0016	11	*B4GALNT4**
<0.2	M	cg01923099	−0.020	0.122	0.0017	11	*KCNQ1DN*
<0.2	M	cg16083838	−0.020	0.087	0.0040	11	*KCNQ1DN*
<0.2	M	cg03983213	−0.009	0.074	0.0042	7	*PTPRN2*
<0.2	M	cg22675922	−0.005	0.060	0.0068	11	*KCNQ1DN*
<0.2	M	cg10726517	−0.003	0.050	0.0101	11	*B4GALNT4**
<0.2	M	cg13538517	−0.017	0.150	0.0119	7	*PTPRN2*
<0.2	M	cg03481077	−0.014	0.193	0.0120	11	*B4GALNT4**
<0.2	M	cg04876474	−0.002	0.043	0.0128	11	*KCNQ1DN*
<0.2	M	cg17682432	−0.003	0.073	0.0229	11	*B4GALNT4**
<0.2	M	cg08242024	−0.007	0.151	0.0330	7	*PTPRN2*
<0.2	P	cg10073842	−0.022	0.156	1.84e-06	15	*MAGEL2*
<0.2	P	cg22872376	−0.014	0.100	1.20e-04	15	*MAGEL2*
<0.2	P	cg23076194	−0.011	0.161	0.0029	1	*DIRAS3*
<0.2	P	cg19447496	−0.011	0.190	0.0225	13	*RB1*
<0.2	P	cg23985641	−0.010	0.177	0.0473	6	*FAM50B*
(0.2–0.8)	I	cg07964163	−0.004	0.782	6.42e-05	20	GNAS
(0.2–0.8)	I	cg24214471	−0.006	0.681	2.87e-04	20	*BLCAP*
(0.2–0.8)	I	cg21733794	−0.008	0.555	6.94e-04	20	*BLCAP*
(0.2–0.8)	I	cg04820254	−0.005	0.723	0.0011	20	*BLCAP*
(0.2–0.8)	I	cg21045560	−0.011	0.351	0.0016	20	*BLCAP*
(0.2–0.8)	I	cg20569652	−0.009	0.282	0.0053	20	*BLCAP*
(0.2–0.8)	I	cg17658854	−0.011	0.207	0.0094	20	*GNAS*
(0.2–0.8)	I	cg01565918	−0.010	0.340	0.0128	20	*GNAS*
(0.2–0.8)	I	cg14235271	−0.011	0.219	0.0365	20	*GNAS*
(0.2–0.8)	M	cg21996245	−0.016	0.202	6.08e-05	11	*B4GALNT4**
(0.2–0.8)	M	cg20846508	−0.012	0.390	1.29e-04	11	*B4GALNT4**
(0.2–0.8)	M	cg24221919	−0.006	0.589	2.18e-04	7	*PTPRN2*
(0.2–0.8)	M	cg18628367	−0.012	0.289	4.13e-04	7	*PTPRN2*
(0.2–0.8)	M	cg03371125	−0.010	0.345	6.71e-04	11	*KCNQ1*
(0.2–0.8)	M	cg04666029	−0.004	0.641	0.0021	11	*KCNQ1*

(0.2–0.8)	M	cg20533553	−0.012	0.221	0.0068	11	*KCNQ1*
(0.2–0.8)	M	cg19698309	−0.010	0.361	0.0094	11	*KCNQ1*
(0.2–0.8)	M	cg20482223	−0.006	0.513	0.0177	7	*SVOPL*
(0.2–0.8)	M	cg19713140	−0.010	0.374	0.0378	7	*PTPRN2*
(0.2–0.8)	M	cg21231189	0.006	0.776	0.0403	7	*PTPRN2*
(0.2–0.8)	M	cg01100465cr	−0.005	0.642	0.0408	7	*PTPRN2*
(0.2–0.8)	M	cg15094119	0.004	0.772	0.0463	7	*PTPRN2*
(0.2–0.8)	M	cg15971656	−0.006	0.374	0.0481	11	*B4GALNT4**
(0.2–0.8)	P	cg20808078	−0.004	0.747	2.87e-06	1	*DIRAS3*
(0.2–0.8)	P	cg25135755	−0.007	0.631	3.14e-06	15	*MAGEL2*
(0.2–0.8)	P	cg09834049	−0.012	0.291	5.29e-06	14	*CDH24**
(0.2–0.8)	P	cg01152488	−0.011	0.440	9.06e-05	15	*MAGEL2*
(0.2–0.8)	P	cg03439898cr	0.009	0.770	0.0025	8	*DLGAP2*
(0.2–0.8)	P	cg20076070	−0.011	0.413	0.0056	8	*DLGAP2*
(0.2–0.8)	P	cg21113768	−0.008	0.464	0.0069	6	*PLAGL1*
(0.2–0.8)	P	cg08082351	−0.005	0.708	0.0106	8	*DLGAP2*
(0.2–0.8)	P	cg26939721	0.005	0.702	0.0147	15	*SNRPN*
(0.2–0.8)	P	cg12818159	0.006	0.785	0.0272	8	*DLGAP2*
(0.2–0.8)	P	cg02566775	−0.005	0.577	0.0352	6	*PLAGL1*
(0.2–0.8)	P	cg09212014	−0.007	0.329	0.0375	14	*DLK1*
(0.2–0.8)	P; M	cg03654058	−0.012	0.331	3.96e-04	11	*KCNQ1OT1; KCNQ1*
(0.2–0.8)	P; M	cg04762676	−0.003	0.764	0.0271	11	*KCNQ1OT1; KCNQ1*
>0.8	I	cg09993814	−0.003	0.830	5.19e-04	20	*BLCAP*
>0.8	I	cg10546626	−0.002	0.891	0.0034	20	GNAS
>0.8	I	cg17820025cr	−0.002	0.832	0.0455	20	*BLCAP*
>0.8	M	cg05821571	−0.003	0.825	5.53e-04	7	*PTPRN2*
>0.8	M	cg04799270	−0.001	0.924	0.0019	7	*PTPRN2*
>0.8	M	cg15012939	−0.003	0.855	0.0027	7	*PTPRN2*
>0.8	M	cg02855778	−0.002	0.904	0.0034	7	*PTPRN2*
>0.8	M	cg02773779	−0.002	0.866	0.0039	7	*PTPRN2*
>0.8	M	cg19100996	−0.002	0.896	0.0061	12	*FBRSL1**
>0.8	M	cg22172494	0.005	0.859	0.0078	11	*H19*
>0.8	M	cg27629384	0.008	0.952	0.0119	7	*PTPRN2*
>0.8	M	cg17416793	0.010	0.942	0.0127	11	*KCNQ1*
>0.8	M	cg05926314	0.005	0.937	0.0129	7	*PTPRN2*
>0.8	M	cg19764489	0.007	0.906	0.0164	11	*KCNQ1*
>0.8	M	cg09350411	−0.002	0.910	0.0222	7	*PTPRN2*
>0.8	M	cg12001456	0.005	0.922	0.0234	7	*PTPRN2*
>0.8	M	cg08376924	0.005	0.913	0.0244	7	*PTPRN2*
>0.8	M	cg27050114cr	0.008	0.884	0.0264	11	*KCNQ1*
>0.8	M	cg06423822	−0.002	0.865	0.0274	7	*PTPRN2*
>0.8	M	cg03647659	−0.001	0.955	0.0320	11	*B4GALNT4**
>0.8	M	cg24652817	−0.002	0.861	0.0363	7	*PTPRN2*
>0.8	P	cg07599819s	−0.002	0.899	0.0012	19	*ZIM2*
>0.8	P	cg03652257	−0.003	0.818	0.0033	7	*GLI3*

>0.8	P	cg24257495	0.005	0.930	0.0038	8	DLGAP2
>0.8	P	cg03156547	−0.002	0.839	0.0065	14	CDH24*
>0.8	P	cg06450373cr	0.008	0.889	0.0087	14	CDH24*
>0.8	P	cg01983373	−0.002	0.890	0.0178	14	DLK1
>0.8	P	cg18100008	−0.001	0.920	0.0228	7	GLI3
>0.8	P	cg03646329	−0.003	0.822	0.0243	13	RB1
>0.8	P	cg11882053	−0.003	0.827	0.0256	13	RB1

Significant age-associated DMCs and allocated genes (*n* = 22) are shown based on Geneimprint and the Imprintome (see methods). Probe IDs, estimated changes of β-values by age, annotated imprinted genes, and expressed allele are indicated. *Predicted imprinted gene; (EA): expressed alleles: (P) Paternally, (M) Maternally, or (I) Isoform Dependent; crcross-reactive probe; sSNP. Results are ordered by subgroup (SG): unmethylated, UM (<0.2); hemi-methylated, HM ((0.2–0.8)); fully methylated, FM (>0.8); and, sorted by adjusted *p*-value (FDR <0.05). Note: If DMCs were mapped to multiple genes, only the ICR-related (predicted) imprinted gene is reported.

Assessment of a potential role in autism development of genes linked to age-related DMCs in sperm

Based on our results and reports by others, age-related DMCs in sperm were frequently seen at genes involved in autism. However, no earlier studies verified if these DNA methylation marks were located at sites important in epigenetic inheritance. First, we verified which of the seven recurrently reported genes (see paragraph above) are prone to imprinting. Two of the seven genes are known to be imprinted: *DLGAP2* has been reported as a paternally expressed gene, and *SLC22A18AS* has been listed as a maternally expressed gene [18]. The estimated effects by age based on the current study at five CpG sites allocated to *DLGAP2* are illustrated (Figure 8 and Supplementary Tables 4 and 5), directions
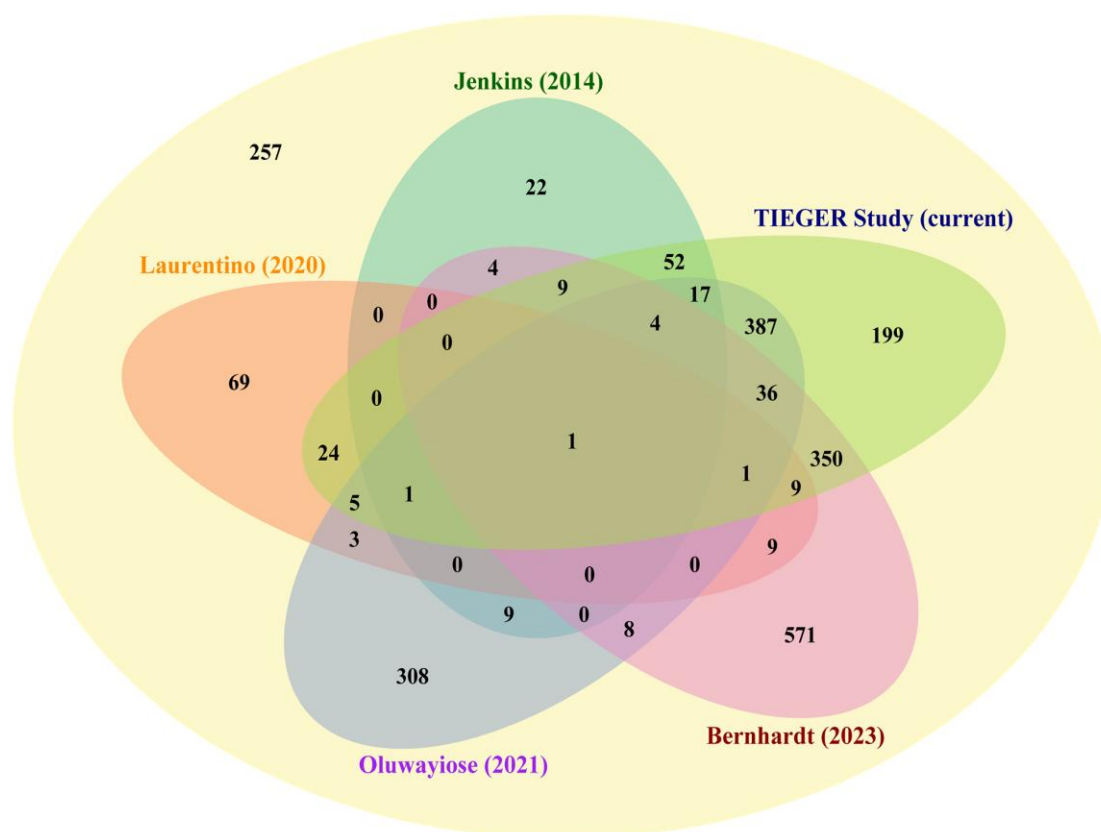
Figure 7. Comparison of our set of age-related DMCs with earlier published datasets. Venn diagram showing the number of genes that have been identified by different authors as being differentially methylated by age in sperm; starting from a recent review by Bernhardt et al. [11], where 2,355 genes have been selected. Next to our results, the following publications have been included in this comparison: Bernhardt et al., Jenkins et al., Laurentino et al., and Oluwayiose et al. [9–12]. Seven genes (center of diagram) were identified by at least four studies (*DLGAP2*, *SLC22A18AS*, *C7orf50*, *UTS2R*, *BEGAIN*, *GRIN1*, and *PCDH15*); two are known imprinted genes (*DLGAP2* and *SLC22A18AS*).

varied by site; while at the three CpG sites allocated to *SLC22A18AS* all were hypomethylated by age (Figure 8 and Supplementary Table 4). This suggests that these genes may play an important role in autism development in children born to fathers of an advanced age. Hence, changes in DNA methylation patterns at these genes could be considered as a potential predictor for autism in the next generation.

