A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP12013.

The Association between Long-Term DDT or DDE Exposures and an Altered Sperm Epigenome—a Cross-Sectional Study of Greenlandic Inuit and South African VhaVenda Men

Ariane Lismer,[1] Xiaojian Shao,[2,3] Marie-Charlotte Dumargne,[4]* Christine Lafleur,[5] Romain Lambrot,[5] Donovan Chan,[6] Gunnar Toft,[7] Jens Peter Bonde,[8,9] Amanda J. MacFarlane,[10] Riana Bornman,[11,12] Natalie Aneck-Hahn,[12] Sean Patrick,[12] Janice M. Bailey,[13] Christiaan de Jager,[11,12] Vanessa Dumeaux,[14,15] Jacquetta M. Trasler,[1,6,16,17] and Sarah Kimmins[1,5,18]*

[1]Department of Pharmacology and Therapeutics, Faculty of Medicine and Health Sciences, McGill University, Montreal, Quebec, Canada
[2]Digital Technologies Research Centre, National Research Council Canada, Ottawa, Ontario, Canada
[3]Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Canada
[4]Department of Animal Science, Faculty of Agricultural and Environmental Sciences, McGill University, Montreal, Quebec, Canada
[5]University of Montreal Hospital Research Centre, Montreal, Quebec, Canada
[6]Child Health and Human Development Program, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada
[7]Steno Diabetes Center Aarhus, Aarhus University Hospital, Aarhus, Denmark
[8]Department of Occupational and Environmental Medicine, Bispebjerg University Hospital, Copenhagen, Denmark
[9]Institute of Public Health, University of Copenhagen, Copenhagen, Denmark
[10]Agriculture Food and Nutrition Evidence Center, Texas A&M University, Fort Worth, Texas, USA
[11]Environmental Chemical Pollution and Health Research Unit, School of Health Systems and Public Health, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
[12]University of Pretoria Institute for Sustainable Malaria Control, School of Health Systems and Public Health, Faculty of Health Sciences, University of Pretoria, South Africa
[13]Research Centre on Reproduction and Intergenerational Health, Department of Animal Sciences, Université Laval, Quebec, Quebec, Canada
[14]Department of Anatomy and Cell Biology, Western University, London, Ontario, Canada
[15]Department of Oncology, Western University, London, Ontario, Canada
[16]Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada
[17]Department of Pediatrics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada
[18]Department of Pathology and Cell Biology, Faculty of Medicine, University of Montreal, Quebec, Canada

Background: The organochlorine dichlorodiphenyltrichloroethane (DDT) is banned worldwide owing to its negative health effects. It is exceptionally used as an insecticide for malaria control. Exposure occurs in regions where DDT is applied, as well as in the Arctic, where its endocrine disrupting metabolite, *p*,*p*′-dichlorodiphenyldichloroethylene (*p*,*p*′-DDE) accumulates in marine mammals and fish. DDT and *p*,*p*′-DDE exposures are linked to birth defects, infertility, cancer, and neurodevelopmental delays. Of particular concern is the potential of DDT use to impact the health of generations to come via the heritable sperm epigenome.

Objectives: The objective of this study was to assess the sperm epigenome in relation to *p*,*p*′-DDE serum levels between geographically diverse populations.

Methods: In the Limpopo Province of South Africa, we recruited 247 VhaVenda South African men and selected 50 paired blood serum and semen samples, and 47 Greenlandic Inuit blood and semen paired samples were selected from a total of 193 samples from the biobank of the INUENDO cohort, an EU Fifth Framework Programme Research and Development project. Sample selection was based on obtaining a range of *p*,*p*′-DDE serum levels (mean = 870.734 ± 134.030 ng/mL). We assessed the sperm epigenome in relation to serum *p*,*p*′-DDE levels using MethylC-Capture-sequencing (MCC-seq) and chromatin immunoprecipitation followed by sequencing (ChIP-seq). We identified genomic regions with altered DNA methylation (DNAme) and differential enrichment of histone H3 lysine 4 trimethylation (H3K4me3) in sperm.

Results: Differences in DNAme and H3K4me3 enrichment were identified at transposable elements and regulatory regions involved in fertility, disease, development, and neurofunction. A subset of regions with sperm DNAme and H3K4me3 that differed between exposure groups was predicted to persist in the preimplantation embryo and to be associated with embryonic gene expression.

Discussion: These findings suggest that DDT and *p*,*p*′-DDE exposure impacts the sperm epigenome in a dose–response-like manner and may negatively impact the health of future generations through epigenetic mechanisms. Confounding factors, such as other environmental exposures, genetic diversity, and selection bias, cannot be ruled out. https://doi.org/10.1289/EHP12013

Introduction

Dichlorodiphenyltrichloroethane (DDT) is a lipophilic organochlorine pollutant that is listed under the Stockholm Convention, with the objective to eliminate its use so as to protect human health and the environment.[1] However its use persists in Africa and India. Consequently, DDT continues to damage health and accumulates in the environment, becoming particularly concentrated in the aquatic food chain.[2,3] Currently, a sanctioned use of DDT is for indoor residual spraying to control the malarial disease vectors *Anopheles gambiae* and *Anopheles arabiensis*.[4] DDT is a potent endocrine disruptor, with DDT isomers exerting different endocrine disrupting actions: *p*,*p*′-Dichlorodiphenyldichloroethylene (*p*,*p*′-DDE) is antiandrogenic,[5] whereas *p*,*p*′-DDT and *o*,*p*′-DDT are estrogenic.[6] The main DDT metabolite, *p*,*p*′-DDE is associated with adverse health outcomes, including low birth weight,[7,8] urogenital malformations,[9] neurotoxicity,[7,10,11] cancer,[7] poor semen quality,[7,12] and increased risk of preterm birth.[7] Animal models[13] and epidemiological studies[14,15] indicate that paternal exposures to organochlorines and DDT, may impact health intergenerationally and, potentially, transgenerationally. These generational effects have been attributed to the epigenome, which refers to the biochemical content associated with DNA that impacts chromatin

*These authors contributed equally to this work.

Address correspondence to Sarah Kimmins. Email: sarah.kimmins@umontreal.ca. And, Vanessa Dumeaux. Email: vdumeaux@uwo.ca. And, Jacquetta M. Trasler. Email: jacquetta.trasler@mcgill.ca. And, Christiaan de Jager. Email: Tiaan.DeJager@up.ac.za

Supplemental Material is available online (https://doi.org/10.1289/EHP12013).

All authors declare they have nothing to disclose.

Received 17 August 2022; Revised 26 September 2023; Accepted 20 December 2023; Published 31 January 2024.

organization, as well as gene expression, and is transmitted via the gametes to alter phenotypes across generations.[16–18] The epigenome includes *a*) DNA methylation (DNAme) occurring at the 5′-position of cytosine residues within 5′-cytosine-phosphate-guanine-3′ (CpG) dinucleotides; *b*) the modifications on nucleosome proteins, histones, such as methylation, acetylation, and phosphorylation, among others; and *c*) small noncoding RNAs. A plausible molecular mechanism for paternal epigenetic inheritance is sperm-mediated transmission of regions bearing environmentally altered DNA and histone methylation.

In mammals, spermatogenesis entails specialized testis-specific gene expression programs that are accompanied by extensive remodeling of the chromatin landscape.[19–22] During this process, the majority of histones are evicted and replaced by protamines, with 1% of histones retained in sperm from mice[19,23] and 15% in sperm from men.[19,23,24] Retained histones are conserved across mammalian species and are found at gene regulatory regions implicated in spermatogenesis, embryogenesis, metabolism, and basic cellular processes.[24–26] We have shown in human[27] and mouse[16,18,28] sperm that the gene activating modification, histone H3 lysine 4 trimethylation (H3K4me3) localizes to genes involved in fertility, metabolism, and development. During spermatogenesis, disrupting the function of histone modifiers genetically[16,18,29,30] or via environmental exposures,[17,28] as well as modifying histone residues in the paternal pronucleus of the zygote,[31,32] provides compelling evidence that sperm-transmitted histones serve critical functions in embryo development and offspring health. A clear indicator that sperm histones may be transmitted at fertilization and retained in the embryo is the localization of sperm-enriched histone variant H3.3 to the paternal pronucleus of the zygote.[31–33] Whether histones in sperm from men, specifically histone H3K4me3, are similarly transmitted to alter embryonic gene expression is unknown. In support of this possibility, H3K4me3 peaks in human sperm occurred at genes that are expressed in the preimplantation embryo.[27]

Human sperm DNA is highly methylated,[34] and disruption of methyl signatures is related to infertility and abnormal offspring development.[35–37] Promoters with high CpG content are enriched for nucleosomes and predominantly hypomethylated in the sperm of mice[23] and humans.[23,24] Although sperm epigenetic modifications that respond to environmental exposures are most frequently studied alone, there are likely interdependent interactions between DNA methylation and chromatin in sperm. For example, there is cooperativity between chromatin and DNA methylation machinery; in mouse embryonic stem cells, DNMT3L acts via the amino terminus of H3K4me to recruit the *de novo* DNA methyltransferase. When H3K4 was methylated, this interaction with DNMT3L was abrogated and DNA methylation was blocked.[38] Whole-genome bisulfite chromatin immunoprecipitation targeting histone H3K4me3 followed by sequencing (ChIP-seq) of human sperm revealed that over 24,000 H3K4me3 peaks coincided with regions of intermediate (between 20% and 80%) and high (>80%) DNA methylation.[27] These overlapping regions occurred at genes implicated in fertility and development.[27] How sperm DNAme and chromatin respond in concert to environmental exposures, including endocrine disrupting chemicals (EDCs), remains underexplored. It is also unknown whether common H3K4me3 and DNAme regions are sensitive to DDT and *p*,*p*′-DDE and could be transmitted and retained in the embryo.

The use of animal models in combination with advances in epigenomic techniques have directly implicated aberrations in the establishment and maintenance of the germline epigenome to birth defects,[16–18] neurodevelopmental disorders,[18] and common diseases, such as diabetes[28] and cancer.[39,40] In rodents, sperm DNAme and histone modifications can be altered not only by environmental exposure to toxicants[41] but also by obesity[28] and nutrient restriction.[17,42,43] These epigenetic aberrations in sperm can subsequently be transmitted to the embryo to alter embryonic gene expression, development, and offspring health.[17,43] Exposure to organochlorines, including DDT in rodents, has been linked with altered DNAme in sperm and associated with poor reproductive outcomes and offspring metabolic disease.[44–46] In a prior study using the same Greenlandic Inuit population as this study, *p*,*p*′-DDE serum levels were examined in relation to sperm DNA methylation by targeted pyrosequencing at Satα tandem repeat sequences, as well as at long interspersed nuclear element 1 (LINE-1) and *Arthrobacter luteus* (Alu) transposable elements. Global DNA methylation levels were also measured by flow cytometric fluorescence. Using these approaches, the results were inconclusive as to whether sperm DNAme is altered by *p*,*p*′-DDE exposures.[47] Therefore, whether the sperm epigenome of men is impacted by DDT and *p*,*p*′-DDE exposures and could be associated with paternal epigenetic transmission remains unresolved.