When screening the 95 age-related imprinted genes based on the Geneimprint database (Supplementary Table 4), we found 28 genes that have been reported in autism development. When considering cross-reactivity of probes, the gene *C6orf145* was excluded from our final list, resulting in 27 ASD-related imprinted genes (Figure 6). Eleven of these imprinted genes have been linked to a reported ICR (Figure 6, lower right panel);
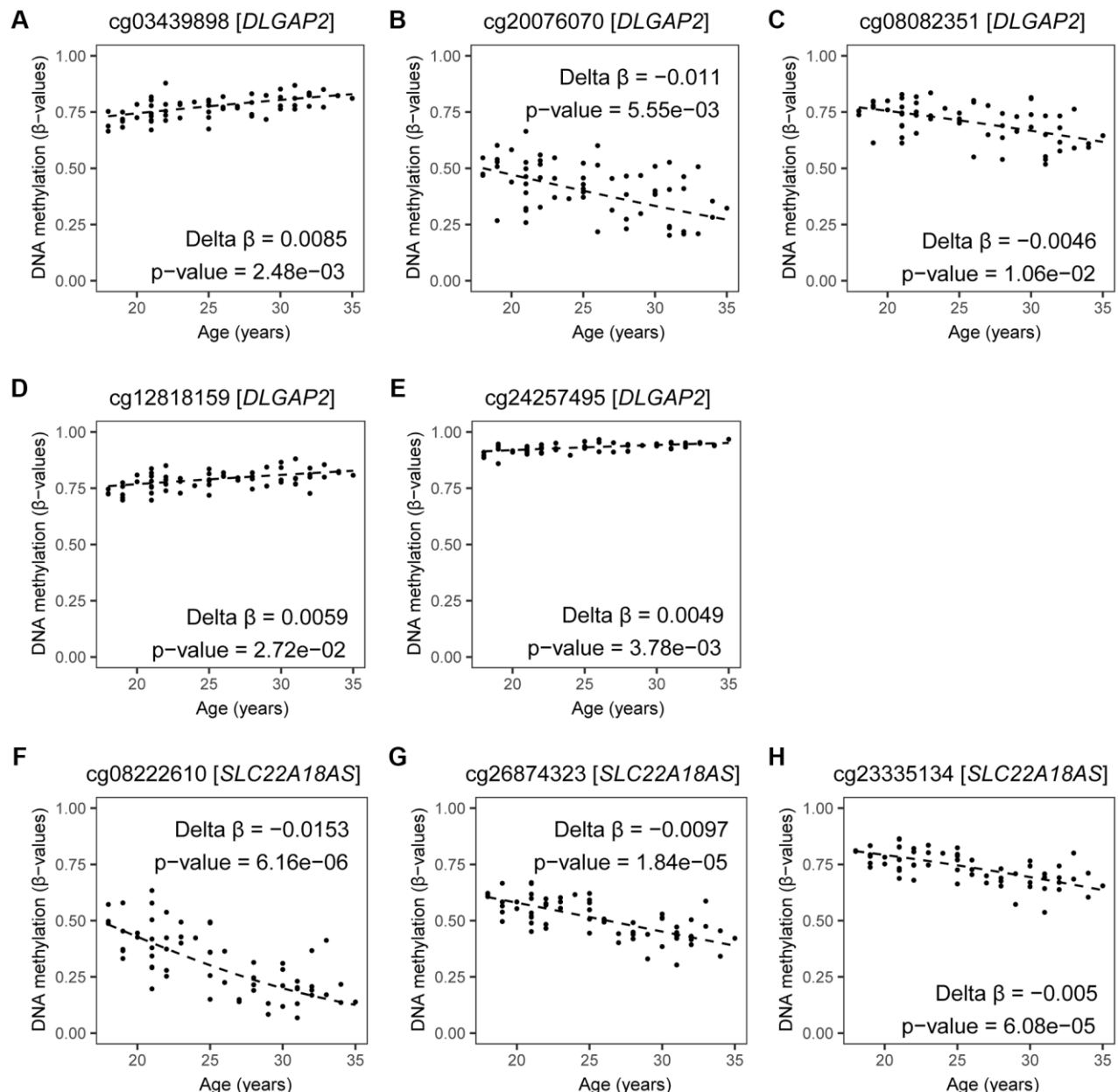
Figure 8. DNA methylation by age at CpGs linked to *DLGAP2* and *SLC22A18AS*. Estimates of DNA methylation in β-values by age for all significant CpG sites annotated to *DLGAP2* (**A–E**) and *SLC22A18AS* (**F–H**) imprinted genes are shown; DNA methylation changes at these genes have been identified by at least three other age-related studies. Fitted regression lines are shown. Regression models included potential confounding factors (BMI and patient status); these were corrected for multiple testing (BH-method). The following CpG sites were annotated to *DLGAP2*: (**A**) cg03439898 (Delta β = 0.0085, *p*-value = 0.0025), (**B**) cg20076070 (Delta β = −0.011, *p*-value = 0.0056), (**C**) cg08082351 (Delta β = −0.0046, *p*-value = 0.7083), (**D**) cg12818159 (Delta β = 0.0059, *p*-value = 0.0272), and (**E**) cg24257495 (Delta β = 0.0049, *p*-value = 0.0038). The following CpG sites were annotated to *SLC22A18AS*: (**F**) cg08222610 (Delta β = −0.0153, *p*-value = 6.158515e-06), (**G**) cg26874323 (Delta β = −0.0097, *p*-value = 1.840627e-05), (**H**) cg23335134 (Delta β = −0.005, *p*-value = 6.076208e-05).

Table 5. ASD-associated (predicted) imprinted genes and scoring.

Gene	Chr	Probe ID	ICR	Co	Multi CpG	Top 90	Ma	Op	Is	Pr	Score of interest
MAGEL2	15	cg10073842	1	0	1	1	1	0	0	1	5
DLGAP2	8	cg03439898	1	1	1	0	0	0	1	0	4
GNAS	20	cg21809160	1	0	1	0	1	0	1	0	4
B4GALNT4	11	cg10798664	1	0	1	0	0	0	0	1	3
GRB10	7	cg27006764	1	0	0	0	0	0	1	1	3
KCNQ1	11	cg03371125	1	0	1	0	0	0	0	1	3
KCNQ1OT1; KCNQ1	11	cg03654058	1	0	1	0	0	0	0	1	3
OTX1	2	cg10487970	0	0	0	0	0	1	1	1	3
PLAGL1	6	cg21113768	1	0	1	0	0	0	0	1	3
PTPRN2	7	cg04937416	1	0	1	0	0	0	1	0	3
PRDM16	1	cg10588310	0	0	1	0	0	1	1	0	3
CDH24	14	cg03156547	1	0	1	0	0	1	0	0	3
SLC26A10	12	cg08177625	0	0	1	0	0	1	0	1	3
ATP10A	15	cg06066676	0	0	1	0	0	0	0	1	2
FBRSL1	12	cg19100996	1	0	0	0	0	1	0	0	2
HOXB3	17	cg10585948	0	0	1	0	0	0	0	1	2
ANO1	11	cg11058904	0	0	0	0	1	0	0	0	1
C6orf145	6	cg18815879[cr]	0	0	0	0	0	0	0	1	1
FOXG1	14	cg18299578	0	0	0	0	0	0	0	1	1
GLIS3	9	cg13804450	0	0	1	0	0	0	0	0	1
LMX1B	9	cg13466694	0	0	1	0	0	0	0	0	1
MAGI2	7	cg00110846	0	0	1	0	0	0	0	0	1
NTM	11	cg12079699	0	0	1	0	0	0	0	0	1
OBSCN	1	cg18477163	0	0	0	0	0	0	1	0	1
UBE3A	15	cg12060334	0	0	1	0	0	0	0	0	1
ADAMTS16	5	cg26892415	0	0	0	0	0	0	0	0	0
DDC	7	cg15001032	0	0	0	0	0	0	0	0	0
RNU5D	5	cg19107296	0	0	0	0	0	0	0	0	0

ASD-associated (predicted) imprinted genes and related probes. Genes were linked to ASD disorders according to previously published databases (see methods). "Score of interest" is given by the following criteria: (ICR) listed by others as being close to an ICR; (Co) confirmed by at least four studies (including the current one); (Multi CpG) age-association was found at more than one CpG of the gene reported; (Top 90) belongs to the most significant top 90 age-related DMCs; (Ma) belongs to the highest magnitude in change by age (Delta-M > 0.1); (Op) opposite direction in methylation change; (Is) located at CpG island; (Pr) located at promoter region. Note: "1" means at least one CpG was found with this characteristic; "0" means none of the DMCs or genes met our criterium. Last column is the sum of eight criteria. [cr]Gene linked to a cross-reactive probe. Genes are listed from highest to lowest scores. For each gene, probe with highest score is shown. Note: one gene (SLC26A10) is reported by the NCBI Gene database as a pseudogene in human.

five of these are included in the SFARI database (see Methods): *MAGEL2* (score 1; high confidence for ASD), *GNAS,* and *DLGAP2* (score 2; strong candidate for ASD), *GRB10* (score 3; suggested evidence for ASD), and *FBRSL1* (unknown score). Based on other databases (including genes found in multiplex ASD families), the following six genes are also listed: *B4GALNT4, KCNQ1, KCNQ1OT1, PLAGL1, CDH24,* and *PTPRN2* (Figure 6) [25, 29]. Finally, these eleven genes form a first set of age-related biomarkers (Figure 6). Based on important functional characteristics of epigenetic inheritance,

current and earlier findings on age-related epigenetic marks in sperm and ASD risk, we ranked our findings by several characteristics (significance, magnitude of change, overlapping findings with earlier reports, multiple age-related CpGs within the same gene, promoter or island content) (Table 5). Except for the *FBRSL1* gene, all 11 genes had a "score of interest" of 3 or higher. We used this numerical value (3) to further select for potential biomarkers within the remaining set of ASD-related imprinted genes without a yet established link to an ICR. Hence, additional genes with

a "score of interest" of at least 3 are: *OTX1*, *PRDM16*, and *SLC26A10*. The latter is listed in the NCBI Gene database as a pseudogene; hence, we do not select for this gene in our final list of potential biomarker genes. Consequently, we found thirteen candidate imprinted genes as potential biomarkers for age-related inheritance of ASD from father to child. These are illustrated in Volcano plots by DNA methylation subgroups (Figure 2), and fitted regression lines are shown for each gene (Figures 8 and 9).
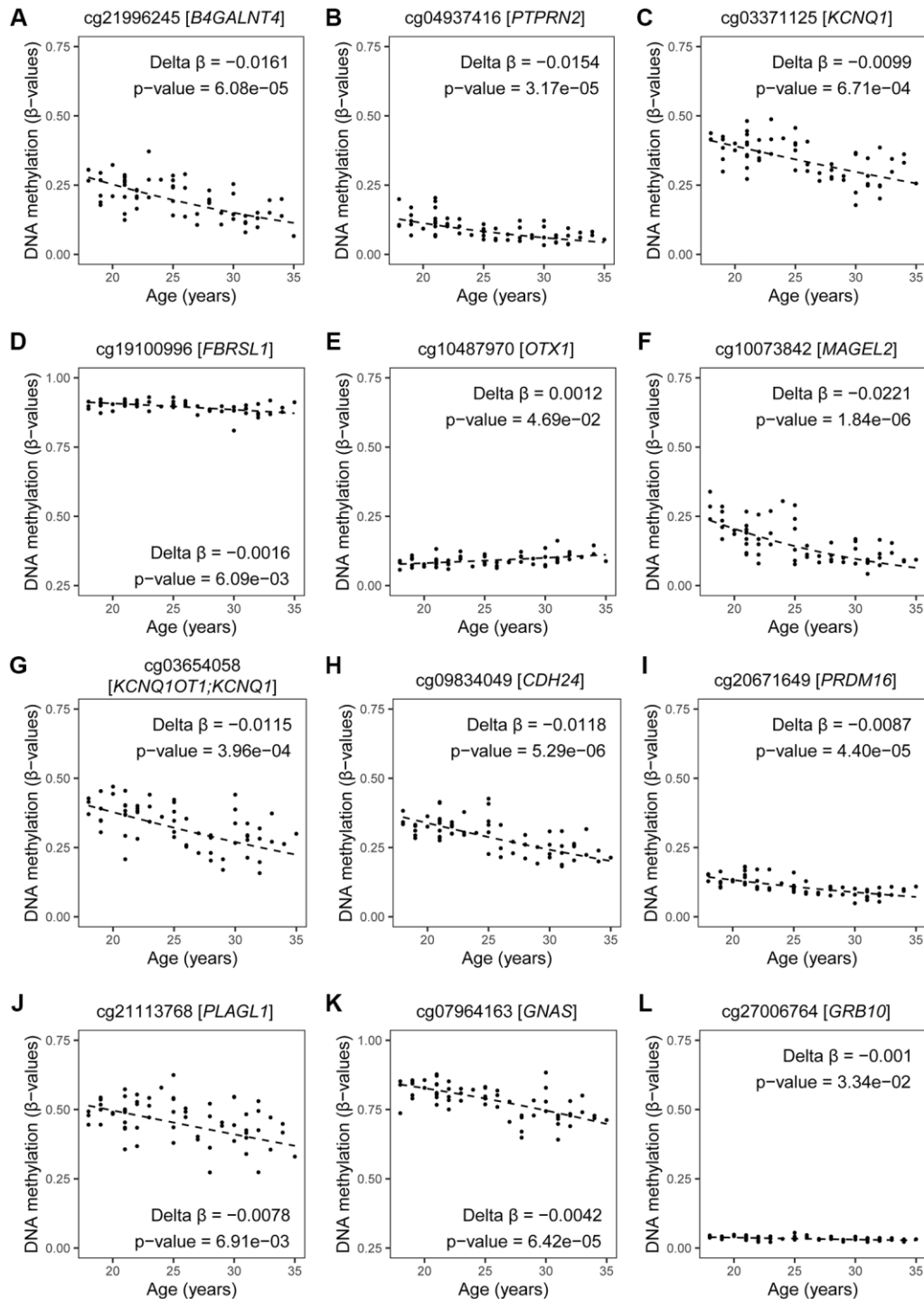
Figure 9. DNA methylation by age at CpGs linked to imprinted genes and related to ASD. Estimates of DNA methylation in β-values by age of CpG sites mapped to our selected list of ASD-related imprinted genes (with exception of CpGs mapped to *DLGAP2*, which are shown in Figure 8). Fitted regression lines are shown. Our regression models included potential confounding factors (BMI and patient status), and we corrected for multiple testing (BH-method). (**A**–**E**) maternally expressed genes; (**F**–**J**) paternally expressed genes; (**K**, **L**) isoform dependent transcribed gene. If more than one CpG site per gene was significant, the one with the lowest *p*-value is shown.

DISCUSSION

This study revealed an overall decrease in DNA methylation at 14,622 CpG sites in sperm by age. In our unique search for age-related signatures in sperm featuring epigenetic inheritance from father to offspring we identified 95 imprinted genes. Only two of these identified genes have been found in similar studies: *DLGAP2* and *SLC22A18AS* [9–12, 30]. *SLC22A18AS*, encoding a lncRNA, is a paternally imprinted gene [18]. While this gene is not listed in SFARI, two independent study cohorts in children with neurodevelopmental disorders showed aberrancies in expression levels or epimutations at regulatory sites of *SLC22A18AS* [22, 23]. Our data show a decrease in DNA methylation percentage by age at all three CpG sites of *SLC22A18AS* (Figure 8). While not selected as a potential biomarker for ASD-risk in the current study, we believe more research is needed to identify this gene's potential role in inheritance of ASD. The *DLGAP2* gene has been listed as a strong ASD candidate in SFARI and its imprinted domain has been validated by others [17, 18]. *DLGAP2* is a paternally expressed gene, but we measured an average DNA methylation of 78% in sperm. Higher methylation percentages than anticipated have also been measured by others [31, 32]. We found that age-related DMCs near *DLGAP2* are either hypermethylated (at islands, and open sea) or hypomethylated (at shores), which suggests that expression of this gene is malleable by environmental stressors. However, not all probes have been mapped to an ICR, and more research on the role of methylation in these regions is needed. Hypomethylation in sperm has also been reported in sperm cannabis users [31]. A role for intergenerational inheritance of these altered DNA methylation patterns at *DLGAP2* by cannabis exposure, has been demonstrated in a rat model [32]. Other studies in rat indicate that the *DLGAP2* methylation status can also be modified by conditions similar to post-traumatic stress in human [33]. According to a review by Rasmussen et al., *DLGAP2* belongs to a family with a direct link to a variety of neurological and psychological disorders. Expression profiles are mainly found in testis and brain, and aberrancies in these expression patterns have been associated with schizophrenia and autism spectrum disorders [34, 35]. This confirms that this gene, important in early development of the brain, is highly susceptible to environmental influences.

After mapping our findings to a list of earlier reported ICRs, 22 genes were retrieved of which 11 have been related to ASD (Figure 6). Notably, 11 out of the 22 genes were also reported by a separate study performed by Court et al., where differentially methylated regions were verified in a range of human tissues (including sperm, embryos and somatic tissues) [36]. Five of these genes verified by Court et al. -providing evidence for inheritable epigenetic states- have been linked to (neuro)developmental disorders: *MAGEL2, KCNQ1OT1, PLAGL1, GNAS, GRB10*. All are included in our selected biomarkers, using an independent ranking system (Figure 6 and Table 5). Among all identified imprinted genes (*n* = 95), the gene *MAGEL2* is matching multiple criteria, such as high magnitude in change by age, being within the top-90, at least one DMC at a promotor region, and having multiple DMCs of highest degree of significance; for instance, at probe cg10073842 we measured a Delta β of −0.022 (or −2.2% per year) (*p* = 1.84e-06). *MAGEL2* is listed in the SFARI database and some DMCs could be allocated to an ICR (Supplementary Tables 5 and 7). The latter was (re)confirmed through data from an independent investigation where Illumina probes were mapped to candidate ICRs [37]. Aberrancies or variations in methylation at these sites have been linked to various diseases [36, 37]. In addition to its role in autism, it has also been reported as a contributor to Prader-Willi, Angelman, and Schaaf-Yang syndromes [38–41]. Another interesting finding is the paternally imprinted gene *PTPRN2*. DMCs related to this gene showed an age-related decrease in methylation; including one DMC within the FM subgroup that was located at an island region. *PTPRN2* plays an important role in vesicle-mediated secretory processes, such as the accumulation of neurotransmitters norepinephrine, dopamine and serotonin in the brain [42]. Copy number variations of this gene have been found to be associated with ASD in a Brazilian cohort [43]. DNA methylation changes have been investigated in other populations, as well as by other exposures than ageing. For instance, hypomethylation at *PTPRN2* was found in cord blood of children exposure to organochlorine compounds (DDE) in families at the Faroe Islands [44]. However, hypomethylation in sperm could not be linked to DDE exposure in the same Faroese population. Instead, the authors found both directions in DNA methylation changes; significant hypo and hypermethylation at multiple DMCs of *PTPRN2*. They concluded that this was suggestive for a regulatory role and potential inheritance of ASD [45]. Similarly, we found nine age-related DMCs at the paternally imprinted gene *B4GALNT4*. These DMCs are either located in the body or the promoter region; we measured an overall decrease in DNA methylation. Liu et al. showed in a rat model that *B4GALNT4* plays a role in brain function and development, including the development of ASD [46]. We further found age-related DNA methylation alterations at promotor regions of several other genes. In the case of *KCNQ1*, this gene is member of the *KCNQ* voltage-gated potassium channel family,

implicated in the pathogenesis of several forms of autism spectrum disorders [47, 48]. Hypomethylation was also measured at *PRDM16* and *SLC26A10*; however, as the latter is reported as a pseudogene in human with unknown function we did not select *SLC26A10*. Aberrant *PRDM16* expression has been linked to cardiovascular diseases and migraine in human [49]. In mice, Prdm16 is expressed in brain throughout middle and late stages of cortical neurogenesis, as well as during early post-natal development. This gene also plays a role in gene regulation through epigenetic mechanisms, and it co-operates with Pax6 [50]. Interestingly, heterozygous spontaneous mutant mice with a mutation in Pax6 show phenotypes corresponding to autism in human; behavioral abnormalities in these mice are more prominent in offspring from aged fathers compared to young fathers [51, 52]. Next, we found age-related alterations at the promoter regions of the following genes: *PLAGL1*, *GRB10*, *KCNQ1OT1* and *OTX1*. DNA methylation differences have been found at imprinted loci of *PLAGL1* in patients with autism *versus* controls [38]. The same study also indicated epivariations at imprinted sites of *GNAS* in patients with autism. Note, according to Carreras-Gallo et al.'s report, the identified age-related DMC of *GNAS* in the current study is part of a reported ICRs [37]. Next, based on animal studies, *GRB10* is also a contributor to ASD pathogenesis. Its transcript forms a complex with the autism-associated protein GIGYF1 to regulate the IGF-1R/ERK signaling pathway, important in brain development [53]. In a study of postmortem temporal cortex samples from ASD subjects, histone modifications could be detected at *GRB10* [54]. The authors suggested that the measured *GRB10* deacetylation could represent an epigenetic mechanism in idiopathic autism linked to pathways that are also affected by rare genetic variants causing syndromic ASD. This supports earlier insights that rare coding variants alone cannot account for all ASD cases, but epigenetic alterations (by environmental conditions) may contribute to many ASD cases [55]. *GRB10* is listed as an isoform dependent imprinted gene because of its complex imprinted expression pattern. It is expressed maternally in most tissues, but it is paternally expressed (or maternally methylated) in the early embryo and in the brain [56]. Deletion of the paternal allele in mice resulted in offspring with an aberrant social behavior, such as social dominance [57]. Alterations at imprint regulatory regions of *KCNQ1OT1* have been linked to autism and Beckwith-Wiedemann and Silver–Russell syndromes [58]. A DNA methylation alteration was found at *KCNQ1OT1* in patients with autism or intellectual disability [59]. An overlap in disease outcomes was also found for *FBRSL1*. This gene is included in the SFARI and Homs

et al.'s databases, but it has also been linked to a disability syndrome, including heart defects and intellectual disability. Evidence suggests this gene orchestrates gene expression in neurogenesis and in the development of autism. An assessment of DNA methylation patterns for this gene has been considered as a future diagnostic tool for ASD [60, 61]. The gene *OTX1* encodes a transcription factor important in brain and sensory organ development [62, 63]. It is also known as a tumor promotor, suggesting its broader involvement in developmental and proliferative processes. In the context of ASD, *OTX1* has emerged as a strong candidate gene (SFARI score 2, meaning that the association has been shown earlier with high confidence) [21]. *CDH24* (cadherin 24) is a paternally imprinted gene. In general, this gene mediates cell adhesion (particularly in the brain) and is involved in the formation of intracellular signaling pathways. But early genome wide association studies revealed some autism-associated cadherins [64].

In summary, imprinted genes selected in the current age-association study as potential epigenetic biomarkers for ASD in offspring are: *OTX1, PRDM16, PTPRN2, B4GALNT4, KCNQ1, KCNQ1OT1, DLGAP2, PLAGL1, GNAS, GRB10, MAGEL2, CDH24*, and *FBRSL1*; these are highlighted in Figure 2 and fitted regression lines are illustrated for at least one probe per gene (Figures 8 and 9). We measured small DNA methylation differences by age; ranging from 0.07% (*B4GALNT4*) to 2.75% (*PTPRN2*) per year. However, if fertilization were to occur with a sperm cell that carries one (or multiple) aberrantly methylated sites at one (or several) of these genes, this developing child may be at risk to have ASD. For instance, at *OTX1* -a gene with an average change per year- our linear estimates shows that a ten-year increase in paternal age could result in 1.2% more sperm cells being hypermethylated. On an individual level this can be translated to 1 200 000 aberrant sperm cells in an ejaculate of 100 million cells, associated with paternal age. While the chance is low that the affected sperm will result in a pregnancy, on a population level 1.2 % more children with an increased risk to develop ASD is an important contribution to public health. However, this is still hypothetical as other factors play a role in ASD development as well, and current knowledge is too limited to understand the effects of one or more affected CpGs at one gene. Even less is known about the combined effect of paternal age on multiple imprinted genes and possible interactions with other determinants.

We detected limited age-related DMCs at CpG islands; only 4.5%, while the expected island content within the array is about 31% [46]. We hypothesize that CpG islands are more protected against changes by ageing or

related factors, because of their importance in gene regulation. Yet, the limited number of significant DNA methylation alterations we have measured at these sites in sperm may be important in evolution and inheritance of specific phenotypes. Notably, our selection of biomarkers was based on several stringent criteria, including ASD-related databases. It is possible that other (unknown) autism-associated genes may have been missed. Consequently, our list of biomarkers is not final and should not be taken strictly. It is a first step towards a more refined and validated set of genes sensitive to epigenetic changes by age. As our sample size is small, other populations of men (on a larger scale and using a broader age-range) should be tested before implementing our findings in clinic. Another reason why we acknowledge its prematurity, is the integration of data from lists of ICRs and imprinted genes that are extensive but not yet fully developed [17]. It has been estimated that the human genome comprises between 300 and 1,000 imprinted genes [65]. Theoretically, this means that male sperm harbors hundreds of important regulatory sites yet to be discovered. As indicated, because of variation in populations, men living in different geographical areas or men with different cultural backgrounds may have different patterns of DNA methylation as they age. For example, in a recent study in the UK, differences were seen in DNA methylation patterns by biological ageing and by housing circumstances [66]. We do not exclude that the use of age is *a proxy* for a mixture of environmental exposures that coincide with ageing. A remarkable finding by Wang et al. shows that depending on the method used to select a sperm cell in fertility treatment procedures, children may have higher risk for autism [67]. Given the need for more research on the exposure impact on sperm epigenetics, age remains an important variable to include in prediction modeling of acquired diseases. A potential weakness of our study is its limited age-range. The current study does not provide an answer to the potential ASD-risk associated with fathering a child at older age (e.g., 40 years old or older). Nevertheless, we detected significant DNA methylation alterations within a range of 18 years (from 18 till 35 years old). This covers the age when most men father their first child. Literature raises the problem of the biological clock of the male germ line in ageing men and related effects on the increasing numbers of children with ASD or other neurodevelopmental disorders. It has been suggested that age of the father gradually increases the risk for autism in offspring [68]. In a meta-analysis it was calculated that every 10 years increase in age results in 21% increase in risk for ASD in offspring (OR: 1.21; 95% CI: 1.18–1.24) [69].

Finally, we do not exclude a potential higher risk for other disorders (than autism) in children from men of advanced age. As hypothesized by other researchers, such as Levine et al., epigenetic biomarkers of ageing could be used to capture an overall risk for future diseases [70]. Because of recurrent findings in literature on autism-related outcomes, we focused on ASD databases in our final analysis, and we did not explore our list of 14,622 DMCs on other potential outcomes. Hence, although our results point to a (first) set of candidate biomarkers for one disorder, ASD, we believe that the sperm epigenome contains more signatures to be explored. A strength of studying imprinting in sperm -and potential offspring outcomes- is the fact that these signatures may persist in somatic cells of the next generation. This is particularly important in research on brain related disorders. Studying causes of autism in humans is complicated, especially because key tissues (brain) are not available for clinical testing. Hence, as indicated earlier by the research group of Craig, researchers can either study postmortem samples, test mechanisms associated with the disorder, or identify and evaluate biomarkers for risk or prognosis of autism [71]. However, our age-associated epigenetic marks in sperm should still be present in somatic cells of offspring. While our study design was limited to sperm analyses only, and inheritance could not be tested, including these biomarkers in future research on children (considering paternal age in children with *versus* without ASD) would facilitate the creation of a causation model for autism, with no (or limited) need of postmortem tissues. In this way, if explored and validated further on a larger scale, it would be possible to estimate the effect (or risk) for ASD in children by preconceptional conditions of the father. Moreover, although risk at individual level may be small, the effect on a population level is an important result to consider, especially in the context of future policy re-commendations. For instance, educational programs would benefit from this knowledge. Specifically, awareness could be increased in young men about the risks for having a child with autism, if they postpone fathership; and in clinic, appropriate preconceptional counseling of couples in which the man is of advanced age could be implemented. Our results may further contribute to new insights on subtypes of autism, reported by the group of Troyanskaya [72, 73]. In their most recent publication, they were able to subdivide autism by specific genetically underlying mechanisms. In brief, some autism subtypes may be caused by *de novo* mutations in DNA at coding regions, while other subtypes may have their origin in changes at non-coding regions or sites with allele-specific regulatory activities. This aligns with our finding and supports the importance of a -yet underestimated role- of epi-genetics in the onset and heterogeneity of neuro-developmental diseases.

MATERIALS AND METHODS

Study participants and sample collection

Male volunteers were recruited as part of The Influence of the Environment on Gametic Epigenetic Reprogramming (TIEGER) study. The cross-sectional design, sample and data collection, and selection criteria of this North Carolina, USA-based study have been published previously; a flow chart of the current study population of 63 men is included in Supplementary materials (Supplementary Figure 1) [15]. In brief, eligibility criteria at recruitment included: non-smoking, no personal history of cancer, no vasectomy or other procedures that may cause infertility, age range of 18–35, and Caucasian; this was applied to keep our population relatively homogenous in this small study sample. In the clinic, semen was analyzed for standard clinical parameters after liquefaction, no later than 60 minutes from collection. These parameters included volume, pH, viscosity, liquification time, presence of white blood cells, concentration, and motility. The World Health Organization's (WHO) Laboratory Manual for the Examination and Processing of Human Semen 5th edition was referenced for normal values [74]. After completion of the clinical sperm analyses, the samples were subjected to two-step ISolate-gradient centrifugation (Irvine Scientific) to select a motile population enriched in normal morphology. This colloidal silica gradient, consisting of a 90% lower layer and 50% upper layer, was prepared by sequentially adding 1.5 ml of each layer to a 15 ml polystyrene conical tube. The sperm sample was pipetted on top of the upper layer and centrifuged at 200× g for 15 minutes. The gradient solution was removed, and pelleted sperm were stored at −80°C for subsequent DNA methylation analysis [15].