Given the improvement in epigenome-wide sequencing approaches for epigenome-wide association studies, we revisited the response of the sperm methylome to DDT and *p*,*p*′-DDE. In this study, we aimed to identify alterations in sperm DNAme and histone H3K4me3 associated with levels of the persistent DDT metabolite measured in blood and how these epigenetic alterations could be implicated in epigenetic inheritance. We studied two geographically diverse exposed populations, Greenlandic Inuit and South African VhaVenda men. To quantitatively identify regions that were altered in DNAme and histone methylation of exposed men, we used a sperm-customized methyl-capture approach followed by sequencing (MCC-seq; Greenlandic and South African populations), and ChIP-seq (South African population exclusively). We then performed differential and functional analyses to define altered DNAme and H3K4me3 regions associated with high serum levels of *p*,*p*′-DDE. *In silico* analyses were used to further explore the possibility that these epigenetic alterations in sperm could be transmitted to the embryo at fertilization, impact embryonic gene expression, and persist throughout embryonic development.

Materials and Methods

Ethical Statements

For the Greenlandic Inuit cohort of the biobank of the EU Fifth Framework Programme Research and Development project, INUENDO, the scientific ethics committee for Greenland approved the research protocol (project reference no. 2014-25/26). For the South African VhaVenda cohort, the research protocol was approved by the Scientific Ethics Committee of the Faculty of Health Sciences, University of Pretoria, and the Limpopo Provincial Government's Department of Health (project reference no. 43/2003). For the Greenland Inuit cohort, the research protocol was approved by Health Canada and Public Health Agency of Canada's research ethics boards (project reference no. REB 2019–0006) and by the ethics committee for the Faculty of Medicine and Health Sciences, McGill University (project reference no. A09-M57-15B).

Greenlandic Inuit Population Blood and Semen Collection

Indigenous Arctic inhabitants consume a traditional marine mammal diet consisting of whale, walrus, and seal and are at a high risk of exposure to *p*,*p*′-DDE through its bioaccumulation in the marine food chains.[48,49] Greenlandic Inuit blood and semen paired samples were selected from the biobank of the INUENDO cohort.[50] The study participants ranged from 20 to 44 years of age (mean age = 31 y), were born in Greenland, and all had proven fertility with confirmation of a pregnant partner. Sample

selection was based on obtaining a range in *p*,*p*′-DDE serum levels (mean = 870.734 ± 134.030 ng/mL; Figure S1 and Table S1) ($n = 47$ selected from a cohort of 193 total for MCC-seq). Participants were recruited between May 2002 and February 2004, and eligible men were born in Greenland. Full details on recruitment and the cohort have been previously described.[50] Data was available on smoking (questionnaire and cotinine levels), DNA fragmentation index and body mass index (BMI) (see aggregate data in Table S1). Cotinine was measured from the paired blood sample, as previously described.[51]

Note, for adherence with General Data Protection Regulations (GDPR), individual data cannot be published. Semen samples from participants who gave informed consent were collected between May 2002 and February 2004 by masturbation in a private room, and blood was collected within 1 wk of semen collection, except for a subgroup from which blood was collected within 1 y. The men were asked to abstain from sexual activities for ≥2 d before collecting the sample. Immediately after collection, semen samples were kept close to the body to maintain a 37°C temperature when transported to the laboratory. Two cryotubes with 0.2-mL aliquots of undiluted raw semen collected 30 min after liquefaction, were prepared from each semen sample, and long-term storage was in a −80°C freezer. The blood samples were centrifuged immediately after collection, and sera were stored in a −80°C freezer for later analysis. Samples were analyzed in the Department of Laboratory Medicine in the Division of Occupational and Environmental Medicine in Lund, Sweden, as previously described.[52,53] Briefly, the *p*,*p*′-DDE was extracted by solid-phase extraction using on-column degradation of the lipids, and analysis was by gas chromatography–mass spectrometry (GC-MS).[54] The relative standard deviations (SDs), calculated from samples analyzed in duplicate at different days, were 1% at 1 ng/mL ($n = 1{,}058$), 8% at 3 ng/mL ($n = 1{,}058$), and 7% at 8 ng/mL ($n = 1{,}058$), and the limit of detection (LOD) was 0.1 ng/mL for *p*,*p*′-DDE.

South African Population Blood and Semen Collection

The Vhembe district is a malaria-endemic area, where housing includes mud or brick or cement dwellings that are sprayed with DDT or not, to control for malaria. Participants volunteered from 12 villages from the Vhembe district of the Limpopo Province of South Africa that either sprayed via indoor residual spraying ($n = 33$ participants) or did not spray ($n = 17$ participants). This prospective study was conducted in the same manner as our prior studies, and full details on the recruitment methods and the questionnaire have been previously described.[55] Semen and blood were collected on the same day in either October 2016, February 2017, or November 2017. Men were excluded from the study if they were <18 or >40 years of age, appeared intoxicated, reported drug use, or had a neuropsychiatric illness. Physical measurements included height and weight. All participants provided informed consent and were interviewed on the same day of sample collection using a yes or no format questionnaire on their use of pesticides (how often, is protection worn, are pesticides used for their job, what type of pesticide) smoking, and drinking (how many alcoholic beverages consumed per day, per week, and what kind) (see aggregate data in Table S2). Fertility status was not queried. Of 247 men enrolled in the study, we selected 50 paired blood and semen samples from men who ranged from 18 to 32 years of age (mean age = 25 y). Sample inclusion was based on normal sperm counts (>15 million/mL), normal sperm DNA fragmentation index, and testing a range of *p*,*p*′-DDE serum levels (mean = 10,462.228 ± 1,792.298 ng/mL) (Figure S1 and Table S2). As noted above, for GDPR adherence, individual data cannot be published. The semen was preserved in Sperm Freeze (LifeGlobal) and stored in liquid nitrogen for transport, followed by long-term storage in a −80°C freezer. Semen analysis was performed according to the World Health Organization (WHO) standard,[56] and the DNA fragmentation index was determined as detailed by de Jager et al.,[55] using the sperm chromatin structure assay[57] (Table S2). Semen analyses involved a macroscopic and microscopic examination of sperm that considered sperm count, sperm motility, and sperm morphology.

Serum lipid measurements and analysis for *p*,*p*′-DDE, *p*,*p*′-DDT, *o*,*p*′-DDE, and *o*,*p*′-DDT were determined by the Center de Toxicologie du Québec (CTQ) of the Institut National de Santé Publique du Québec (INSPQ) from serum samples. This CTQ method and the following details are from Tillaut et al.[58] Two microliters of serum samples were enriched with internal polychlorinated biphenyl (PCB)-labeled standards: PCB 141-$^{13}C_{12}$, PCB 153-$^{13}C_{12}$, PCB 180-$^{13}C_{12}$, PCB 194-$^{13}C_{12}$, hexachlorobenzene-$^{13}C_6$, α-HCH-$^{13}C_6$, *trans*-nonachlor-$^{13}C_{10}$, oxychlordane-$^{13}C_{10}$, *p*,*p*′-DDE-$^{13}C_{12}$, Parlar 26-$^{13}C_{10}$, Parlar 50-$^{13}C_{10}$, α-endosulfane-$^{13}C_9$, dieldrin-$^{13}C_{12}$, and endrin-$^{13}C_{12}$, and alcohol-denatured proteins. Internal standards were provided by the CTQ INSPQ.

> The POP [persistent organic pollutant] compounds (PCBs and OCPs [organochlorine pesticides]) were extracted with hexane from the aqueous matrix using a liquid–liquid extraction in the presence of a saturated ammonium sulfate solution. These extracts were cleaned up on deactivated 0.5% florisil columns. Elution was broken down into 2 steps: the first fraction was eluted with a mixture of dichloromethane:hexane (25:75; 9 mL) and contain[ed] all compounds except heptachlor epoxide, endrin, dieldrin, [α-endosulfan and β-endosulfan], which were eluted in the second fraction with a mixture of acetone (dichloromethane, 2:98; 4 mL). The solvent of the first fraction was evaporated, taken up in 125 µL of hexane concentrated to 20 µL, and analyzed for the PCBs and OCPs on an Agilent 6890 Network or 7890A [GC] equipped with an Agilent 7683 or 7693 series automatic injector and an Agilent 5973 Network or 5975C [MS] (Agilent Technologies Inc.; Mississauga, ON, Canada). The GC was fitted with an Agilent 60 m DB-XLB column (0.25 mm i.d., 0.25 µm film thickness) to the MS and with an Agilent Ultra-1 50 m (0.20 mm i.d., 0.33 µm film thickness) to the ECD [electron-capture dissociation]. The carrier gas was helium, and the injections were 3 µL in splitless mode. Fraction 2 was also evaporated, taken up in 20 µL of acetonitrile, and analyzed on a GC-MS [as] described above. The injection was 2 µL in splitless mode. All the MSs were operated in selected ion monitoring, using negative ion chemical ionization, with methane (99.97%) as the reagent gas. Total cholesterol (TC), free cholesterol (FC), triglycerides and phospholipids levels were also measured in these samples by enzymatic methods (in g/L) and allowed to calculate the total lipid level as $1.677 * (\text{total cholesterol} - \text{FC}) + \text{FC} + \text{triglycerides} + \text{phospholipids}$.[58]

The LODs of OCPs ranged between 0.01 and 0.05 µg/L, depending on the analyte. The internal reference materials used to control the quality of the analyses were the certified reference material SRM-1958 from the National Institute of Standards and Technology and the noncertified reference material W-16-04 was from the Artic Monitoring and Assessment Program (AMAP) External Quality Assessment Scheme (CTQ INSPQ). The overall quality and accuracy for the analytical method was monitored by the participation to the interlaboratory program as the AMAP External Quality Assessment Scheme (CTQ INSPQ) for the

analytes *p*,*p*′-DDT, *p*,*p*′-DDE, TC, triglycerides, and total lipids, as well as the German External Quality Assessment Scheme for the analytes *p*,*p*′-DDT and *p*,*p*′-DDE.

DNA Isolation for MCC-Seq on Greenlandic Inuit and South African Sperm

Approximately 10 million spermatozoa were lysed overnight at 37°C in a buffer containing a final concentration of 150 mM Tris-hydrochloride (Tris-HCl), 10 mM ethylenediaminetetraacetic acid (EDTA), 40 mM dithiothreitol (DTT), 2 mg/mL proteinase K, and 0.1% sarkosyl detergent. DNA was then extracted using the QIAamp DNA Mini kit (Qiagen) according to the manufacturer's protocols. MCC-seq was performed as previously described,[59] which involves an initial preparation of whole-genome bisulfite sequencing (WGBS) libraries followed by targeted capturing of our regions of interest. Briefly, 1–2 μg of DNA was sonicated (Covaris) to obtain fragments of 300–400 bp and was analyzed on a Bioanalyzer 1000 Chip (Agilent). Following this, the KAPA Biosystems' protocols were used for DNA end repair, 3′-end adenylation, adapter ligation, and cleanup steps. Using the Epitect Fast bisulfite kit (Qiagen), samples were bisulfite converted according to the manufacturer's protocol, followed by quantification with OliGreen (Life Technology). Polymerase chain reaction (PCR) amplification with 9–12 cycles using the KAPA HiFi HotStart Uracil+ DNA Polymerase (Roche/KAPA Biosystems) was performed according to suggested protocols. The final WGBS libraries were purified using Agencourt XP AMPure beads (Beckman Coulter), validated on Bioanalyzer High Sensitivity DNA Chips (Agilent), and quantified using PicoGreen (ThermoFisher). After WGBS library preparations of all samples, CpGs targeted by our recently developed human sperm panel[59] were captured using the SeqCap Epi Enrichment System protocol by RocheNimbleGen. Our custom SeqCap Epi panel covers 3,179,096 CpG sites that are found mainly in intergenic (35%), intronic (33%), and promoter–transcriptional start site (TSS) (19%) regions. Of the captured CpGs, 937,141 had been selected based on their environmental sensitivity, high variability, and intermediate-DNA methylation levels (20%–80%) in a WGBS data set of pooled sperm from 30 men, and the remainder CpGs corresponded to sites from the commonly used Infinium MethylationEPIC BeadChip array that targets gene promoters/CpG island regions and enhancers.[59] Equal amounts of multiplexed libraries (84 ng of each; 12 samples per capture) were combined to obtain 1 μg of total input library and was hybridized at 47°C for 72 h to the capture panel, specifically, the human sperm capture panel that was recently developed by our group. This was followed by washing, recovery, PCR amplification of the captured libraries, and final purification according to the manufacturer's recommendations. Quality, concentration, and size distribution of the final captured libraries were determined using Bioanalyzer High Sensitivity DNA Chips (Agilent). The capture libraries were sequenced with a 200-cycle S2 kit (100-bp paired-end sequencing) on the NovaSeq 6000 following the NovaSeq XP workflow (Tables S3 and S4).