DNA methylation measurements in sperm, data preprocessing, and validation

Sperm genomic DNA was extracted and prepared for Illumina HumanMethylation450 BeadChip as previously described [15]. Methylation levels range from 0% to 100%; where 0% means no sperm cell is methylated at a specific CpG site, and 100% means that all sperm cells in a sample are methylated. Then, we scaled our data in an interval of (0, 1). Hence, DNA methylation values are represented as proportions ranging from 0 (if "unmethylated") to 1 (if "methylated"). Our preprocessing scenario included sample quality control, filtering of low intensities probes, Beta Mixture Quantile normalization (BMIQ), and logit transformation of default β-values into M-values; $M = \log(\beta/1-\beta)$ (Supplementary Figure 2) [75–77]. Probes with low quality signal (detection p-value

greater than 0.01 in at least one sample) were filtered out. Our final dataset represents outcomes of 482,287 probes (out of 485,512), referred as "450K". Validation tests were performed through data obtained after bisulfite pyrosequencing. These protocols, including PCR conditions and primer sequences were reported previously [15, 78, 79]. In brief, using the Pyrosequencing WorkStation, the single strand was isolated and then underwent pyrosequencing using a PyroMark Q96 MD pyrosequencing instrument (Qiagen). The region upstream from *IGF2* exon 3, including three CpG dinucleotides (chr 11p15.5), was tested, as well as the DMR for *H19*, including four CpG sites (chr 11 p15.5). Assay validation data and the results of sensitivity tests for pyrosequencing have been published [15, 78, 79]. At *IGF2*, an increase of 1 year of age showed a decrease in DNA methylation (β-value = −0.10; SE = 0.04; p = 0.02). At *H19*, an increase of 1 year of age corresponded to a significant increase in DNA methylation (β-value = +0.14; SE = 0.05; p = 0.009). These results are similar to our findings in the current approach, when using the 450K array (e.g., all DMCs at *IGF2* are negatively associated with age with a mean delta β-value of −0.01; at *H19*, one age-associated DMC was found, with a delta β-value of +0.0054).

Statistical analysis

Two main analyses were assessed to better understand how age may affect sperm DNA methylation. We analyzed 1) age *versus* global DNA methylation, and 2) age *versus* DNA methylation at each CpG site of the 450K. A workflow illustrates our approach (Figure 1). The mean β-value was measured at all CpG sites of the 450K. Using this mean we calculated the global DNA methylation in sperm for each subject. Pearson correlation coefficient was used to measure the correlation between global DNA methylation and age, and Mann-Whitney U-test was used to evaluate the difference in global DNA methylation for different subgroups of men. We used the mean β-value (in all subjects, for each CpG site) to classify each CpG site by its potential biological relevance, as suggested in literature [80]. Consequently, a CpG site with a mean β-value <0.20 was defined as "unmethylated" (UM), as "hemi-methylated" (HM) if 0.20≤ mean β-value ≤0.80, and as "fully methylated" (FM) if mean β-value >0.80. To estimate age-related associations, outcomes of the 450K -expressed as M-values- were included as dependent variables in a linear regression model. Age, BMI and patient status were added as independent variables. We used moderated t-statistic and corrected for multiple testing using the Benjamini-Hochberg (BH) method. Significant threshold was defined as FDR <0.05. Notably, the use of M-values in regression

modeling -while less intuitive than β-values- provides a better estimate of age-related effects on DNA methylation as data of M-values are approximately homoscedastic. We back-transformed delta M-values to delta β-values, using Kruppa et al.'s "intercept method"; as follows:

$$\Delta\beta = \frac{2^{(M_0 + \text{delta M})}}{1 + 2^{(M_0 + \text{delta M})}} - \frac{2^{M_0}}{1 + 2^{M_0}};$$

where M0 indicates the baseline DNA methylation (intercept) [81, 82]. Volcano plots and fitted regression lines (based on estimated β-values) were used to graphically represent the results of our linear regression models (Figures 2, 8 and 9). Sensitivity tests were performed, removing outliers. We also repeated our analyses by excluding patients, and by adding additional co-factors in our models, such as sperm quality. These tests did not result in differences in our final results. In a post-hoc analysis we also repeated our analyses excluding 73,735 CpGs reported at SNPs [19]. We investigated cross-reactivity of probes using DMRcate package, based on Chen et al.'s list [20, 83]. SNPs and cross-reactive probes are indicated in our tables with "s" and "cr", respectively. To investigate the chromosomal distribution of all age-associated DMCs, a Miami plot was generated, showing the direction of DNA methylation change by age per chromosome (hypermethylated CpG sites, Delta Beta >0 vs. hypomethylated CpG sites, Delta Beta <0). We used the Illumina annotation manifest to explore DMCs by island content, their functional genomic allocation and associated genes [84]. To investigate the potential role of differentially methylated CpG sites (DMCs) in imprinting, we identified genes corresponding to age-associated sites that were previously listed in the Geneimprint database, and as nearest transcript of a candidate imprint control region (ICR) [17]. Moreover, we verified if DMCs allocated to those genes correspond to probes identified by Carreras Gallo et al. as part of related ICRs [37]. In order to compare our results with earlier reports, we used datasets from Bernhardt et al. [11], Jenkins et al. [9], Laurentino et al. [10], and Oluwayiose et al. [12]. Finally, we used the following databases to elaborate our study on ASD: The Simons Foundation Autism Research Initiative (SFARI) database (which includes a gene scoring system to reflect the strength of evidence for its genetic involvement in ASD) and few other studies reporting ASD-related gene databases; including also ASD risk genes based on families with multiple affected children [29, 30, 85]. We also used the following open access resources: the Human Protein Atlas and the NCBI gene database [42, 86]. We ranked our findings by the following characteristics (Table 5): Top 90 most significant (yes/no), delta-M >0.1 (yes/no), opposite direction (yes/no), located at CpG island (yes/no), located at promoter region (yes/no), and listed by at least three other studies (yes/no). The sum was called "score of interest" (Table 5).

Our analysis was conducted in R (version 4.2.2.). Data preprocessing and annotation were conducted using Bioconductor packages, including missMethyl for GO terms enrichment analysis [87–90]. Upon request, full lists of data will be made available upon documentation of approval by a recognized Institutional Review Board (IRB).

Abbreviations

ASD: Autism Spectrum Disorders; BH: Benjamini-Hochberg; BMIQ: Beta Mixture Quantile; DMC: Differentially Methylation CpG Site; FM: Fully Methylated; GO: Gene Ontology; HM: Hemi-methylated; ICR: Imprint Control Region; SFARI: Simons Foundation Autism Reference Initiative; SNP: Single Nucleotide Polymorphisms; TIEGER: The Influence of the Environment on Gametic Epigenetic Reprogramming; UM: Unmethylated.

AUTHOR CONTRIBUTIONS

EC, JD and AS developed the research question and wrote the manuscript. JD was equally involved in the interpretation and discussion of the research results. MB performed the data cleaning and initiated data analysis. EC performed the major statistical analysis. EC designed and performed post-hoc comparisons. CD contributed to the interpretation of the data and results on outcomes related to psychiatric disorders. AS and CH developed the epidemiological study design. SM developed the experimental study design and oversaw DNA methylation data generation. AS initiated the overall study hypothesis of the TIEGER study and supervised the overall interpretation of the data analysis. RLJ, CH, SM, PA and TP edited the manuscript. All authors have critically reviewed the manuscript and have given final approval for publication.

CONFLICTS OF INTEREST

We declare no competing interests. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

ETHICAL STATEMENT AND CONSENT

The TIEGER Study was performed with approval of the Duke University Institutional Review Board (reference number: Pro00036645). Informed consent was obtained from all participants.

FUNDING

REFERENCES

1. Van Opstal J, Fieuws S, Spiessens C, Soubry A. Male age interferes with embryo growth in IVF treatment. Hum Reprod. 2021; 36:107–15.
   https://doi.org/10.1093/humrep/deaa256
   PMID:33164068

2. Katz-Jaffe MG, Parks J, McCallie B, Schoolcraft WB. Aging sperm negatively impacts in vivo and in vitro reproduction: a longitudinal murine study. Fertil Steril. 2013; 100:262–8.e1-2.
   https://doi.org/10.1016/j.fertnstert.2013.03.021
   PMID:23579004

3. Milekic MH, Xin Y, O'Donnell A, Kumar KK, Bradley-Moore M, Malaspina D, Moore H, Brunner D, Ge Y, Edwards J, Paul S, Haghighi FG, Gingrich JA. Age-related sperm DNA methylation changes are transmitted to offspring and associated with abnormal behavior and dysregulated gene expression. Mol Psychiatry. 2015; 20:995–1001.
   https://doi.org/10.1038/mp.2014.84
   PMID:25092244

4. Pan N, Lin LZ, Wang X, Guo CH, Jing J, Li XH. Association between paternal age at childbirth and autism spectrum disorder in offspring. Zhongguo Dang Dai Er Ke Za Zhi. 2022; 24:863–8.
   https://doi.org/10.7499/j.issn.1008-8830.2203146
   PMID:36036123

5. Miller B, Messias E, Miettunen J, Alaräisänen A, Järvelin MR, Koponen H, Räsänen P, Isohanni M, Kirkpatrick B. Meta-analysis of paternal age and schizophrenia risk in male versus female offspring. Schizophr Bull. 2011; 37:1039–47.
   https://doi.org/10.1093/schbul/sbq011
   PMID:20185538

6. Grosvenor LP, Croen LA, Lynch FL, Marafino BJ, Maye M, Penfold RB, Simon GE, Ames JL. Autism Diagnosis Among US Children and Adults, 2011-2022. JAMA Netw Open. 2024; 7:e2442218.
   https://doi.org/10.1001/jamanetworkopen.2024.42218
   PMID:39476234

7. Ha AS, Scott M, Zhang CA, Li S, Langroudi AP, Glover F, Basran S, Del Giudice F, Shaw GM, Eisenberg ML. Sociodemographic Trends and Perinatal Outcomes in Fathers 50 Years and Older. JAMA Netw Open. 2024; 7:e2425269.
   https://doi.org/10.1001/jamanetworkopen.2024.25269
   PMID:39088214

8. Vervoort I, Delger C, Soubry A. A multifactorial model for the etiology of neuropsychiatric disorders: the role of advanced paternal age. Pediatr Res. 2022; 91:757–70.
   https://doi.org/10.1038/s41390-021-01435-4
   PMID:33674740

9. Jenkins TG, Aston KI, Pflueger C, Cairns BR, Carrell DT. Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. PLoS Genet. 2014; 10:e1004458.
   https://doi.org/10.1371/journal.pgen.1004458
   PMID:25010591

10. Laurentino S, Cremers JF, Horsthemke B, Tüttelmann F, Czeloth K, Zitzmann M, Pohl E, Rahmann S, Schröder C, Berres S, Redmann K, Krallmann C, Schlatt S, et al. A germ cell-specific ageing pattern in otherwise healthy men. Aging Cell. 2020; 19:e13242.
    https://doi.org/10.1111/acel.13242
    PMID:32951333

11. Bernhardt L, Dittrich M, Prell A, Potabattula R, Drummer C, Behr R, Hahn T, Schorsch M, Müller T, Haaf T. Age-related methylation changes in the human sperm epigenome. Aging (Albany NY). 2023; 15:1257–78.
    https://doi.org/10.18632/aging.204546
    PMID:36849136

12. Oluwayiose OA, Wu H, Saddiki H, Whitcomb BW, Balzer LB, Brandon N, Suvorov A, Tayyab R, Sites CK, Hill L, Marcho C, Pilsner JR. Sperm DNA methylation mediates the association of male age on reproductive outcomes among couples undergoing infertility treatment. Sci Rep. 2021; 11:3216.

https://doi.org/10.1038/s41598-020-80857-2
PMID:33547328

13. Jenkins TG, Aston KI, Cairns B, Smith A, Carrell DT. Paternal germ line aging: DNA methylation age prediction from human sperm. BMC Genomics. 2018; 19:763.
https://doi.org/10.1186/s12864-018-5153-4
PMID:30348084

14. Jenkins TG, James ER, Alonso DF, Hoidal JR, Murphy PJ, Hotaling JM, Cairns BR, Carrell DT, Aston KI. Cigarette smoking significantly alters sperm DNA methylation patterns. Andrology. 2017; 5:1089–99.
https://doi.org/10.1111/andr.12416
PMID:28950428

15. Soubry A, Guo L, Huang Z, Hoyo C, Romanus S, Price T, Murphy SK. Obesity-related DNA methylation at imprinted genes in human sperm: Results from the TIEGER study. Clin Epigenetics. 2016; 8:51.
https://doi.org/10.1186/s13148-016-0217-2
PMID:27158277

16. Shnorhavorian M, Schwartz SM, Stansfeld B, Sadler-Riggleman I, Beck D, Skinner MK. Differential DNA Methylation Regions in Adult Human Sperm following Adolescent Chemotherapy: Potential for Epigenetic Inheritance. PLoS One. 2017; 12:e0170085.
https://doi.org/10.1371/journal.pone.0170085
PMID:28146567

17. Jima DD, Skaar DA, Planchart A, Motsinger-Reif A, Cevik SE, Park SS, Cowley M, Wright F, House J, Liu A, Jirtle RL, Hoyo C. Genomic map of candidate human imprint control regions: the imprintome. Epigenetics. 2022; 17:1920–43.
https://doi.org/10.1080/15592294.2022.2091815
PMID:35786392

18. Jirtle R. Imprinted Genes by Species. 2025.

19. Naeem H, Wong NC, Chatterton Z, Hong MK, Pedersen JS, Corcoran NM, Hovens CM, Macintyre G. Reducing the risk of false discovery enabling identification of biologically significant genome-wide methylation status using the HumanMethylation450 array. BMC Genomics. 2014; 15:51.
https://doi.org/10.1186/1471-2164-15-51
PMID:24447442

20. Peters TJ, Buckley MJ, Chen Y, Smyth GK, Goodnow CC, Clark SJ. Calling differentially methylated regions from whole genome bisulphite sequencing with DMRcate. Nucleic Acids Res. 2021; 49:e109.
https://doi.org/10.1093/nar/gkab637
PMID:34320181

21. Abrahams BS, Arking DE, Campbell DB, Mefford HC, Morrow EM, Weiss LA, Menashe I, Wadkins T, Banerjee-Basu S, Packer A. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol Autism. 2013; 4:36.
https://doi.org/10.1186/2040-2392-4-36
PMID:24090431

22. Kuwano Y, Kamio Y, Kawai T, Katsuura S, Inada N, Takaki A, Rokutan K. Autism-associated gene expression in peripheral leucocytes commonly observed between subjects with autism and healthy women having autistic children. PLoS One. 2011; 6:e24723.
https://doi.org/10.1371/journal.pone.0024723
PMID:21935445

23. Iourov IY, Vorsanova SG, Korostelev SA, Zelenova MA, Yurov YB. Long contiguous stretches of homozygosity spanning shortly the imprinted loci are associated with intellectual disability, autism and/or epilepsy. Mol Cytogenet. 2015; 8:77.
https://doi.org/10.1186/s13039-015-0182-z
PMID:26478745

24. Potabattula R, Prell A, Dittrich M, Nava C, Depienne C, Bejaoui Y, El Hajj N, Hahn T, Schorsch M, Haaf T. Effects of paternal and chronological age on *BEGAIN* methylation and its possible role in autism. Aging (Albany NY). 2023; 15:12763–79.
https://doi.org/10.18632/aging.205275
PMID:38019471

25. Ruzzo EK, Pérez-Cano L, Jung JY, Wang LK, Kashef-Haghighi D, Hartl C, Singh C, Xu J, Hoekstra JN, Leventhal O, Leppä VM, Gandal MJ, Paskov K, et al. Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. Cell. 2019; 178:850–66.e26.
https://doi.org/10.1016/j.cell.2019.07.015
PMID:31398340

26. Bajaj V, Markandaya M, Krishna L, Kumar A. Paternal imprinting of the SLC22A1LS gene located in the human chromosome segment 11p15.5. BMC Genet. 2004; 5:13.
https://doi.org/10.1186/1471-2156-5-13
PMID:15175115

27. Ugur C, Sertcelik M, Uneri O, Senses Dinc G, Sekmen E, Solmaz E. Evaluation of serum urotensin-II levels of children with ADHD and autism spectrum disorder. Anatolian Journal of Psychiatry. 2018; 18:7.

28. Valentino F, Bruno LP, Doddato G, Giliberti A, Tita R, Resciniti S, Fallerini C, Bruttini M, Lo Rizzo C, Mencarelli MA, Mari F, Pinto AM, Fava F, et al. Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations. Brain Sci. 2021; 11:936.
https://doi.org/10.3390/brainsci11070936
PMID:34356170

29. Homs A, Codina-Solà M, Rodríguez-Santiago B,

Villanueva CM, Monk D, Cuscó I, Pérez-Jurado LA. Genetic and epigenetic methylation defects and implication of the ERMN gene in autism spectrum disorders. Transl Psychiatry. 2016; 6:e855.
https://doi.org/10.1038/tp.2016.120
PMID:27404287

30. Jenkins TG, Aston KI, Cairns BR, Carrell DT. Paternal aging and associated intraindividual alterations of global sperm 5-methylcytosine and 5-hydroxymethylcytosine levels. Fertil Steril. 2013; 100:945–51.
https://doi.org/10.1016/j.fertnstert.2013.05.039
PMID:23809503

31. Murphy SK, Itchon-Ramos N, Visco Z, Huang Z, Grenier C, Schrott R, Acharya K, Boudreau MH, Price TM, Raburn DJ, Corcoran DL, Lucas JE, Mitchell JT, et al. Cannabinoid exposure and altered DNA methylation in rat and human sperm. Epigenetics. 2018; 13:1208–21.
https://doi.org/10.1080/15592294.2018.1554521
PMID:30521419

32. Schrott R, Acharya K, Itchon-Ramos N, Hawkey AB, Pippen E, Mitchell JT, Kollins SH, Levin ED, Murphy SK. Cannabis use is associated with potentially heritable widespread changes in autism candidate gene *DLGAP2* DNA methylation in sperm. Epigenetics. 2020; 15:161–73.
https://doi.org/10.1080/15592294.2019.1656158
PMID:31451081

33. Chertkow-Deutsher Y, Cohen H, Klein E, Ben-Shachar D. DNA methylation in vulnerability to post-traumatic stress in rats: evidence for the role of the post-synaptic density protein Dlgap2. Int J Neuropsychopharmacol. 2010; 13:347–59.
https://doi.org/10.1017/S146114570999071X
PMID:19793403

34. Rasmussen AH, Rasmussen HB, Silahtaroglu A. The DLGAP family: neuronal expression, function and role in brain disorders. Mol Brain. 2017; 10:43.
https://doi.org/10.1186/s13041-017-0324-9
PMID:28870203

35. Sánchez AM, Delgado MS, Ciurana SG, Medrano J, Poo-Llanillo ME, Ishida M, Moore GE, Platas II, Simón C, Monk D. Epigenetic Symmetry of DLGAP2: Pre-Implantation Maternal Methylation Switches to a Random Monoallelic Profile in Somatic Tissues. OBM Genetics. 2018; 2:026.
https://doi.org/10.21926/obm.genet.1803026

36. Court F, Tayama C, Romanelli V, Martin-Trujillo A, Iglesias-Platas I, Okamura K, Sugahara N, Simón C, Moore H, Harness JV, Keirstead H, Sanchez-Mut JV, Kaneki E, 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; 24:554–69.
https://doi.org/10.1101/gr.164913.113
PMID:24402520

37. Carreras-Gallo N, Dwaraka VB, Jima DD, Skaar DA, Mendez TL, Planchart A, Zhou W, Jirtle RL, Smith R, Hoyo C. Creation and Validation of the First Infinium DNA Methylation Array for the Human Imprintome. bioRxiv. 2024; 2024.01.15.575646.
https://doi.org/10.1101/2024.01.15.575646
PMID:38293193

38. Garg P, Sharp AJ. Screening for rare epigenetic variations in autism and schizophrenia. Hum Mutat. 2019; 40:952–61.
https://doi.org/10.1002/humu.23740
PMID:30900359

39. Li J, Chen W, Li D, Gu S, Liu X, Dong Y, Jin L, Zhang C, Li S. Conservation of Imprinting and Methylation of *MKRN3*, *MAGEL2* and *NDN* Genes in Cattle. Animals (Basel). 2021; 11:1985.
https://doi.org/10.3390/ani11071985
PMID:34359112

40. Chen H, Victor AK, Klein J, Tacer KF, Tai DJ, de Esch C, Nuttle A, Temirov J, Burnett LC, Rosenbaum M, Zhang Y, Ding L, Moresco JJ, et al. Loss of MAGEL2 in Prader-Willi syndrome leads to decreased secretory granule and neuropeptide production. JCI Insight. 2020; 5:e138576.
https://doi.org/10.1172/jci.insight.138576
PMID:32879135

41. Gigliucci V, Busnelli M, Santini F, Paolini C, Bertoni A, Schaller F, Muscatelli F, Chini B. Oxytocin receptors in the Magel2 mouse model of autism: Specific region, age, sex and oxytocin treatment effects. Front Neurosci. 2023; 17:1026939.
https://doi.org/10.3389/fnins.2023.1026939
PMID:36998737

42. HPA. Human Protein Atlas.

43. Costa CIS, da Silva Montenegro EM, Zarrei M, de Sá Moreira E, Silva IMW, de Oliveira Scliar M, Wang JYT, Zachi EC, Branco EV, da Costa SS, Lourenço NCV, Vianna-Morgante AM, Rosenberg C, et al. Copy number variations in a Brazilian cohort with autism spectrum disorders highlight the contribution of cell adhesion genes. Clin Genet. 2022; 101:134–41.
https://doi.org/10.1111/cge.14072
PMID:34664255

44. Leung YK, Ouyang B, Niu L, Xie C, Ying J, Medvedovic M, Chen A, Weihe P, Valvi D, Grandjean P, Ho SM. Identification of sex-specific DNA methylation changes driven by specific chemicals in cord blood in a Faroese

birth cohort. Epigenetics. 2018; 13:290–300.
https://doi.org/10.1080/15592294.2018.1445901
PMID:29560787

45. Maggio AG, Shu HT, Laufer BI, Bi C, Lai Y, LaSalle JM, Hu VW. Elevated exposures to persistent endocrine disrupting compounds impact the sperm methylome in regions associated with autism spectrum disorder. Front Genet. 2022; 13:929471.
https://doi.org/10.3389/fgene.2022.929471
PMID:36035158

46. Liu Y, Di Y, Zheng Q, Qian Z, Fan J, Ren W, Wei Z, Tian Y. Altered expression of glycan patterns and glycan-related genes in the medial prefrontal cortex of the valproic acid rat model of autism. Front Cell Neurosci. 2022; 16:1057857.
https://doi.org/10.3389/fncel.2022.1057857
PMID:36568890

47. Sands TT, Miceli F, Lesca G, Beck AE, Sadleir LG, Arrington DK, Schönewolf-Greulich B, Moutton S, Lauritano A, Nappi P, Soldovieri MV, Scheffer IE, Mefford HC, et al. Autism and developmental disability caused by *KCNQ3* gain-of-function variants. Ann Neurol. 2019; 86:181–92.
https://doi.org/10.1002/ana.25522
PMID:31177578

48. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L, Stevens C, Reichert J, Mulhern MS, et al, and Autism Sequencing Consortium, and iPSYCH-Broad Consortium. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020; 180:568–84.e23.
https://doi.org/10.1016/j.cell.2019.12.036
PMID:31981491

49. Van Wauwe J, Kemps H, Vrancaert P, Mahy A, Schellingen R, Grootaert MOJ, Beerens M, Luttun A. PRDM16, a new kid on the block in cardiovascular health and disease. Cardiovasc Res. 2025; 121:1156–72.
https://doi.org/10.1093/cvr/cvaf089
PMID:40439125

50. Baizabal JM, Mistry M, García MT, Gómez N, Olukoya O, Tran D, Johnson MB, Walsh CA, Harwell CC. The Epigenetic State of PRDM16-Regulated Enhancers in Radial Glia Controls Cortical Neuron Position. Neuron. 2018; 98:945–62.e8.
https://doi.org/10.1016/j.neuron.2018.04.033
PMID:29779941

51. Umeda T, Takashima N, Nakagawa R, Maekawa M, Ikegami S, Yoshikawa T, Kobayashi K, Okanoya K, Inokuchi K, Osumi N. Evaluation of Pax6 mutant rat as a model for autism. PLoS One. 2010; 5:e15500.
https://doi.org/10.1371/journal.pone.0015500
PMID:21203536

52. Yoshizaki K, Furuse T, Kimura R, Tucci V, Kaneda H, Wakana S, Osumi N. Paternal Aging Affects Behavior in Pax6 Mutant Mice: A Gene/Environment Interaction in Understanding Neurodevelopmental Disorders. PLoS One. 2016; 11:e0166665.
https://doi.org/10.1371/journal.pone.0166665
PMID:27855195

53. Xing M, Zhang Q, Song W. GIGYF1-disturbed IGF-1R recycling: a potential contributor to autism spectrum disorder pathogenesis? J Clin Invest. 2022; 132:e163553.
https://doi.org/10.1172/JCI163553
PMID:36189799

54. Sun W, Poschmann J, Cruz-Herrera Del Rosario R, Parikshak NN, Hajan HS, Kumar V, Ramasamy R, Belgard TG, Elanggovan B, Wong CCY, Mill J, Geschwind DH, Prabhakar S. Histone Acetylome-wide Association Study of Autism Spectrum Disorder. Cell. 2016; 167:1385–97.e11.
https://doi.org/10.1016/j.cell.2016.10.031
PMID:27863250

55. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment interactions in psychiatric disorders: a new concept of clinical genetics. Clin Epigenetics. 2012; 4:1.
https://doi.org/10.1186/1868-7083-4-1
PMID:22414323

56. Plasschaert RN, Bartolomei MS. Tissue-specific regulation and function of Grb10 during growth and neuronal commitment. Proc Natl Acad Sci U S A. 2015; 112:6841–7.
https://doi.org/10.1073/pnas.1411254111
PMID:25368187

57. Garfield AS, Cowley M, Smith FM, Moorwood K, Stewart-Cox JE, Gilroy K, Baker S, Xia J, Dalley JW, Hurst LD, Wilkinson LS, Isles AR, Ward A. Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. Nature. 2011; 469:534–8.
https://doi.org/10.1038/nature09651
PMID:21270893

58. Chiesa N, De Crescenzo A, Mishra K, Perone L, Carella M, Palumbo O, Mussa A, Sparago A, Cerrato F, Russo S, Lapi E, Cubellis MV, Kanduri C, et al. The KCNQ1OT1 imprinting control region and non-coding RNA: new properties derived from the study of Beckwith-Wiedemann syndrome and Silver-Russell syndrome cases. Hum Mol Genet. 2012; 21:10–25.
https://doi.org/10.1093/hmg/ddr419
PMID:21920939

59. Mayo S, Monfort S, Roselló M, Oltra S, Orellana C,

Martínez F. In Pursuit of New Imprinting Syndromes by Epimutation Screening in Idiopathic Neurodevelopmental Disorder Patients. Biomed Res Int. 2015; 2015:341986.
https://doi.org/10.1155/2015/341986
PMID:26106604

60. Pauli S, Berger H, Ufartes R, Borchers A. Comparing a Novel Malformation Syndrome Caused by Pathogenic Variants in *FBRSL1* to AUTS2 Syndrome. Front Cell Dev Biol. 2021; 9:779009.
https://doi.org/10.3389/fcell.2021.779009
PMID:34805182

61. Bukvic N, De Rinaldis M, Chetta M, Trabacca A, Bassi MT, Marsano RM, Holoubkova L, Rivieccio M, Oro M, Resta N, Kerkhof J, Sadikovic B, Viggiano L. De Novo Pathogenic Variant in *FBRSL1*, Non OMIM Gene Paralogue *AUTS2*, Causes a Novel Recognizable Syndromic Manifestation with Intellectual Disability; An Additional Patient and Review of the Literature. Genes (Basel). 2024; 15:826.
https://doi.org/10.3390/genes15070826
PMID:39062605

62. Terrinoni A, Micheloni G, Moretti V, Caporali S, Bernardini S, Minieri M, Pieri M, Giaroni C, Acquati F, Costantino L, Ferrara F, Valli R, Porta G. *OTX* Genes in Adult Tissues. Int J Mol Sci. 2023; 24:16962.
https://doi.org/10.3390/ijms242316962
PMID:38069286

63. Huang B, Li X, Tu X, Zhao W, Zhu D, Feng Y, Si X, Chen JG. OTX1 regulates cell cycle progression of neural progenitors in the developing cerebral cortex. J Biol Chem. 2018; 293:2137–48.
https://doi.org/10.1074/jbc.RA117.001249
PMID:29273633