MCC-Seq Data Preprocessing

R (version 3.6.0; R Development Core Team) was used coupled to the Bioconductor packages from the 3.9 release. The sperm MCC-seq data were processed using the Genpipes pipeline.[60] Specifically, the high-throughput bisulfite sequencing reads (i.e., paired-end FASTQ reads) of MCC-seq were first trimmed using Trimmomatic (version 0.36)[61] by applying the following filtering criteria: *a*) remove low quality bases (i.e., phred33 ≥30) and short reads (length $n > 50$ bp), and *b*) remove the Illumina adapters. The trimmed reads were then aligned to the *in silico* bisulfite-converted reference genome with hg19/GRCh37 version using Bismark (version 0.18.1).[62] The alignments were performed with the Bowtie2 aligner (version 318 2.3.1)[63] in paired-end mode and with the nondirectional protocol setting, as well as with the other default parameters. The aligned Binary Alignment Map (BAM) files per lane were merged, and duplicated reads were further removed with Picard (version 2.9.0; MusicBrainz). Finally, the cytosine-level methylation calls were obtained using Bismark where counts of methylated and unmethylated reads at each cytosine position in the genome were recorded. DNA methylation level of each CpG was then defined as the number of methylated reads divided by the total number of sequenced reads, where the latter is the sum of the methylated and unmethylated reads. In addition, CpG sites that were located within the Data Analysis Center (DAC) Blacklisted Regions or Duke Excluded Regions [generated by the Encyclopedia of DNA Elements (ENCODE) project][64] or overlapped with single-nucleotide polymorphisms (SNPs) (i.e., dbSNP 137) were removed. CpG sites with low read coverage (i.e., <20×) were also discarded. Sperm sample purity (absence of somatic cell contamination) was assessed by examination of sequencing data for imprinted loci.

Single-Nucleotide Polymorphism and Genotype Analysis

BiSNP (version 0.82.2)[65] was run on the deduplicated BAM files to call variants (including homozygous alternate and heterozygous genotypes). A set of unique variants called from all individuals was used as the potential SNP set. The homozygous reference genotypes of individuals on these SNPs were extracted from the aligned BAM files by requiring ≥10× read coverage aligned to the reference allele. SNPs with genotypes inferred from all individuals were kept for downstream analysis. Principal component analysis (PCA) on genotype profiles of South African and Greenlandic participants relative to African, American, Asian, and European profiles were used to investigate the genetic background distribution of the samples used in this study (Figure S1C–E). Chromosome 1 genotype data from 1,000K genomes (Imputed HapMap 3; https://www.broadinstitute.org/medical-and-population-genetics/hapmap-3) were used as the reference genotype profile (the human population genetic background) to compare with that of South African and Greenlandic cohorts. To assess the relationship between the *p*,*p*′-DDE serum group and individual genotype profiles, we selected the common SNPs on chromosome 1 between the South African and Greenlandic men, as well as the 1,000K genome data sets (Imputed hapmap V3; total = 550 SNPs), performed PCAs on the SNPs for both cohorts, and colored each sample by *p*,*p*′-DDE serum tertiles (Figure S1D,E).

Differentially Methylated CpGs and Merged Differentially Methylated CpGs Identification

Generalized linear regression models were built using the methylation proportion inferred from the combination of methylated reads and unmethylated reads as a binomially distributed response variable to look for associations between DNAme in sperm and *p*,*p*′-DDE serum levels. For the downstream analyses, we opted for the continuous regression *p*,*p*′-DDE model, but for visualization purposes, samples from each cohort were separated into *p*,*p*′-DDE serum level tertiles (Figure 1D). For some CpGs, the number of individuals with sufficient sequencing coverage (≥20×) was low (e.g., <30 samples); these CpGs were removed from our analyses, to minimize the impact of low measurement accuracy. Nonvariable CpGs (SD = 0) were also removed to reduce the multiple testing burden. For the South African cohort, from a total number of 2,354,599 CpGs with 20× coverage, and covered by at least 1 sample, we obtained 1,573,641 CpGs with 20× coverage, and these CpGs were covered by >30 samples (66.8% of total CpGs were retained after removing low-coverage CpGs).

Furthermore, 3,327 CpGs showed nonvariable methylation (0.14% of total CpGs, or 0.2%, after removing low-coverage CpGs). For the Greenlandic cohort, from a total number of 2,458,895 CpGs with 20× coverage and covered by at least 1 sample, we obtained 1,728,019 CpGs with 20× coverage, and these CpGs were covered by >30 samples (70.3% of total CpGs were retained after removing low-coverage CpGs). Furthermore, 4,034 CpGs showed nonvariable methylation (0.16% of total CpGs, or 0.23%, after removing low-coverage CpGs). R function glm() and the binomial family were used to fit each model, and *p*-values for variables of interest were obtained accordingly. The obtained *p*-values were then corrected by estimating the false discovery rate (FDR) *q*-values using the Bioconductor/R package qvalue (version 2.16).[66] We defined significant associated differentially methylated CpGs (DMCs) when *q*-values were <0.01. Furthermore, consecutive DMCs that spanned 500 bp with methylation changes in the same direction were merged to define differentially methylated regions.

MCC-Seq Hotspot Cluster Analysis

"Hotspot," or cluster, analysis was performed by calculating the ratio of DMCs with DNAme gain or loss over the total number of CpGs found within 1 Mb nonoverlapping bins over the genome; densities >10% (termed clusters) were extracted for further analysis. To investigate genetic effects on DNAme of CpGs within the clusters, methylation quantitative trait locus (meQTLs) analyses were performed. The genotype profiles of SNPs within all the candidate clusters, as well as all the DMCs, were extracted. By considering possible SNP *cis*-effects within 250 kb of a CpG (i.e., a 500-kb window), meQTLs were calculated using Bioconductor/R package MatrixEQTL (version 2.3) with default parameters.[67] The reported *p*-values were corrected using the Benjamini-Hochberg FDR (Tables S5 and S6).[68]

ChIP-Seq on South African Sperm

Tubes containing South African sperm were thawed at room temperature and pelleted by centrifugation at 2,500 × *g* for 5 min at 4°C to remove the freezing medium supernatant. Each pellet was dissolved in 1 mL of Ferticult Flushing medium. Samples were centrifuged at 2,500 × *g* for 5 min at 4°C, the supernatant was removed, and an additional 500 µL of Ferticult Flushing medium was added to the sperm pellet. The sperm were incubated at room temperature for 45 min to allow for a swim-up. Fractions containing motile sperm were retrieved for downstream ChIP. ChIP on human sperm was performed as previously described,[27,69] with slight modifications. Sperm sample purity (absence of somatic cell contamination) was assessed visually in a hemocytometer counting chamber. Briefly, per sample, 12 million spermatozoa were resuspended in 300 mL of buffer 1 [15 mM Tris-HCl, 60 mM potassium chloride, 5 mM magnesium chloride ($MgCl_2$), and 0.1 mM egtazic acid (EGTA)] that contained 0.3 M sucrose and 10 mM DTT. The samples were then each split into six tubes (2 million spermatozoa cells/tube) and 50 mL of buffer 1 was supplemented with 0.5% NP-40, and 1% sodium deoxycholate and added to each tube. After a 30-min incubation on ice, 100 mL of micrococcal nuclease (MNase) buffer (85 mM Tris-HCl, pH 7.5, 3 mM $MgCl_2$ and 2 mM $CaCl_2$) containing 0.3 M of sucrose and MNase (30 units of MNase per 2 million sperm; Roche Nuclease S7) was added to each tube. The tubes were immediately placed at 37°C for precisely 5 min. The MNase treatment was terminated by adding 2 mL of 0.5 M EDTA and leaving the tubes on ice for 10–20 min. The tubes were subsequently centrifuged at 17,000 × *g* for 10 min at room temperature to separate the debris and protamines (pellet) from the sheared chromatin (supernatant). Supernatants from the same sample were pooled into a 1.5-mL tube and 1× protease inhibitor (Roche) was added to the chromatin. The immunoprecipitation for H3K4me3 (Cell Signaling Technology) was carried out overnight at 4°C using Protein A Dynabeads (ThermoFisher Scientific). The mononucleosomal fraction (147 bp) was size selected using Agencourt XP AMPure beads (Beckman Coulter). Libraries were prepared using the Qiagen Ultralow Input Library Kit (Qiagen; #180495), as per the manufacturer's recommendations, and samples were sequenced using paired-end 100-bp reads with the NovaSeq 6000 platform ($n = 49$; Table S7).

ChIP-Seq Data Preprocessing

Raw reads were trimmed with the TrimGalore wrapper script around the sequence-grooming tool cutadapt (version 0.50; Felix Krueger) with the following quality trimming and filtering parameters: –length 50 -q 5 –stringency 1 -e 0.1`.[70] The trimmed reads were mapped onto the hg19/GRCh37 reference genome downloaded from University of California, Santa Cruz (UCSC) genome browser using Bowtie2 (version 2.3.5.1), as previously described.[27,63] We excluded reads that exhibited more than three mismatches. SAMtools (version 1.9)[71] was then used to convert SAM files and index BAM files. BigWig coverage tracks and binding heatmaps were generated from the aligned reads using deepTools2 (version 3.2.0).[72] The coverage was calculated as the number of reads extended to 150-bp fragment size per 25-bp bin and normalized using reads per kilobase per million mapped reads (RPKM) not located on the X chromosome.

H3K4me3 Differential Enrichment Analysis

For all bioinformatics analyses downstream of the preprocessing steps, R (version 3.6.0; R Development Core Team) was used coupled to the Bioconductor packages from the 3.9 release. We previously identified the H3K4me3 enrichment location in sperm using a high-quality reference human data set.[27] The exceptionally high sequencing depth of this data set allowed us to obtain a robust and high-resolution map of H3K4me3 peaks in sperm ($n = 50{,}117$ peaks). We computed the sum of read counts under peaks, confirmed the robustness of the signals in the present data set, and excluded peaks with low counts (<5 reads) in this study ($n = 48{,}499$ H3K4me3 peaks; Figure S4A,B). Differential binding analysis was conducted to identify regions exhibiting different levels of H3K4me3 binding in sperm of individuals with the highest level of *p*,*p*′-DDE serum levels (third tertile, ter3; *p*,*p*′-DDE >14,102 ng/mL) compared with individuals with the lowest level of *p*,*p*′-DDE serum levels (first tertile, ter1; *p*,*p*′-DDE <1,132 ng/mL). To do so we used the Bioconductor/R package edgeR (version 3.26.0)[73] using the negative binomial distribution and shrinkage estimates of the dispersions to model read counts. Regions with FDR[68] of <0.2 were defined as significantly differentially enriched regions (deH3K4me3) between the two *p*,*p*′-DDE serum level groups.