64. Redies C, Hertel N, Hübner CA. Cadherins and neuropsychiatric disorders. Brain Res. 2012; 1470:130–44.
https://doi.org/10.1016/j.brainres.2012.06.020
PMID:22765916

65. Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL, Hartemink AJ. Computational and experimental identification of novel human imprinted genes. Genome Res. 2007; 17:1723–30.
https://doi.org/10.1101/gr.6584707
PMID:18055845

66. Clair A, Baker E, Kumari M. Are housing circumstances associated with faster epigenetic ageing? J Epidemiol Community Health. 2023; 78:40–6.
https://doi.org/10.1136/jech-2023-220523
PMID:37816534

67. Wang L, Chen M, Yan G, Zhao S. DNA Methylation Differences Between Zona Pellucida-Bound and Manually Selected Spermatozoa Are Associated With Autism Susceptibility. Front Endocrinol (Lausanne). 2021; 12:774260.
https://doi.org/10.3389/fendo.2021.774260
PMID:34858344

68. Schubert C. Male biological clock possibly linked to autism, other disorders. Nat Med. 2008; 14:1170.
https://doi.org/10.1038/nm1108-1170a
PMID:18989289

69. Wu S, Wu F, Ding Y, Hou J, Bi J, Zhang Z. Advanced parental age and autism risk in children: a systematic review and meta-analysis. Acta Psychiatr Scand. 2017; 135:29–41.
https://doi.org/10.1111/acps.12666
PMID:27858958

70. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018; 10:573–91.
https://doi.org/10.18632/aging.101414
PMID:29676998

71. Loke YJ, Hannan AJ, Craig JM. The Role of Epigenetic Change in Autism Spectrum Disorders. Front Neurol. 2015; 6:107.
https://doi.org/10.3389/fneur.2015.00107
PMID:26074864

72. Zhou J, Park CY, Theesfeld CL, Wong AK, Yuan Y, Scheckel C, Fak JJ, Funk J, Yao K, Tajima Y, Packer A, Darnell RB, Troyanskaya OG. Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk. Nat Genet. 2019; 51:973–80.
https://doi.org/10.1038/s41588-019-0420-0
PMID:31133750

73. Litman A, Sauerwald N, Green Snyder L, Foss-Feig J, Park CY, Hao Y, Dinstein I, Theesfeld CL, Troyanskaya OG. Decomposition of phenotypic heterogeneity in autism reveals underlying genetic programs. Nat Genet. 2025; 57:1611–9.
https://doi.org/10.1038/s41588-025-02224-z
PMID:40634707

74. World Health Organization Laboratory Manual for the Examination and Processing of Human Semen. WHO Press: Geneva, Switzerland; 2010; 287.

75. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL, Fan JB, Shen R. High density DNA methylation array with single CpG site resolution. Genomics. 2011; 98:288–95.
https://doi.org/10.1016/j.ygeno.2011.07.007
PMID:21839163

76. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, Beck S. ChAMP: 450k

Chip Analysis Methylation Pipeline. Bioinformatics. 2014; 30:428–30.
https://doi.org/10.1093/bioinformatics/btt684
PMID:24336642

77. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013; 29:189–96.
https://doi.org/10.1093/bioinformatics/bts680
PMID:23175756

78. Nye MD, Hoyo C, Huang Z, Vidal AC, Wang F, Overcash F, Smith JS, Vasquez B, Hernandez B, Swai B, Oneko O, Mlay P, Obure J, et al. Associations between methylation of paternally expressed gene 3 (PEG3), cervical intraepithelial neoplasia and invasive cervical cancer. PLoS One. 2013; 8:e56325.
https://doi.org/10.1371/journal.pone.0056325
PMID:23418553

79. Murphy SK, Huang Z, Hoyo C. Differentially methylated regions of imprinted genes in prenatal, perinatal and postnatal human tissues. PLoS One. 2012; 7:e40924.
https://doi.org/10.1371/journal.pone.0040924
PMID:22808284

80. Krausz C, Sandoval J, Sayols S, Chianese C, Giachini C, Heyn H, Esteller M. Novel insights into DNA methylation features in spermatozoa: stability and peculiarities. PLoS One. 2012; 7:e44479.
https://doi.org/10.1371/journal.pone.0044479
PMID:23071498

81. Kruppa J, Sieg M, Richter G, Pohrt A. Estimands in epigenome-wide association studies. Clin Epigenetics. 2021; 13:98.
https://doi.org/10.1186/s13148-021-01083-9
PMID:33926513

82. Xie C, Leung YK, Chen A, Long DX, Hoyo C, Ho SM. Differential methylation values in differential methylation analysis. Bioinformatics. 2019; 35:1094–7.
https://doi.org/10.1093/bioinformatics/bty778
PMID:30184051

83. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013; 8:203–9.
https://doi.org/10.4161/epi.23470
PMID:23314698

84. Hansen KD. IlluminaHumanMethylation450kanno.ilmn12.hg19: Annotation for Illumina's 450k methylation arrays. R package version 0.6.0. 2016.

85. Feinberg JI, Bakulski KM, Jaffe AE, Tryggvadottir R, Brown SC, Goldman LR, Croen LA, Hertz-Picciotto I, Newschaffer CJ, Fallin MD, Feinberg AP. Paternal sperm DNA methylation associated with early signs of autism risk in an autism-enriched cohort. Int J Epidemiol. 2015; 44:1199–210.
https://doi.org/10.1093/ije/dyv028
PMID:25878217

86. Bethesda (MD): National Library of Medicine (US) NCfBI. 2004.

87. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014; 30:1363–9.
https://doi.org/10.1093/bioinformatics/btu049
PMID:24478339

88. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to pre-processing Illumina 450K methylation array data. BMC Genomics. 2013; 14:293.
https://doi.org/10.1186/1471-2164-14-293
PMID:23631413

89. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47.
https://doi.org/10.1093/nar/gkv007
PMID:25605792