Dose–Response-Like Trend Analyses

For MCC-seq data sets, dose–response-like relationships were assessed via linear regressions conducted between the average percentage DNAme at DNAme gain or loss DMCs relative to $\log_{10}$ serum *p*,*p*′-DDE concentration in Greenlandic and South African populations, respectively (Figure 1C). Serum *p*,*p*′-DDE concentrations served as a surrogate measure of individual exposure to *p*,*p*′-DDE in both populations and is irrespective of the source of exposure (DDT via indoor residual spraying or diet). Analyses were conducted for each population individually. We plotted the average percentage of DNAme gain or loss across the significant DMCs against the individual serum *p*,*p*′-DDE concentration

for each population (Figure 1C). In each case, we observed linear relationships between the independent variable (average DNAme gain or loss) and the dependent variable ($\log_{10}$ DDE serum p,p'-DDE concentration). We fit a linear model for each of these relationships, verified that model assumptions were met, and found relationships between the two variables to be significant in each case ($p < 0.001$; Figure 1C). R^2, or the coefficient of determination (ranging from 0.2 to 0.5), provides the proportion of variance in average DNAme gain or loss that can be explained by the serum p,p'-DDE concentrations.

Given that ChIP-seq is an antibody-based technique, quantification of the data is less sensitive than for the MCC-seq and performing a linear regression on the data is unsuitable. We thus assessed for chromatin dose–response-like trends by assigning samples to p,p'-DDE serum level tertiles and comparing levels of sperm H3K4me3 enrichment at deH3K4me3 peaks across tertiles. For ChIP-seq data sets, dose–response-like analyses were thus assessed via Kolmogorov–Smirnov tests to compare the empirical cumulative distribution functions for regions with H3K4me3 gain or loss in sperm of South African men exposed to increasing categorical levels of p,p'-DDE, namely tertiles ter1, ter2, and ter3 (Figure 3F,G). We plotted the empirical cumulative distribution functions that provide the proportion of peaks with normalized counts below a certain value (x-axis). For peaks with H3K4me3 gain, we observed that ter1 p,p'-DDE serum levels stood out in terms of the fraction of peaks being with the lowest counts compared with ter2 and ter3 (Figure 3F). For peaks with H3K4me3 loss, we observed that ter2 was located in between ter1 and ter3, with ter3 exhibiting the fraction of peaks with the lowest counts (Figure 3G). In other words, the gains in H3K4me3 appeared to occur similarly for the ter2 and ter3 levels, with a modest increase from ter2 to ter3, whereas the losses in H3K4me3 were more pronounced as the level of p,p'-DDE serum levels gradually increased from ter1 to ter3.

Transposable Element and Putative Fetal Enhancer Annotations

For the transposable element analysis, hg19 RepeatMasker library was downloaded from http://www.repeatmasker.org/species/hg.html (hg19 - Feb 2009 - RepeatMasker open-4.0.5 - Repeat Library 20140131), and the genomic ranges for classes of transposable and other repetitive DNA elements were cataloged. Genomic ranges for putative tissue-specific fetal enhancers were retrieved from the Enhancer Atlas 2.0 (http://www.enhanceratlas.org).[74]

Classification of Transposable Elements by Percentage Divergence Score to Infer Transposable Element Age

The hg19 RepeatMasker library includes divergence scores for each transposable element. For DNAme analysis, long terminal repeat–endogenous retrovirus (LTR-ERV1) [enriched at merged DMCs (mDMCs) gaining and losing DNAme in Greenland and South African sperm] transposable elements described in the RepeatMasker library were sorted into percentage divergence quantiles (Figure S3G). For H3K4me3 analysis, LINE-1, LTR-ERV1-maltose operon regulator (MaLR) (enriched at H3K4me3 gaining regions in South African sperm), LINE-2, short interspersed nuclear element–mammalian-wide interspersed repeat (SINE-MIR), and SINE-Alu (enriched at H3K4me3 losing regions in South African sperm) transposable elements described in the RepeatMasker library were similarly sorted into percentage divergence quantiles (Figure S4G–K). Transposable elements belonging to the first quarter were characterized as young; second quarter, as mid-young; third quarter, as mid-old; and fourth quarter, as old.[75]

Enrichment and Gene Ontology Analyses on mDMCs and deH3K4me3 Peaks

The significance of functional Gene Ontology (GO) enrichment for mDMCs gaining or losing in DNAme and peaks with H3K4me3 gain or loss was estimated using a weighted Fisher test ("weight01" algorithm) as implemented in the R/Bioconductor topGO package (version 2.34.0).[76] All pathways presented in the Excel Tables were statistically significant ($p < 0.05$). To illustrate the GO results in Figures 2–5, we selected significant pathways from the Excel Tables based on common data trends, nonredundancy of pathways, and relevance to population and child health in Greenland and South Africa. Significant overlaps of mDMCs or deH3K4me3 with specific genomic location including genic annotation from R/Bioconductor annotatr (version 1.8.0),[77] overlapping specific families of transposable elements from the RepeatMasker library hg19 open-4.0.5, or putative enhancers from the Enhancer Atlas 2.0 were calculated using a permutation test framework implemented in the R/Bioconductor regioneR package (version 1.30.0).[78] Random regions (of the same size as the tested regions) were resampled from the background list of MCC-seq regions or H3K4me3 peaks. We asked whether regions that bear both changes in DNAme and H3K4me3 overlapped specific genomic location compared with regions bearing changes in only DNAme or H3K4me3 (Figure 5D–F). In that case, random regions were resampled from the list of mDMCs or deH3K4me3, respectively.

Identification of DNAme Sperm-to-Embryo Persistent Regions

Reduced representation bisulfite sequencing (RRBS) preimplantation embryo data sets were retrieved from Guo et al.[79] (GSE49828). Sperm DNAme CpGs were classified into low (<20%)-, intermediate (20%–80%)-, or high (>80%)-DNAme levels and compared with preimplantation embryo CpG DNAme levels. DNAme sperm-to-zygote persistent CpGs retained the same levels of DNAme from sperm to the zygote, whereas DNAme sperm-to-inner cell mass (ICM) persistent CpGs retained the same levels of DNAme from sperm through to the ICM. Low-, intermediate-, or high-DNAme persistent CpGs that spanned 500 bp were then merged to generate DNAme sperm-to-zygote and sperm-to-ICM persistent regions. We identified 18,562 low-, 2,132 intermediate-, and 5,745 high-DNAme sperm-to-zygote persistent regions and 14,235 low-, 488 intermediate-, and 694 high-DNAme sperm-to-ICM persistent regions. Significant overlaps of DNAme sperm-to-embryo persistent regions with specific genomic location (including genic annotation, overlapping specific families of transposable elements, or putative enhancers) were calculated using a permutation test framework implemented in the R/Bioconductor regioneR package.[78] Random regions (of the same size as the tested regions) were resampled from the background list of MCC-seq regions overlapping RRBS regions profiled by Guo et al.[79]

Preimplantation Embryo H3K4me3 Data Set Preprocessing

H3K4me3 ChIP-seq preimplantation embryo raw reads were retrieved from Xia et al.[80] (GSE124718). Reads were trimmed using Trimmomatic in paired-end mode (version 0.38).[61] The trimmed reads were mapped onto the hg19/GRCh37 reference genome downloaded from the UCSC genome browser using Bowtie2, as previously described.[27]

Preimplantation Embryo H3K4me3 Enrichment

Regions enriched for H3K4me3 in 4-cell embryos, 8-cell embryos, and ICM, were identified using the R/Bioconductor package csaw (version 1.18.0).[81] Reads with a mapping quality score of >20

were counted in 150-bp sliding windows for each library across the genome after exclusion of blacklisted regions.[64] To estimate the overall background signal, reads were also counted in 2,000-bp contiguous bins for each library across the genome. Regions enriched for H3K4me3 were identified by filtering windows with nonspecific background enrichment and by merging the contiguous 150-bp windows that were remaining. All parameters were optimized independently for each stage of preimplantation embryogenesis after visual assessment of tracks using Integrative Genome Viewer (IGV). Windows with a $\log_2$-fold change of >7 (for H3K4me3 4-cell embryo data), >18 (for H3K4me3 8-cell embryo data), or >6 (for H3K4me3 ICM data) were merged. Maximum peak size was set at 20,000 bp for all embryo stages. This yielded 42,630, 25,457, and 18,284 H3K4me3 peaks in 4-cell embryos, 8-cell embryos, and ICM, respectively (Figure S6).

Association between Promoter Sperm, Embryo H3K4me3 Counts, and Embryo Gene Expression

To identify promoters that were enriched for sperm, 4-cell, 8-cell, and ICM H3K4me3, we plotted the density distribution of the $\log_2$ (mean counts >8) and identified the cutoff value between low and high abundant signal based on the local minimum of the bimodal distribution (Figure S7D–G). Of note, 4-cell embryo H3K4me3 density distribution did not follow a bimodal trend; therefore, the cutoff value between low and high abundant signal was determined by visually assessing the distribution of counts against its density (Figure S7E). In each embryo stage, genes with an RPKM of >1 were considered expressed.

Identification of H3K4me3 Sperm-to-Embryo Persistent Peaks

To identify persistent sperm-to-4-cell H3K4me3 peaks, we overlapped sperm H3K4me3 peaks to 4-cell H3K4me3 peaks. To identify persistent sperm-to-ICM H3K4me3 peaks, we first overlapped persistent sperm-to-4-cell H3K4me3 peaks to 8-cell H3K4me3 peaks to generate persistent sperm-to-8-cell H3K4me3 peaks. Then, we overlapped persistent sperm-to-8-cell H3K4me3 peaks to ICM H3K4me3 peaks, which yielded sperm-to-ICM H3K4me3 peaks.

Data Visualization Tools

The R/Bioconductor package ggplot2 (version 3.3.2)[82] was used for data set visualization in all figures. Tracks visualization was done with the IGV browser (version 2.11.2).[83]

Results

p,p′-*DDE Serum Levels and Genetic Diversity in Greenlandic Inuit and South African VhaVenda Men*

Sperm samples from men of two geographically distinct populations exposed to DDT or its metabolite -DDE were used for this study to investigate the association between -DDE serum levels and DNAme in Greenlandic Inuit, and histone and DNA methylation in sperm from South African men (see the "Materials and Methods" section for participant details). To study the consequences of *p*,*p*′-DDE bioaccumulation in Northern populations on sperm DNAme and H3K4me3, we used 47 paired serum and semen samples from Greenlandic Inuit men of the INUENDO cohort (Figure S1A and Table S1).[50,84,85] The men ranged from 20 to 44 years of age (mean age = 31 y) and had serum levels of the DDT metabolite, *p*,*p*′-DDE, between 39.4 and 5,000 ng/mL lipid (Figure S1A and Table S1).

To investigate the relationship between direct exposure to DDT through indoor residual spraying ($n = 33$ participants from villages that use indoor residual spraying and $n = 17$ participants from villages that do not) and the sperm epigenome, we used 50 semen samples from a cohort of men who were recruited from 12 villages that were either sprayed or not in the Thulamela Local Municipality of the Vhembe District Municipality in the Limpopo Province of South Africa (Figure S1B and Table S2). The men were between 18 and 32 years of age (mean age = 25 y). Body burdens of *p*,*p*′-DDE ranged from 33 to 58,544 ng/mL of lipids with a mean of 32,000 ng/mL of lipids (Figure S1B and Table S2). Sample selection was based on sperm quality requirements for ChIP-seq (normal DNA fragmentation index and sperm count >10 million). In comparison with the Greenlandic Inuit men, the South African VhaVenda men had on average 12 times higher serum *p*,*p*′-DDE levels (Table S1).

To detect population stratifications and potential technical artifacts within our cohorts for the epigenome-wide *p*,*p*′-DDE association analyses, we assessed the genetic diversity of the Greenlandic and South African populations (see the "Materials and Methods" section). PCA on genotype profiles of Greenlandic Inuit and South African VhaVenda cohorts compared with human reference populations (including African, American, Asian, and European), showed unique population ancestries for both cohorts while demonstrating the population homogeneity within each cohort (Figure S1C). Consequently, the results indicate that the likelihood of ethnic subgroups being the source of observed epigenetic variation is low and, as such, we did not include the top principal components as covariates to adjust for population structure in our epigenomics analyses. To further support the lack of association between individual genetic makeup and *p*,*p*′-DDE serum level groups, we performed PCAs on the genotype profiles of Greenlandic and South African men and used different colors/shapes to distinguish individuals by tertiles of serum *p*,*p*′-DDE level (Figure S1D,E; see the "Materials and Methods" section). As indicated by this analysis, the genetic background of individual men did not cluster based on paired serum *p*,*p*′-DDE tertiles and was thus not correlated with their paired *p*,*p*′-DDE serum level groups (Tables S3 and S4).

Serum p,p′*-DDE Levels in Greenlandic and South African Men, and Association with Sperm CpG Methylation*

Sperm DNA methylation was assessed using MCC-seq (Table S2) targeting a select set of regions and providing sequencing-based information on millions of CpG sites. Specifically, we used the human 5-methylcytosine-capture sequencing sperm capture panel recently developed by our group.[59] This panel interrogates <3.18 million CpGs in the genome, including the >850,000 sites present on the Infinium MethylationEPIC BeadChip, an array-based technique targeting commonly assessed gene promoter/CpG island regions and enhancers, which is widely used in DNA methylation analyses. In addition, the human sperm capture panel targets ~1 million intermediate CpGs, which are environmentally sensitive sequences that demonstrate higher variability and possess intermediate levels of methylation (between 20% and 80%) in human sperm. The quantitative association between sperm DNAme and serum *p*,*p*′-DDE was assessed using a generalized linear regression model adjusted for potential confounders (smoking, age, and BMI). We applied a continuous analysis to identify DMCs with DNAme gain or loss in sperm that were associated with *p*,*p*′-DDE serum levels. Here, we defined a CpG as a DMC with DNAme gain if the DNA methylation increased with *p*,*p*′-DDE levels. Conversely, we defined a CpG as a DMC with DNAme loss if the DNA methylation decreased when serum *p*,*p*′-DDE increased. To this end, we identified 24,177 DMCs in Greenlandic sperm and 34,698 DMCs in South African sperm (Figure 1A; Excel Tables S1 and S2). In both populations, >75% of DMCs were found at intermediate CpGs (methylation levels

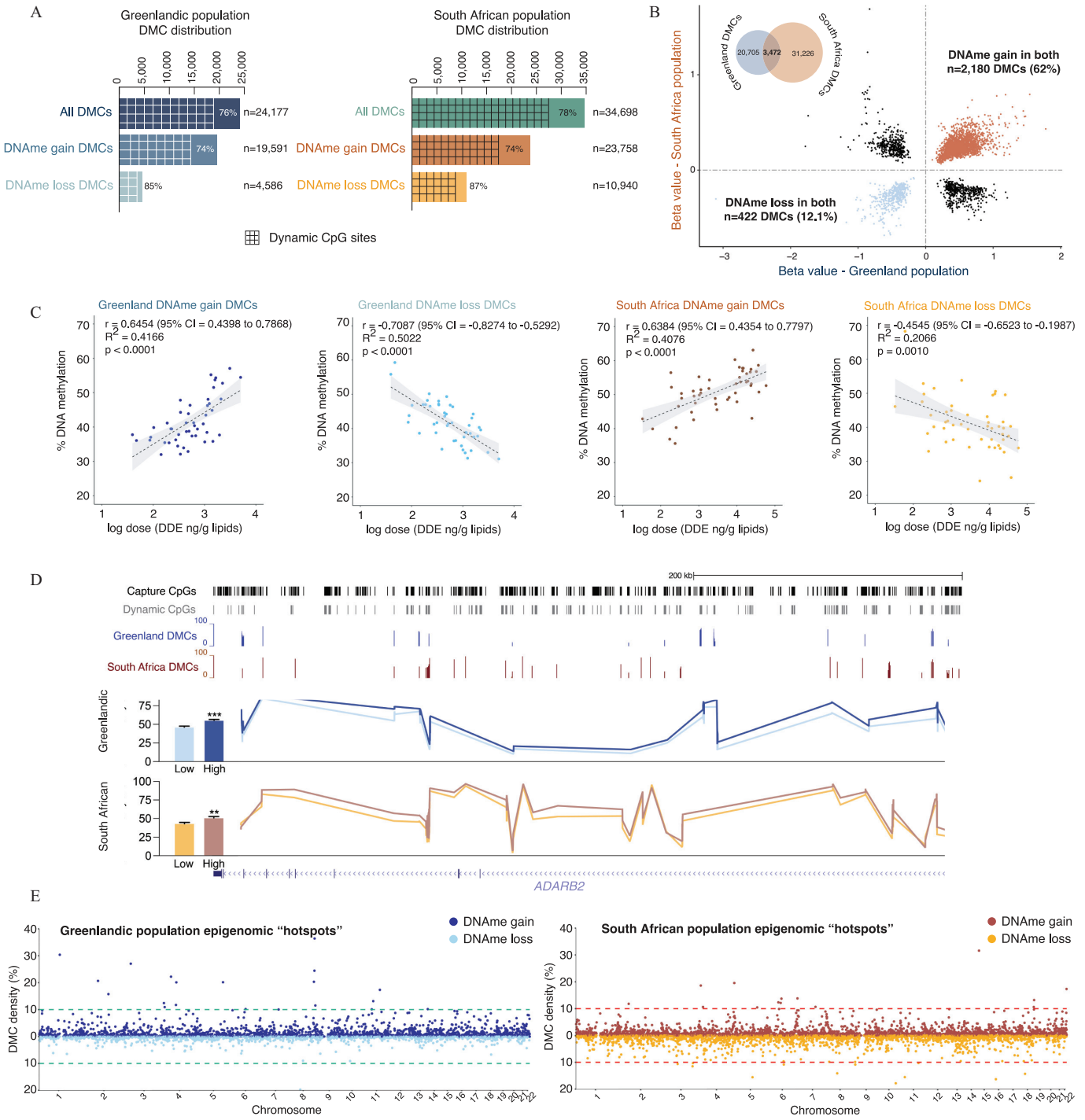

Figure 1. Association between *p*,*p*′-dichlorodiphenyldichloroethylene (*p*,*p*′-DDE) serum levels and human sperm DNA methylation (DNAme). (A) Number of differentially methylated CpGs (DMCs), DNA methylation (DNAme) gain DMCs, and DNAme loss DMCs, in the sperm of DDT-exposed Greenland or South African men. Number of intermediate DMCs (DNAme between 20% and 80%) are indicated by grids and percentages on the bar graphs. See Excel Tables S1 and S2. (B) Scatter plot of overlapping DMCs in Greenland and South African sperm (3,472 overlapping DMCs). Orange dots correspond to DMCs that gain DNAme in both Greenland and South African sperm (2,180 DMCs; 62% of total DMCs). Blue dots correspond to DMCs that lose DNAme in both Greenland and South African sperm (422 DMCs; 12.1% of total DMCs). (C) Average percentage DNAme at DNAme gain or loss DMCs in Greenland or South African sperm relative to $\log_{10}$ serum *p*,*p*′-dichlorodiphenyldichloroethylene (*p*,*p*′-DDE) concentration (in ng/mL) for each individual. Linear regression line is plotted in dashed black, and confidence interval is shaded in light gray. See Excel Tables S1 and S2. (D) Tracks at the *ADARB2* locus showing percentage DNAme levels in Greenland and South African sperm categorized based on low or high serum *p*,*p*′-DDE levels (Greenland: low in light blue <350 ng/mL, $n = 17$ and high in dark blue >900 ng/mL, $n = 18$; South Africa: low in yellow <1,200 ng/mL, $n = 19$ and high in brown >14,000 ng/mL, $n = 16$). All CpGs captured by MethylC-Capture-sequencing (MCC-seq) are represented in black, and intermediate CpGs are in gray. CpGs captured in the Greenland sperm data set are in blue, and CpGs captured in the South African sperm data set are in red. (E) Manhattan plots on hotspot analysis for Greenland (blue) or South African (orange) sperm. Cluster analysis was performed by calculating the ratio of DMCs with DNAme gain or loss over the total number of CpGs found within 1 Mb sliding windows over the genome; densities >10% (termed clusters) were extracted for further analysis. See Tables S5 and S6. Note: DDT, dichlorodiphenyltrichloroethane; *p*,*p*′-DDE, *p*,*p*′-dichlorodiphenyldichloroethylene.

between 20% and 80%; Figure 1A). DMCs with a DNAme gain were more abundant than DMCs with a DNAme loss in Greenlandic and South African sperm (19,591 and 23,758 DMCs with a DNAme gain and 4,586 and 10,940 DMCs with a DNAme loss, respectively; Fisher test $p < 0.0001$; Figure 1A). Interestingly, 3,472 DMCs overlapped between both populations (Fisher test $p < 0.0001$), with 62% of the overlapping DMCs gaining DNAme (2,180 DMCs; Fisher test $p < 0.0001$; Figure 1B). Previous studies have shown that DNAme across various cell types exhibit a dose–response trend to toxicants, such as arsenic and cigarettes.[86,87] To determine whether DNAme in sperm exhibited dose–response trends relative to *p,p′*-DDE, we investigated the linear relationship between *p,p′*-DDE from the men's serum and DMCs from their paired sperm samples (Figure 1C). For each individual in the Greenlandic or South African cohorts, we calculated the average DNA methylation levels across DMCs gaining or losing in DNAme (percentage DNAme from 0% to 100% methylated). Percentage DNAme level at DMCs in Greenlandic or South African sperm, and serum *p,p′*-DDE values, showed a linear dose–response trend (Figure 1C; $r = 0.6454$ for Greenland DNAme gain DMCs, -0.7087 for Greenland DNAme loss DMCs, 0.6384 for South Africa DNAme gain DMCs, -0.4545 for South Africa DNAme loss DMCs, and $p < 0.0001$) across all categories.

An example of dose–response difference in DNAme is shown for a gene of interest, adenosine deaminase, RNA-Specific, B2 (*ADARB2*) (Figure 1D). *ADARB2* is important for the regulation of RNA editing[88] and its perturbed function has been implicated in childhood cancer and neurological diseases.[89,90] A total of 162 and 160 DMCs were identified throughout the *ADARB2* locus in the Greenlandic and South African sperm, respectively, of which 32 DMCs were common between both populations (Figure 1D). The DMCs at the *ADARB2* locus were predominantly at intermediate CpGs and displayed intermediate levels of DNAme (129 and 132 intermediate DMCs in the Greenlandic and South African samples, respectively) (Figure 1D). We categorized samples based on low or high *p,p′*-DDE serum levels (Greenland: low <350 ng/mL, $n = 17$, and high >900 ng/mL, $n = 18$; South Africa: low <1,200 ng/mL, $n = 19$, and high >14,000 ng/mL, $n = 16$). Subjects with high *p,p′*-DDE serum levels had higher levels of sperm DNAme than those with low serum levels across the *ADARB2* locus (*t*-test, $p < 0.05$; Figure 1D).