90. Maksimovic J, Oshlack A, Phipson B. Gene set enrichment analysis for genome-wide DNA methylation data. Genome Biol. 2021; 22:173.
https://doi.org/10.1186/s13059-021-02388-x
PMID:34103055
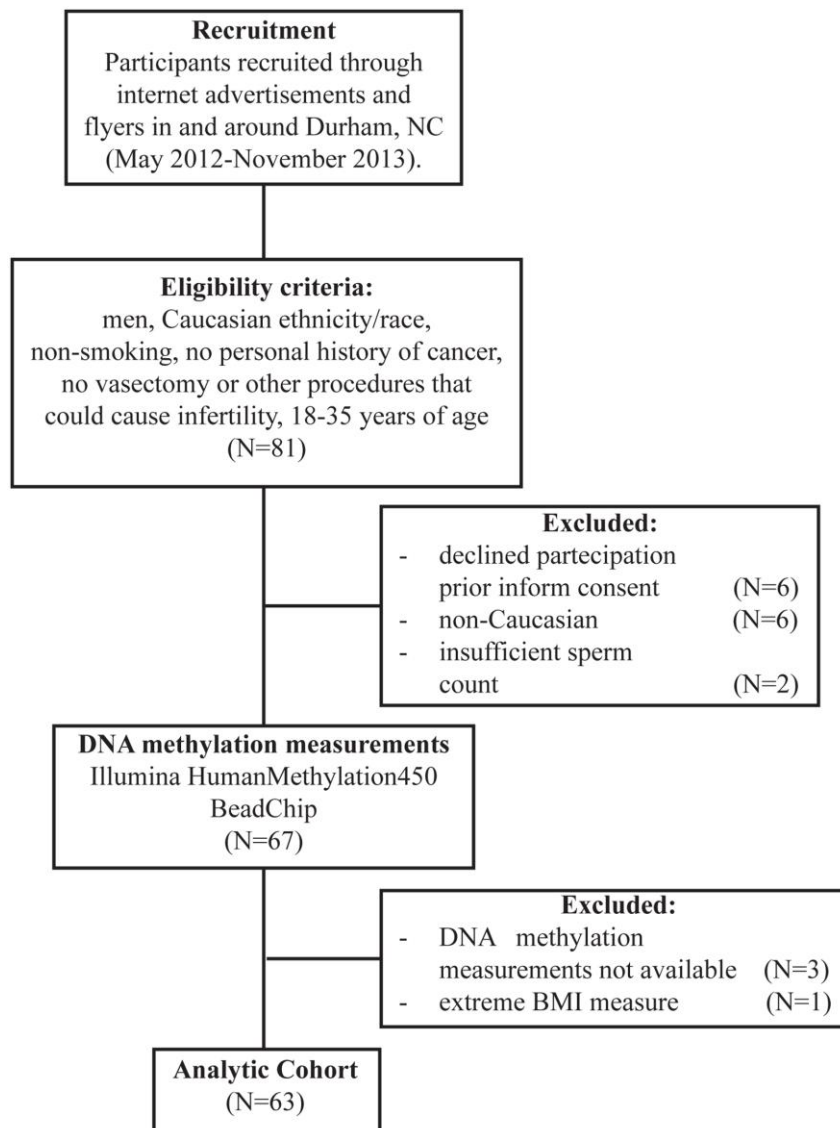
Supplementary Figure 1. Flow chart of study design. The influence of the Environment on Gametic Epigenetic Reprogramming (TIEGER) study design and inclusion/exclusion criteria are illustrated.
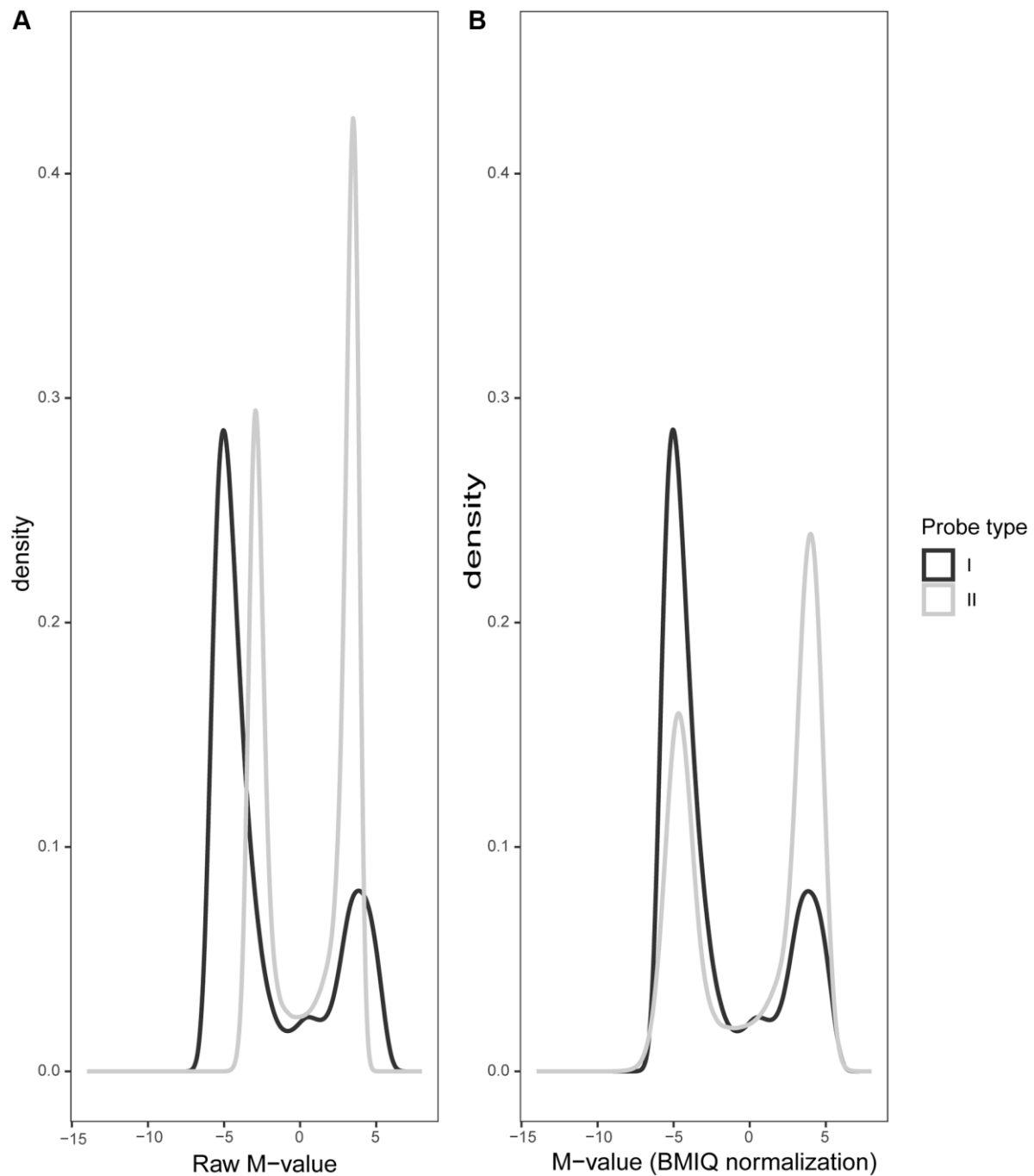
Supplementary Figure 2. Distribution of M-values per probe type. Distribution of: (**A**) raw M-values, and (**B**) M-values obtained from BMIQ-normalized β-values. This is shown for the Infinium I (black) and Infinium II (gray) probe types.
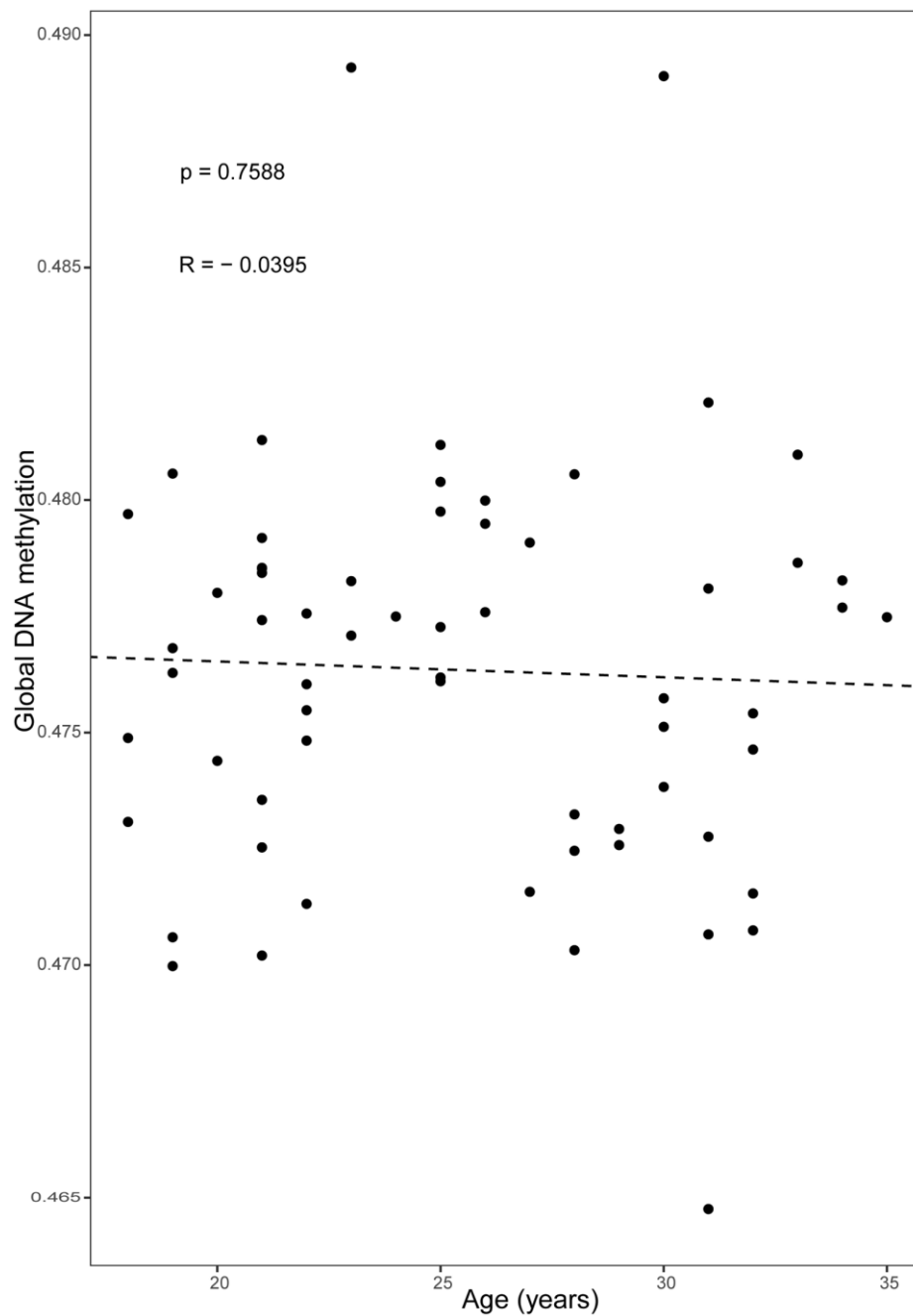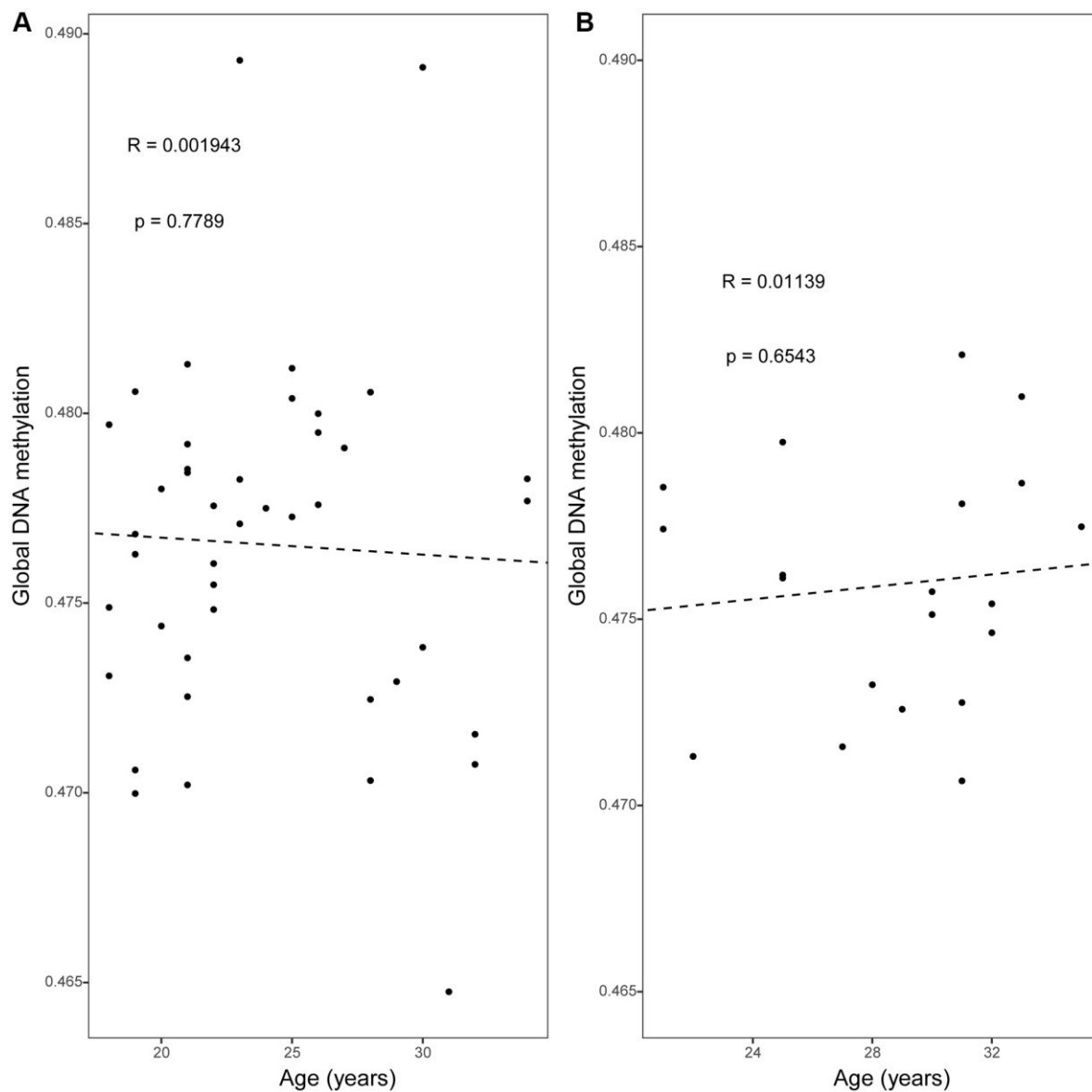
Supplementary Figure 3. Global DNA methylation in human sperm by age. Global DNA methylation per subject was calculated as the average DNA methylation at all CpG sites of the 450K. The estimated Pearson correlation coefficient (R = −0.0395, *p*-value = 0.7588) is reported. After removing two outliers, these results remained unchanged (R = −0.0724, *p*-value = 0.5793).
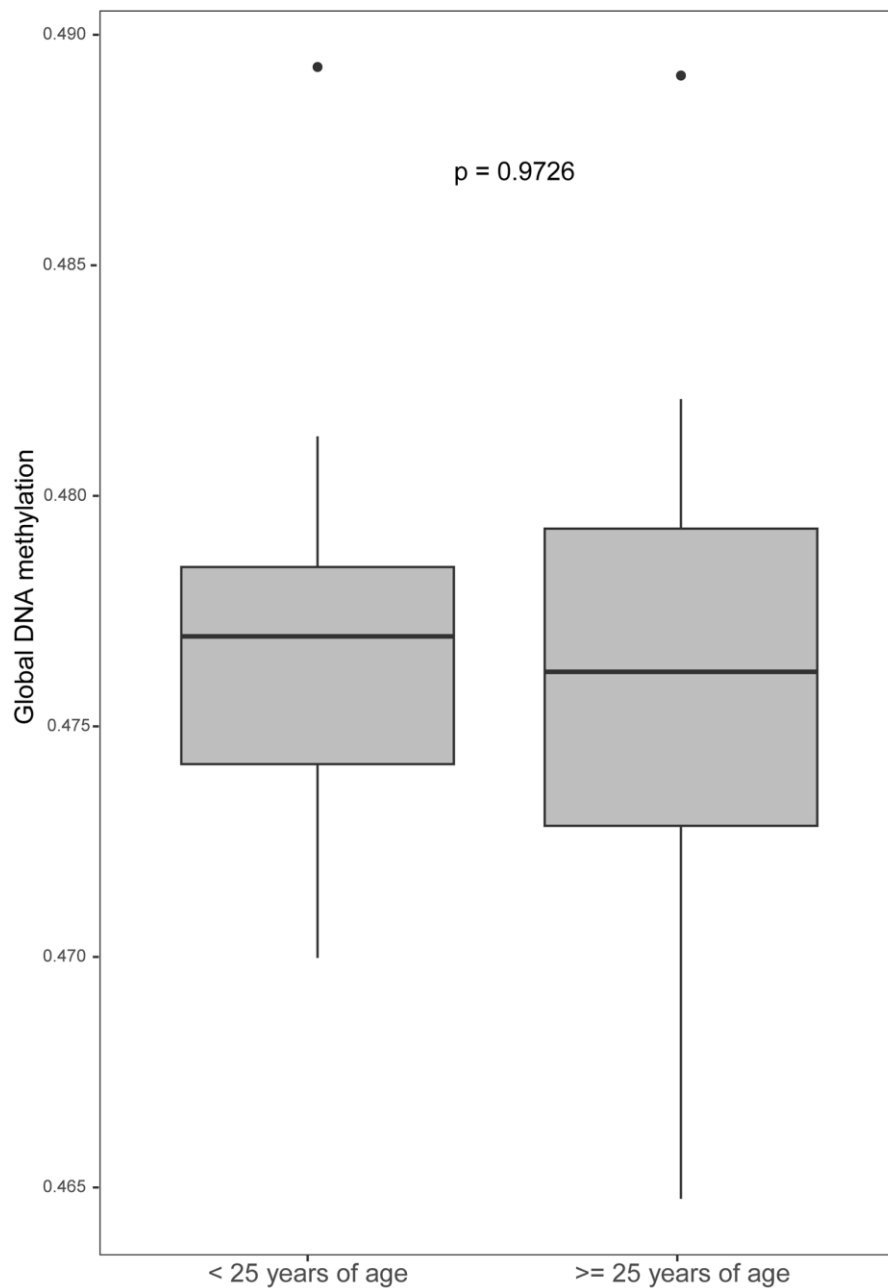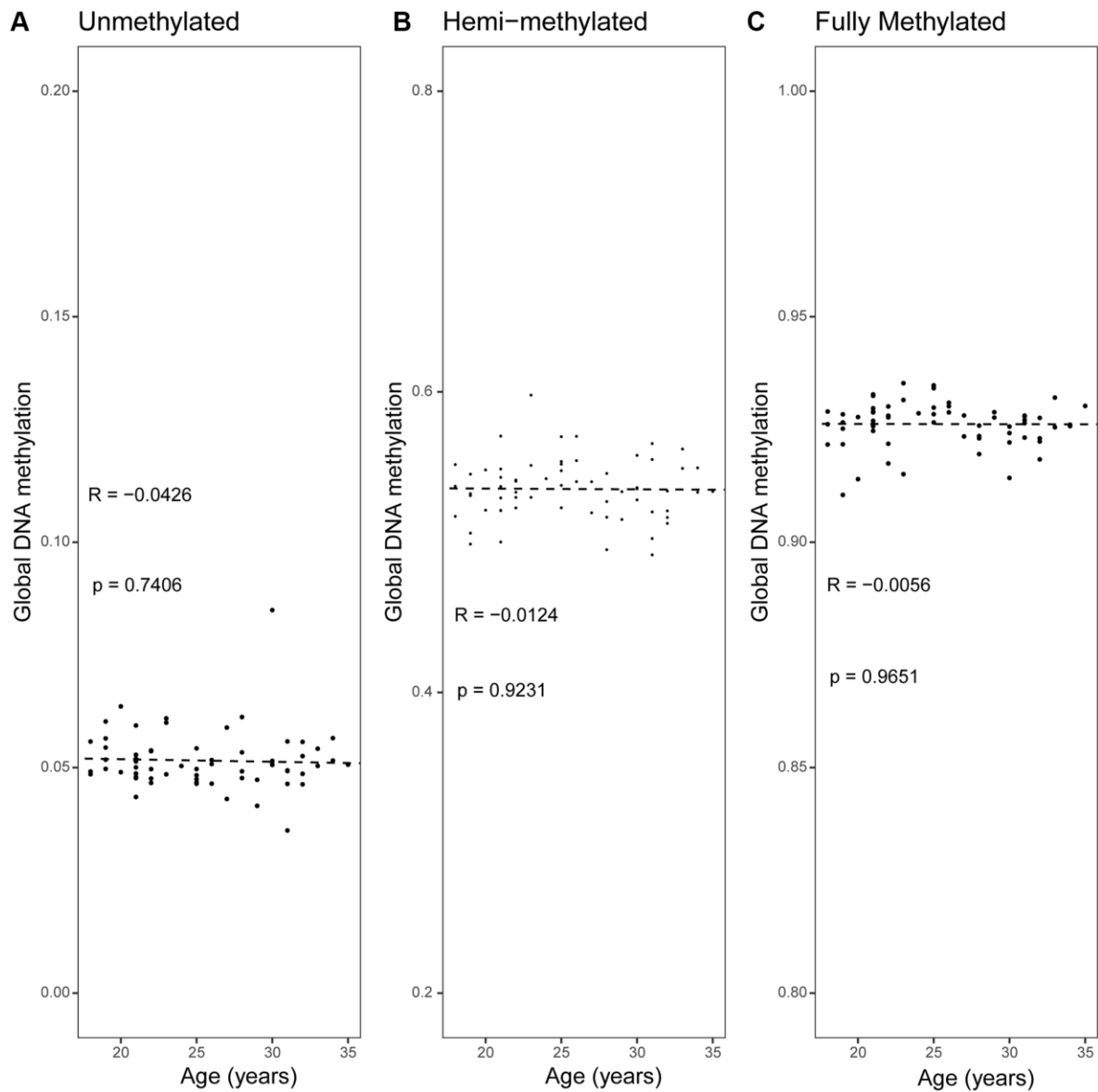
Supplementary Figure 4. Global DNA methylation in sperm of men above and below 25 years old. Boxplot comparing global DNA methylation levels in sperm from men below the median age (25 years old) *versus* above the median age in our study population. Global DNA methylation for each subject was calculated as the average DNA methylation at all CpG sites of the 450K. The differences in outcomes were evaluated by the Wilcoxon rank sum exact test (W = 493, *p*-value = 0.9726). After removing two outliers, our results remained unchanged (W = 452, *p*-value = 0.9254).
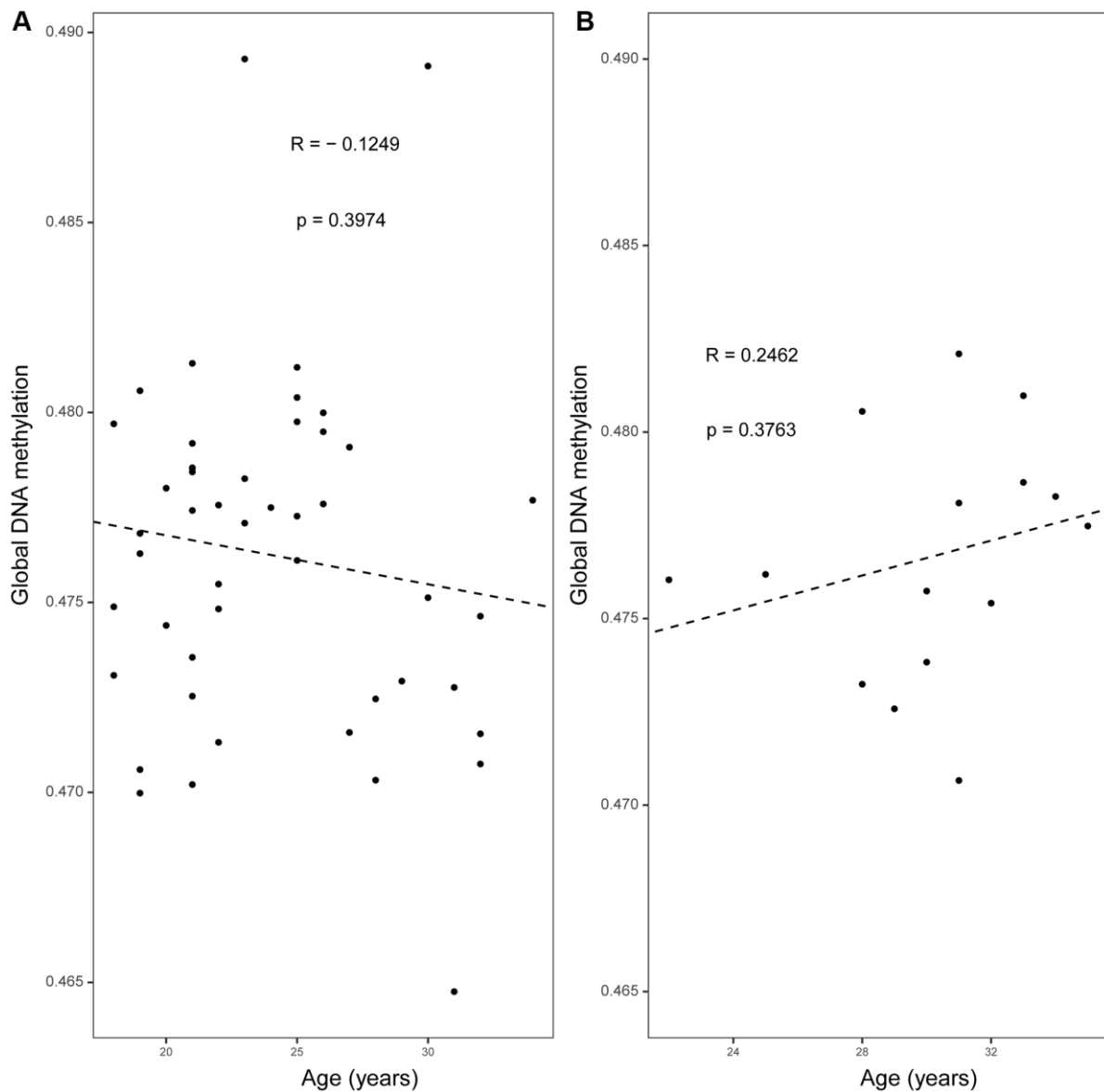
Supplementary Figure 5. Global DNA methylation in human sperm by age and stratified by patient status. Global DNA methylation, calculated as the average DNA methylation percentage at all CpG sites of the 450K, is represented per subject by age, in (**A**) non-patient subgroup (*n* = 48) and (**B**) patient subgroup (*n* = 15). Linear regression analysis shows no significant association between global DNA methylation and age in both subgroups (non-patients: *p*-value = 0.3974; patients: *p*-value = 0.3763).