As we previously reported for age-related alterations in sperm DNAme, DMCs can be found densely clustered in some areas of the genome, suggesting the possibility of genomic hotspots for environmental exposures.[91] We identified several chromosomal regions of high DMC density (>10%; termed clusters), in the sperm of both populations that were not confounded by genetic variations (Figure 1E; Tables S5 and S6). A total of 21 clusters of differential methylation were identified in the Greenlandic sperm (20 with DNAme gain, 1 with DNAme loss; Table S5), and 26 in the South African sperm (14 with DNAme gain, 12 with DNAme loss; Table S6). Two DNAme gain and 1 DNAme loss clusters were common between both cohorts. In the South African population, a cluster containing 443 DMCs overlapped to the imprinting control small nuclear ribonucleoprotein polypeptide N (SNRPN) region, which is implicated Prader Willi and Angelman syndromes[92] (Table S6). Many DMCs were also annotated to the small nucleolar RNAs (SNORDs) family of genes, which have been shown to have altered DNAme in sperm samples from the Early Autism Risk Longitudinal Investigation cohort[93] (Tables S5 and S6).

Functional Assessment of Differentially Methylated Regions in Greenlandic and South African Sperm

To gain functional insight into how *p,p′*-DDE-associated alterations in sperm DNAme may impact sperm function and embryo development, mDMCs were identified and their relationship to genome function and embryo development was explored. An mDMC was defined as a region with two or more DMCs with a consistent DNAme gain or loss and spanning a maximum of 500 bp (mDMCs; Figure 2A; Excel Tables S3 and S4). This merging of DMCs yielded 6,787 mDMCs in Greenlandic sperm and 12,849 mDMCs in South African sperm (Figure S2; Excel Tables S3 and S4). Overall, 1,187 mDMCs in Greenland sperm overlapped mDMCs in South African sperm, and 1,189 mDMCs in South African sperm overlapped mDMCs in Greenlandic sperm (Figure 2B). In line with the DMC analysis, mDMCs with increased DNAme were more abundant than mDMCs, with decreased DNAme in relationship to *p,p′*-DDE serum levels for both populations (4,928 DNAme gain mDMC in Greenlandic sperm and 8,259 DNAme gain mDMC in South African sperm; Figure 2C). mDMCs were predominantly enriched at >10 kb from a TSS in intergenic space for both populations (Fisher test $p < 0.0001$; Figure 2D; Figure S2A). GO analysis on genes that overlapped a sperm mDMC revealed significantly enriched pathways that were common to both populations, highlighting the consistency of *p,p′*-DDE-associated effects on the sperm epigenome across different populations. Common identified pathways were those involved in the development and function of the nervous system, such as axon guidance and regulation of synapse assembly and neurogenesis (Figure 2E–H; Excel Tables S5–S8). These pathways included genes *DVL1*, *ROBO1*, and *TRIO*. Notably, all are implicated in neurodevelopmental disorders, such as autism and dyslexia, which have been reported as adverse outcomes from developmental exposures to EDCs.[94,95] Developmental processes were also identified and included common genes in both populations, such as *PARD3* (neural tube defects), *MYLK* (aorta development), and *HOXB2* (skeletal patterning) (Figure 2E–H; Excel Tables S5–S8). Other significantly enriched pathways common to both populations and relevant to documented complex diseases in DDT- and DDE-exposed populations included heart morphogenesis and immune system function (Excel Table S5–S8).[96–99]

Overlap between Sperm-to-Embryo Persistent DNAme Regions and Sperm mDMCs Associated with p,p′*-DDE Serum Levels*

During embryo development, a DNAme reprogramming wave occurs from the zygote to the blastocyst stage that is characterized by a loss of gamete-specific DNA methylation patterns inherited from the sperm or the oocyte.[100] By the blastocyst stage, 20% of CpGs still retain gamete-inherited DNA methylation in mice[101] and humans.[102] Imprinted regions and certain classes of transposable elements have been shown to resist DNA methylation reprogramming, and their silencing is critical during preimplantation embryogenesis.[103,104] To this end, we aimed to investigate whether *p,p′*-DDE serum levels were associated with sperm mDMCs that are predicted to persist in the embryo and could, therefore, potentially be implicated in epigenetic inheritance. To do so, we performed a stepwise analysis initially focused on identifying sperm-transmitted DNAme that potentially escapes reprogramming. First, we identified CpGs that retained the same level of DNAme from sperm to the zygote and from sperm to the ICM of the blastocyst using our data and an existing data set.[79] CpGs that retained sperm-specific DNAme patterns in the embryo may correspond to CpGs that escape reprogramming (Figure S3 and Excel Tables S9–S14). By this approach we identified *a*) CpGs with DNAme that was consistent from sperm to the zygote, corresponding to CpGs that could influence the first wave of zygotic gene expression (Figure S3A,B); and *b*) CpGs with DNAme that persisted across preimplantation embryogenesis until the first lineage specification in the blastocyst (Figure S3C,D).

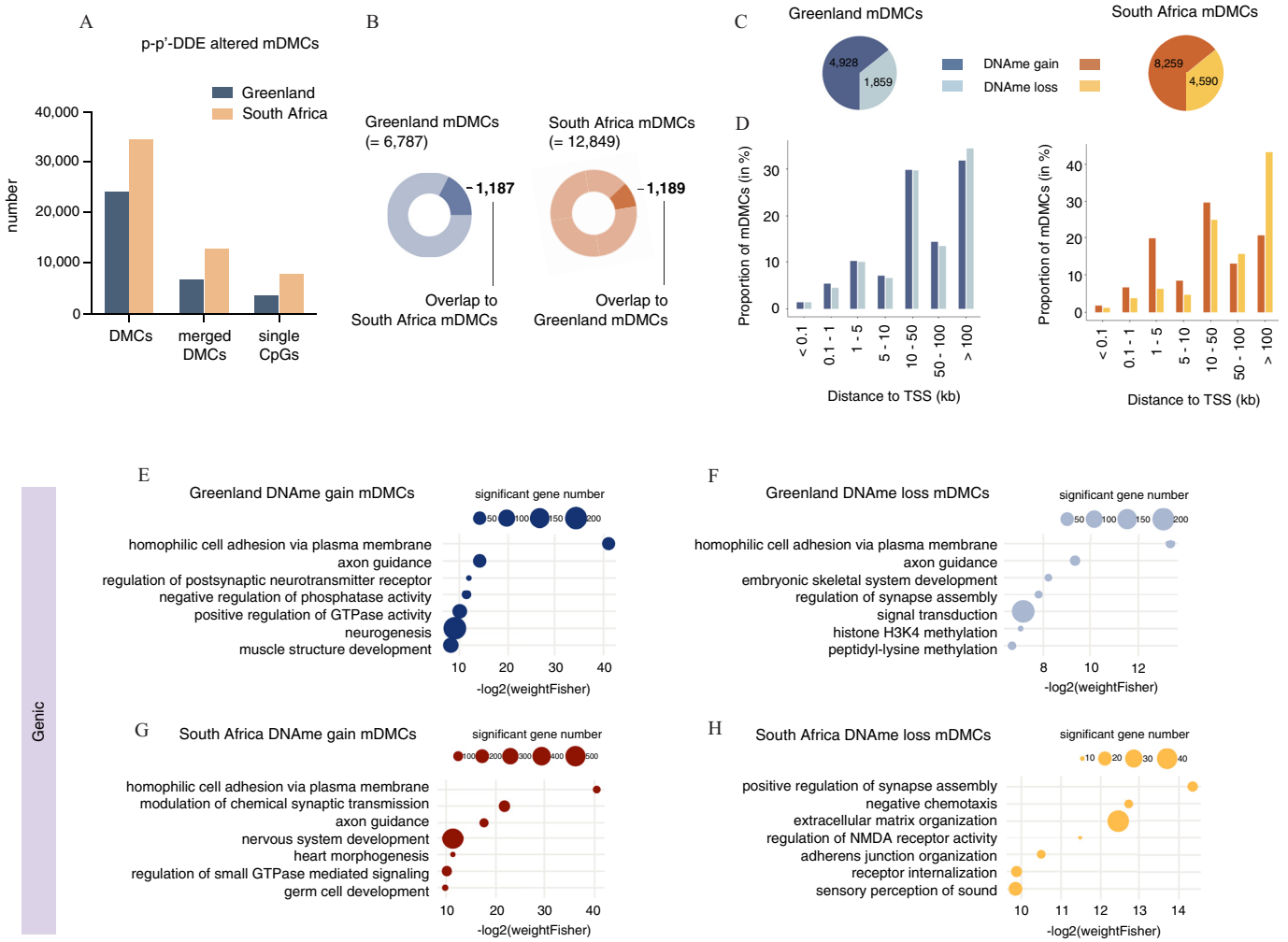

Figure 2. Sperm differentially methylated regions association with TEs and regions that retain DNA methylation (DNAme) during embryogenesis. (A) Number of differentially methylated CpGs (DMCs), merged DMCs (mDMCs), and single CpGs in Greenland (dark blue) and South African (light orange) sperm. mDMCs were called by merging DMCs separated by a maximum distance of 500 bp. See Excel Tables S3 and S4. (B) Proportion of Greenlandic sperm mDMCs (6,787) that overlap South African sperm mDMCs (overlap = 1,187) and South African sperm mDMCs (overlap = 12,849) that overlap Greenlandic sperm mDMCs (overlap = 1,189). See Excel Tables S3 and S4. (C) Number of DNAme gain or loss mDMCs in Greenlandic or South African sperm. See Excel Tables S3 and S4. (D) Genomic distribution of mDMCs with DNAme gain or loss in Greenlandic or South African sperm relative to the transcriptional start site (TSS). (E–H) Selected significant pathways from GO analysis on genes at an mDMC with DNAme gain or loss in Greenlandic or South African sperm (weighed Fisher $p < 0.05$). Size of circle corresponds to the number of genes from a significant pathway that overlap an mDMC. See Excel Tables S5–S8. (I) Enrichment of DNAme gain (dark blue for Greenlandic; dark red for South African) or loss (light blue for Greenlandic; bright yellow for South African) mDMCs in Greenlandic (white hashes) or South African (white dots) sperm at identified DNAme embryonic persistent regions [see Figure S3, primordial germ cell (PGC) escapees, and transposable element annotations (RepeatMasker hg19 library 20140131)]. Positive enrichments were determined by z-scores using the Bioconductor package regioneR. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions (of the same size) resampled from the targeted MCC-seq regions. (J) Distribution of DNAme gain (dark blue for Greenlandic; dark red for South African) or loss (light blue for Greenlandic; bright yellow for South African) mDMCs in Greenlandic (white hashes) or South African (white dots) sperm relative to age quarters of LTR-ERV1 TEs (significantly enriched in the four mDMC categories; Figure 2H). Age of LTR-ERV1 TEs was determined by partitioning the TEs' percentage divergence scores into quarters, where first quarter = low percentage divergence and young LTR-ERV1; second quarter = mid-low percentage divergence and mid-young LTR-ERV1; third quarter = mid-high percentage divergence and mid-old LTR-ERV1; fourth quarter = high percentage divergence and old LTR-ERV1 (see Figure S3B). (K) Enrichment of DNAme gain (dark blue for Greenlandic; dark red for South African) or loss (light blue for Greenlandic; bright yellow for South African) mDMCs in Greenland (white hashes) or South African (white dots) sperm at putative tissue-specific enhancer classes. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions (of the same size) resampled from the targeted MCC-seq regions. Note: dyn, dynamic; GO, Gene Ontology; GTPase, guanosine triphosphatase; hESC, human embryonic stem cell; ICM, inner cell mass; MCC-seq, MethylC-Capture-sequencing; NMDA, NMDA, *N*-methyl-D-aspartate; *p*,*p*′-DDE, *p*,*p*′-dichlorodiphenyldichloroethylene; TEs, transposable elements.