Supplementary Figure 6. Global DNA methylation in human sperm by age and stratified by obesity status. Global DNA methylation, calculated as the average DNA methylation percentage at all CpG sites of the 450K, is represented per subject by age, in (**A**) normal weight subgroup (*n* = 43) and (**B**) overweight/obese weight subgroup (*n* = 20). Linear regression analysis shows no significant association between global DNA methylation and obesity status (normal weight men: *p*-value= 0.7789; overweight/obese men: *p*-value = 0.6543).

Supplementary Figure 7. Global DNA methylation in human sperm by age subdivided by DNA methylation subgroups. Linear regression models show no association between global DNA methylation levels and age by DNA methylation subgroups: (**A**) unmethylated CpG sites (R = −0.0426 and *p*-value = 0.7406), (**B**) hemi-methylated CpG sites (R = −0.0124 and *p*-value = 0.9231), and (**C**) fully methylated CpG sites (R = −0.0056 and *p*-value = 0.9651).
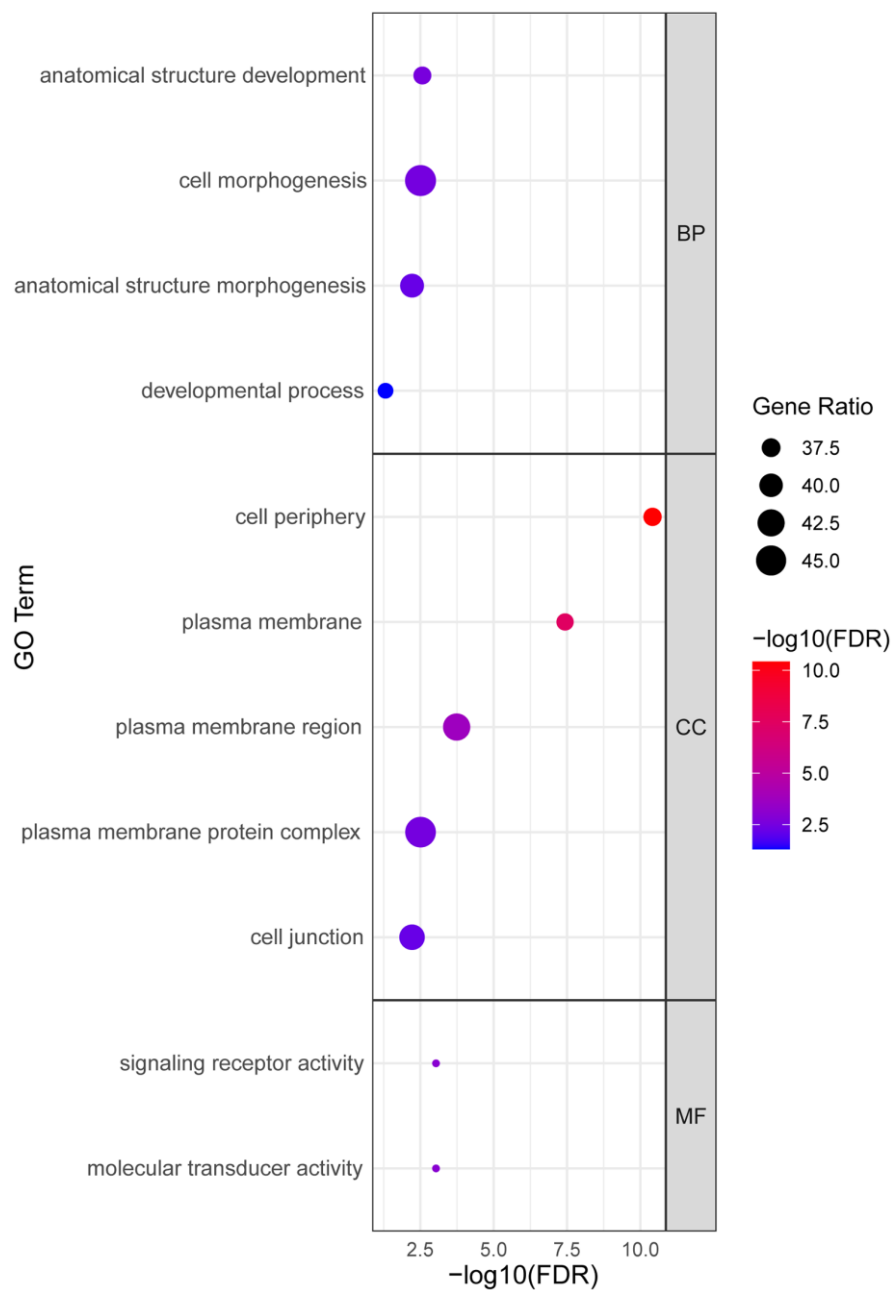
Supplementary Figure 8. Gene Ontology terms enrichment analysis. Eleven significant GO terms (FDR <0.05). X-axis shows the logarithmic transformation of adjusted *p*-values. Left Y-axis: GO term description. Right Y-axis: biological process (BP), cellular component (CC), and molecular function (MF). Gene ratio represents the percentage of genes that are differentially methylated out of the total number of genes in our 450K in the specific GO term.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 4–7.

Supplementary Table 1. Distribution of CpGs by island content within the 450K array.

CpG island content	Total of analyzed CpG sites		DNA methylation subgroups					
			UM		HM		FM	
	n	*%*	*N*	*%*	*N*	*%*	*n*	*%*
Island	149,092	30.91	131,110	87.94	4,875	3.27	13,107	8.79
Shore	111,367	23.09	67,762	60.85	15,613	14.02	27,992	25.13
Shelf	46,845	9.71	2,419	5.16	5,792	12.36	38,634	82.47
OpenSea	174,983	36.28	25,141	14.37	22,111	12.64	127,731	73.00
	482,287	100	226,432		48,391		207,464	

All CpG sites of the analytic 450K are targeted by a specific region of CpG neighborhood content, then presented by DNA methylation subgroups: Unmethylated (UM), Hemi-methylated (HM), or Fully methylated (FM) (UM: β-value <0.2; HM: β-value (0.2–0.8); FM: β-value >0.8).

Supplementary Table 2. Distribution of CpGs by island content after age-related analysis.

CpG island content	Differentially Methylated CpG sites								
	Total			UM		HM		FM	
	n	*%*[1]	*%*[2]	*n*	*%*	*n*	*%*	*n*	*%*
Island	722	4.94	0.15	325	13.46	280	4.38	117	2.01
	320[+] (44.32%); 402[-] (55.68%)			96[+]; 229[−]		143[+]; 137[−]		81[+]; 36[−]	
	Ratio: 1.26			Ratio: 2.38		Ratio: 0.95		Ratio: 0.44	
Shore	7,182	49.12	1.49	1,753	72.62	3,729	58.36	1,700	29.22
	278[+] (3.87%); 6,904[-] (96.13%)			66[+]; 1,687[−]		123[+]; 3,606[−]		89[+]; 1,611[−]	
	Ratio: 24.83			Ratio: 25.56		Ratio: 29.3		Ratio: 18.10	
Shelf	1,575	10.77	0.33	93	3.85	521	8.15	961	16.52
	258[+] (16.38%); 1,317[-] (83.62%)			2[+]; 91[-]		71[+]; 450[−]		185[+]; 776[−]	
	Ratio: 5.10			Ratio: 45.45		Ratio: 6.34		Ratio: 4.19	
Open Sea	5,143	35.17	1.06	243	10.07	1,860	29.11	3,040	52.25
	3,675[+] (71.46%);1,468[-] (28.54%)			33[+]; 210[−]		1,278[+]; 82[−]		2,364[+]; 676[−]	
	Ratio: 0.40			Ratio: 6.36		Ratio: 0.45		Ratio: 0.29	
All	14,622	100	3.03	2,414	100	6,390	100	5,818	100
	4,531[+] (30.99%); 10,091[-] 69.01%)			197[+]; 2,217[−]		1,615[+]; 4,775[−]		2,719[+]; 3,099[−]	

Significant age-related DMCs are grouped by DNA methylation subgroups: Unmethylated (UM), Hemi-methylated (HM), or Fully methylated (FM) (UM: β-value <0.2; HM: β-value (0.2–0.8); FM: β-value >0.8). Adjusted *p*-values are used (BH method; FDR ≤0.05). (+) positive correlation with age; (–) negative correlation with age. The ratio is calculated as a negative *versus* a positive correlation. (%) calculated relative to the total number of sites within subgroups. [1]% calculated in relation to the total number of DMCs (*n* = 14,622); [2]% calculated to the total number of CpGs tested at the 450K array.

Supplementary Table 3. Distribution of age-associated DMCs in sperm by functional genomic allocations.

	Differentially Methylated CpG sites							
	Total		UM		HM		FM	
Functional genomic distribution	*n*	*%*	*N*	*%*	*N*	*%*	*n*	*%*
Promoter	2,710	18.54	669	27.71	1,297	20.30	744	12.79
1stExon	160	1.09	15	0.62	50	0.78	95	1.63
5′UTR	873	5.97	154	6.38	365	5.71	354	6.08
Body	4,266	29.18	614	25.43	1,755	27.46	1,897	32.61
3′UTR	519	3.55	69	2.86	216	3.38	234	4.02
Intergenic	4,586	31.36	565	23.41	2,010	31.46	2,011	34.57
Multiple allocation	1,508	10.31	328	13.59	697	10.91	483	8.30
Total	14,622	100	2,414	100	6,390	100	5,818	100

Distributions are shown for all significant DMCs and by subgroups: Unmethylated (UM), Hemi-methylated (HM), or Fully methylated (FM) (UM: β-value <0.2; HM: β-value (0.2–0.8); FM: β-value >0.8). CpG sites allocated to multiple regions are reported as a separate category (multiple allocation).

Supplementary Table 4. Age-associated DMCs allocated to (predicted) 95 imprinted genes according to the Geneimprint database.

Supplementary Table 5. Age-associated DMCs allocated to 318 genes identified as nearest transcript of candidate ICRs.

Supplementary Table 6. Genes susceptible to ageing in sperm according to current and earlier studies.

Supplementary Table 7. Age-associated DMCs and allocated (predicted) imprinted genes linked to autism spectrum disorders and ranked by scoring criteria for biomarker selection.