We then merged two or more low-, intermediate-, or high-DNAme persistent CpGs spanning a maximum of 500 bp to classify DNAme sperm-to-zygote and sperm-to-ICM persistent regions (Figure S3A–D and Excel Tables S9–S14). This yielded 18,562 low-, 2,132 intermediate-, and 5,745 high-DNAme sperm-to-zygote persistent regions (Figure S3B and Excel Tables S9–S11) and 14,235 low-, 488 intermediate-, and 694 high-DNAme sperm-to-ICM persistent regions (Figure S3D; Excel Tables S12–S14). Of note, the sperm-to-ICM persistent regions consisted of the subset of sperm-to-zygote persistent regions that retain their DNAme levels throughout preimplantation embryogenesis. Intermediate- and high-DNAme persistent regions were enriched at CpG shores, exons, CpG shelves, and introns (Figure S3E; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations). Intermediate-DNAme persistent regions were more enriched in intergenic space than high-DNAme persistent regions (Figure S3E). Low-DNAme

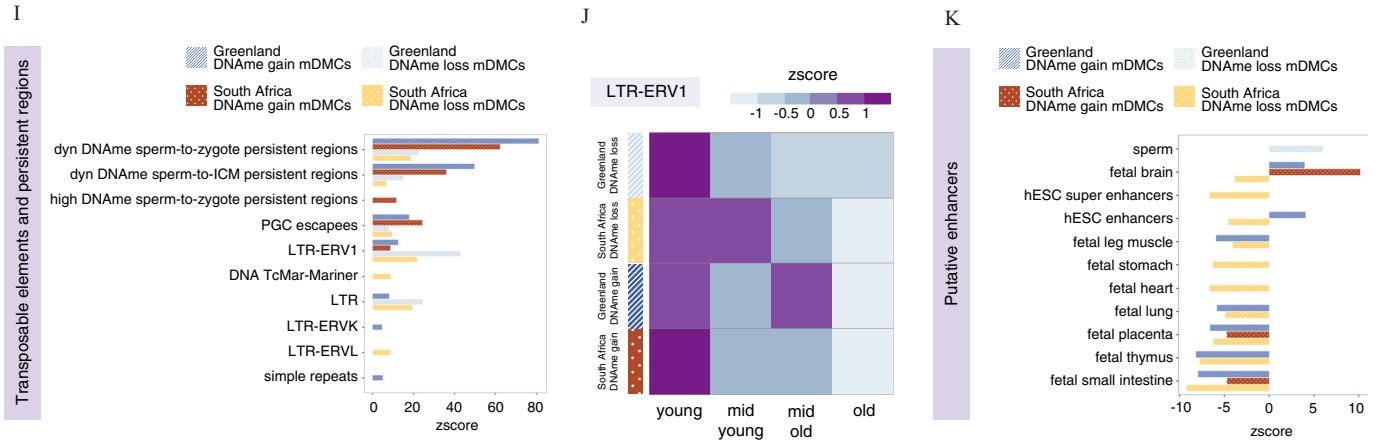

Figure 2. (*Continued.*)

persistent regions were not enriched at any genic or CpG annotations (Figure S3E). Remarkably, a significant proportion of sperm-to-zygote and sperm-to-ICM persistent regions annotated to transposable elements (TEs), with the LTR-ERV1, LINE-2, and LTR-ERVL-MaLR TE families being overrepresented across intermediate- and high-DNAme persistent regions (Figure S3F; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations). Analysis of functional elements also revealed an enrichment of human embryonic stem cell (hESC) enhancers and super-enhancers at intermediate- and high-DNAme persistent regions (Figure S3F; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations).[105]

Last, to identify p,p'-DDE-associated mDMCs that may be epigenetically inherited, we determined whether mDMCs co-localized to DNAme persistent mDMCs (Figure 2I). Intermediate-DNAme sperm-to-zygote and sperm-to-ICM persistent mDMCs, as well as previously described primordial germ cell escapees,[106] were enriched across mDMCs, with DNAme gain or loss in Greenlandic and South African sperm (Figure 2I; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations), supporting the possibility that intermediate-DNAme persistent mDMCs impacted by p,p'-DDE may escape reprogramming in the embryo. mDMCs were also enriched for certain TE families, with an overrepresentation of the LTR-ERV1 family at mDMCs, with DNAme gain and loss in Greenlandic and South African sperm (Figure 2I). The age of a TE can be inferred from the level of divergence observed between the TE sequence and the canonical full-length element sequence.[107] Indeed, this value reflects the duration the TE has been integrated into the genome.[107] Importantly, more newly integrated young TEs are typically more active than old TEs.[75] Consequently, epigenetic changes at more recently integrated TEs may elicit more dramatic effects on gene expression. We categorized all LTR-ERV1 TEs of the human genome into four quarters based on their percentage divergence scores (Figure S3G), Interestingly, young LTR-ERV1s were the predominant age quarter that overlapped mDMCs (Figure 2J; Figure S3H). Because the majority of mDMCs were found in intergenic space, we intersected them to putative fetal tissue enhancers (Figure 2K). mDMCs with DNAme gain in Greenland and South African sperm were overrepresented at fetal brain putative enhancers, whereas mDMCs with DNAme loss in Greenland sperm were enriched at sperm putative enhancers (Figure 2K; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations).[74] All other studied classes of putative fetal enhancers were underrepresented at mDMCs, highlighting the functional specificity of p,p'-DDE–associated mDMCs in fertility and neurodevelopment (Figure 2K). Taken together the analysis of MCC-seq data suggests that intermediate-DNAme CpGs in sperm are more sensitive to alterations associated with p,p'-DDE serum levels and may escape epigenetic reprogramming in the embryo at young TEs.

Association of p,p′-DDE Serum Levels and Sperm H3K4me3 Enrichment at Developmental Loci

Next, we set out to investigate whether p,p'-DDE serum levels were associated with differential enrichment of H3K4me3 (deH3K4me3) in human sperm and if changes occurred at genomic regions associated with the observed population health abnormalities in DDT-exposed regions. To do so, we performed 100-bp paired-end sequencing on high-quality H3K4me3 chromatin immunoprecipitation libraries from the sperm of each participant of the South African VhaVenda cohort. On average, 98% of reads aligned to the human genome, yielding ~ 115 million reads per sperm sample, of which 88 million were uniquely mapped (Table S7). We identified 48,499 H3K4me3 peaks in South African sperm samples (Figure S4A,B). Sperm samples were categorized based on tertiles of their corresponding serum p,p'-DDE levels as low (ter1), intermediate (ter2), or high (ter3) levels (Figure S1B), and a differential enrichment analysis was performed to identify differential H3K4me3 peaks between categories of p,p'-DDE levels.[81] Comparison between high (ter3) vs. low (ter1) p,p'-DDE levels identified 1,865 peaks with differentially enriched H3K4me3 (deH3K4me3; FDR <0.2) of which 851 peaks gained H3K4me3 enrichment and 1,014 peaks lost H3K4me3 enrichment in the high (ter3) p,p'-DDE serum level group (Figure 3A; Excel Table S15). Peaks gaining H3K4me3 were enriched in intergenic regions (Figure S4C; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations) and were predominantly >10 kb from the TSS (Figure 3B). Conversely, peaks with H3K4me3 loss were primarily genic occurring at promoters, exons, and at CpG-rich regions (Figure S4D; z-score >0; $p < 0.0001$; $n = 10{,}000$ permutations) and were mostly <5 kb from the TSS (Figure 3C). This suggests that peaks with H3K4me3 loss were more likely to directly impact gene expression in the next generation (Figure 3C).

In rodents, H3K9ac was identified as dose responsive to traffic air pollution in lung and blood cells.[108] Whether sperm chromatin in humans shows a dose–response-like trend to any toxicant is unknown. To investigate this possibility, we first depicted the average H3K4me3 enrichment in deH3K4me3 peaks across the three levels of p,p'-DDE serum levels. Consistent with a dose–response effect, H3K4me3 enrichment was most apparent for the extreme levels of serum p,p'-DDE at deH3K4me3 (ter3 vs. ter1; Figure 3D–G; Figure S1B). Indeed, testing altered

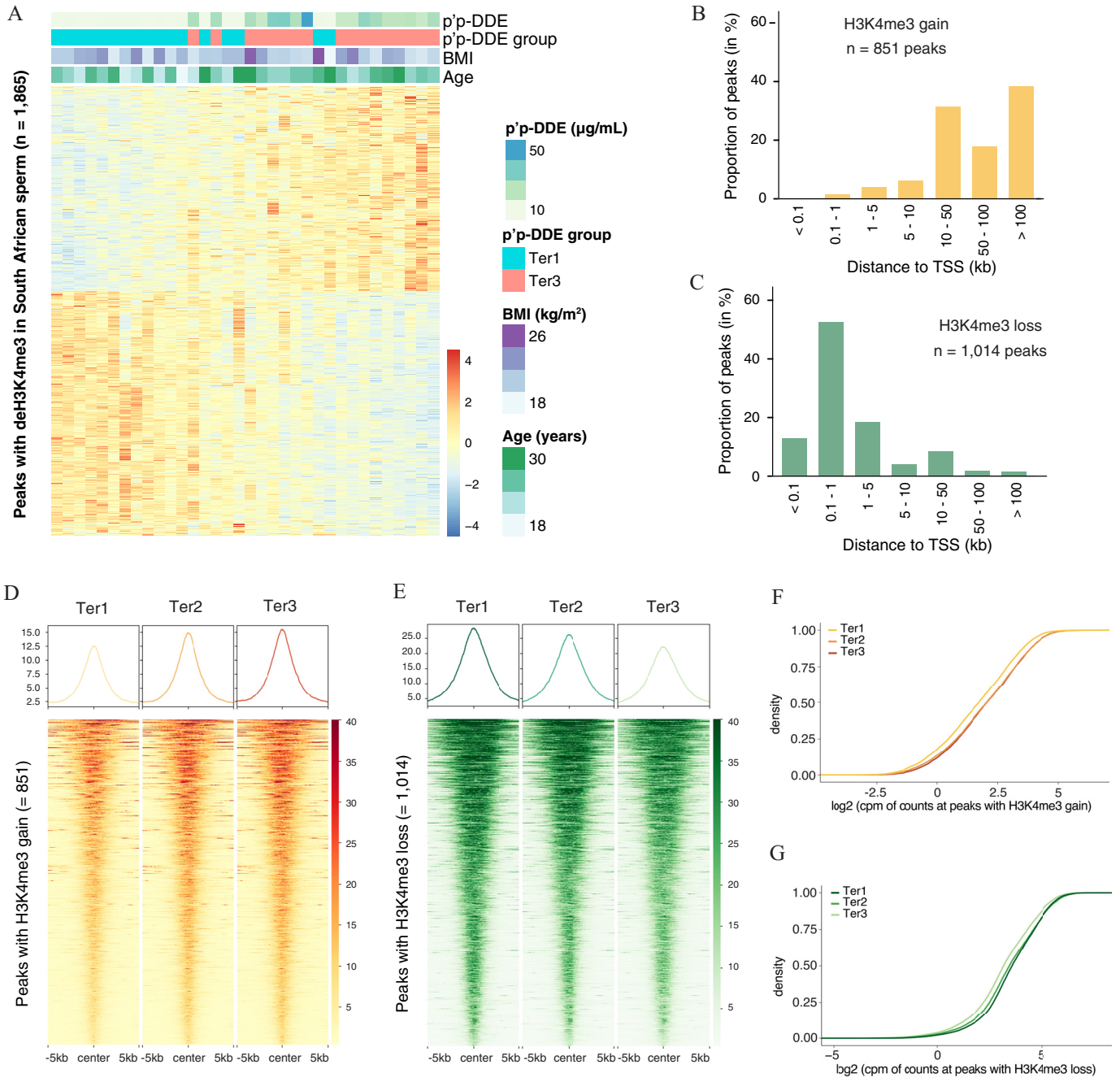

Figure 3. Histone 3 lysine 4 trimethylation (H3K4me3) enrichment in sperm of South African men with different *p*,*p*′-DDE serum levels. (A) Heatmap of normalized H3K4me3 counts at the 1,865 peaks with differentially enriched H3K4me3 (deH3K4me3; 851 peaks with H3K4me3 gain and 1,014 peaks with H3K4me3 loss; FDR <0.2) in sperm from South African men with low (ter1) or high (ter3) serum *p*,*p*-DDE levels. Serum *p*,*p*-DDE concentration, tertile, body mass index (BMI), and age of participants is indicated by colored boxes above the heatmap. (B) Genomic distribution of peaks with H3K4me3 gain in South African sperm relative to the TSS. (C) Genomic distribution of peaks with H3K4me3 loss in South African sperm relative to the TSS. (D–E) H3K4me3 signal intensity heatmaps at ±5 kb the center of the peaks with H3K4me3 gain in South African sperm (851 peaks) relative to the three *p*,*p*-DDE tertile levels. (E) H3K4me3 signal intensity heatmaps at ±5 kb the center of the peaks with H3K4me3 loss in South African sperm (1,014 peaks) relative to the three *p*,*p*-DDE tertile levels. (F) Estimator of the cumulative distribution function (ECDF) plot for $\log_2$ cpm of H3K4me3 counts at regions with H3K4me3 gain in ter1 (light yellow), ter2 (medium orange), and ter3 (dark red) sperm samples. Dose–response trends were assessed by via Kolmogorov–Smirnov tests with $p < 0.0001$ for ter1 vs. ter2 at H3K4me3 gain regions, $p < 0.05$. (G) ECDF plot for $\log_2$ cpm of H3K4me3 counts at regions with H3K4me3 gain in ter1 (dark green), ter2 (medium green), and ter3 (light green) sperm samples. Dose–response trends were assessed by via Kolmogorov–Smirnov tests with $p < 0.0001$ for ter1 vs. ter2 at H3K4me3 gain regions, $p < 0.0001$. (H) Representative Integrative Genome Viewer (IGV) tracks of peak with H3K4me3 gain in South African sperm at the *FSIP1* genic region. (I) Representative IGV tracks of peak with H3K4me3 loss in South African sperm at the *BRD1* promoter. Data for (A–I) are reported in Excel Table S15. Selected significant pathways for (J–K) are from GO analysis on genes with H3K4me3 gain (J) or loss (K) in South African sperm (weighed Fisher $p < 0.05$). Size of circle corresponds to the number of genes from a significant pathway that overlap a peak with H3K4me3 gain. Shade intensity of the circle indicates $-\log_2$ (weightFisher) value of significant pathway. See Excel Tables S16 and S17. Note: cpm, counts per minute; GO, Gene Ontology; *p*,*p*′-DDE, *p*,*p*′-dichlorodiphenyldichloroethylene; ter, tertile; TSS, transcriptional start site.

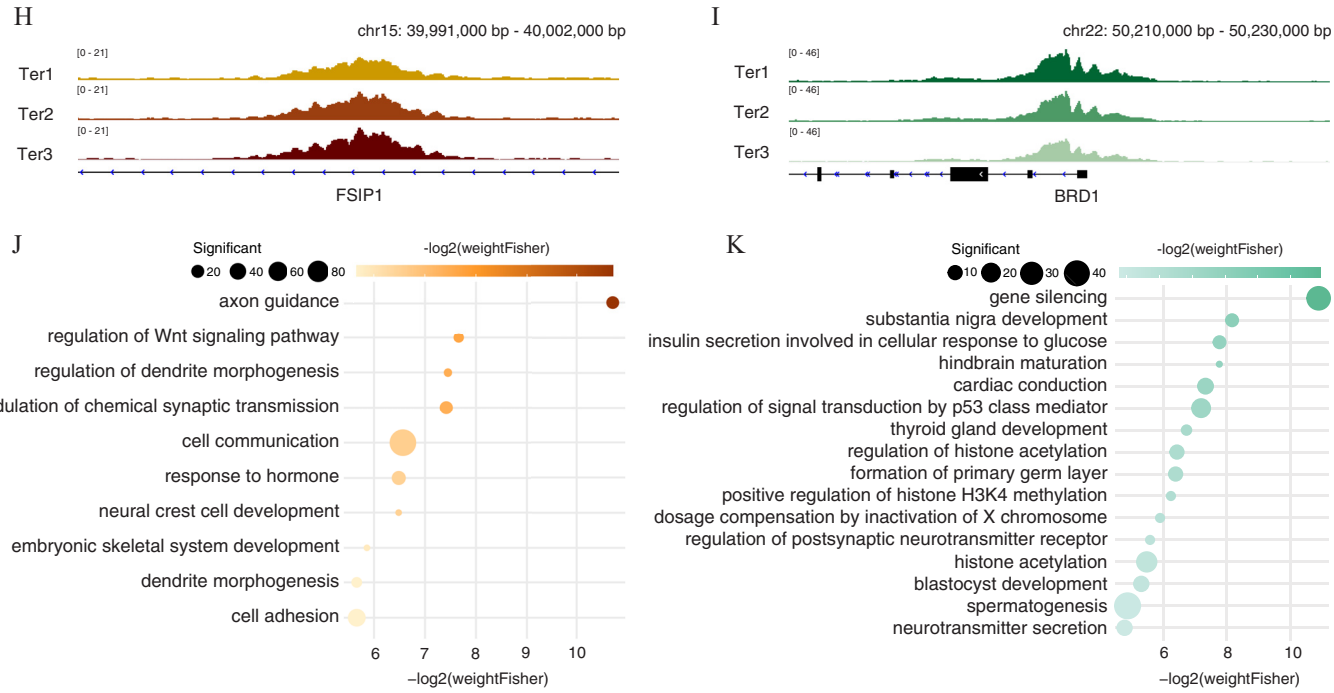

Figure 3. (*Continued.*)

regions for a dose-like response across all *p*,*p*′-DDE serum level categories revealed significant effects for regions that gained H3K4me3 (ter1 vs. ter2; Kolmogorov–Smirnov $p < 0.0001$ and ter2 vs. ter3; Kolmogorov–Smirnov $p < 0.05$; Figure 3F; Figure S4E), as well as for regions that lost H3K4me3 (ter1 vs. ter2 and ter2 vs. ter3; Kolmogorov–Smirnov $p < 0.0001$; Figure 3G; Figure S4F). Examples of peaks with a dose–response-like trends across the three tertiles were *FSIP1* gaining in H3K4me3 (important for acrosomal reaction and sperm flagellation; Figure 3H),[109] and *BRD1* losing in H3K4me3 chromatin interacting protein involved in brain development (Figure 3I).[110]

Co-Localization of Peaks with Sperm H3K4me3 Gain or Loss to TEs and Gene Regulatory Regions

GO analyses on genes overlapping peaks with H3K4me3 gains revealed enrichment at genes involved in neural development and cell signaling, including hormone responses (Figure 3J; Excel Table S16). Peaks with H3K4me3 loss were associated with genes and tissues implicated in metabolism, development, spermatogenesis, chromatin remodeling, and the endocrine system (Figure 3K; Excel Table S17). Genes of interest for their roles in development and disease included *NRP1*, *SOX6*, and *SEMA5A* at peaks with H3K4me3 gain and *KDM6B*, *SOX17*, and *BRCA1* at peaks with H3K4me3 loss. Sperm-transmitted intergenerational effects of DDT exposure may be mediated by alterations of H3K4me3 at TEs in sperm. In fact, TEs that escape epigenetic reprogramming and that are regulated by chromatin features, including H3K4me3, may behave as enhancers and/or promoters in the embryo.[111–115] We identified a high degree of specificity in the TE families that overlapped with peaks with gain or loss of H3K4me3 in sperm (Figure 4A,B). Peaks with H3K4me3 gain were significantly enriched at LINE-1 elements, DNA repeats, and LTR-ERVL-MaLR (z-score >0, $p < 0.0001$, $n = 10{,}000$ permutations; Figure 4A). Conversely, TE and repeat families overrepresented at peaks with H3K4me3 loss included low complexity repeats, SINE-Alu [such as SINE-(full-length poly(A) and mRNA) FLAM-A, SINE-FLAM-C, and SINE-FRAM], SINE-MIR, and LINE-2 (z-score >0, $p < 0.0001$, $n = 10{,}000$ permutations; Figure 4B).

Complementing the DNAme TE analysis, we selected the families of TEs that were overrepresented in peaks with gain or loss of H3K4me3 and categorized them into young, mid-young, mid-old, and old based on their percentage divergence scores (see the "Materials and Methods" section; Figure S4G–K). Regions that gained H3K4me3 in sperm at LINE-1 and LTR ERVL-MaLR transposable elements, predominantly overlapped older TEs and were located >10 kb from the TSS (Figure 4C,E). In contrast, LINE-2, SINE-MIR, and SINE-Alu that intersected peaks that showed reduced H3K4me3 in sperm were classified as younger TEs and mostly enriched at <5 kb from the TSS (Figure 4D,F). Next, we performed a GO analysis for young LINE-2, SINE-MIR, and SINE-Alu elements that overlapped a gene with H3K4me3 loss in sperm (Figure 4G; Excel Tables S18–S20). The genes overlapping young LINE-2 and SINE-MIR TEs were involved in noncoding and coding RNA processing, metabolism, and basic cellular processes (Figure 4G; Excel Tables S18–S20). Young SINE-Alu TEs overlapping a gene with H3K4me3 loss in sperm were enriched for genes involved in craniofacial and skeletal development (Figure 4G; Excel Table S20). We next intersected peaks with H3K4me3 gain or loss to putative sperm and fetal enhancers (Figure 4H). Peaks with H3K4me3 gain were overrepresented at sperm putative enhancers but not at other enhancer classes. Conversely, peaks with H3K4me3 loss were enriched across a broad number of embryo development putative enhancers and underrepresented at sperm putative enhancers (Figure 4H). Examples of *p*,*p*′-DDE–associated deH3K4me3 peaks in South African sperm that are enriched for TEs and putative enhancers can be found in Figure S5.

p,p′-*DDE Serum Levels Association with Sperm H3K4me3 and Regions That Overlap Preimplantation Embryo H3K4me3*

To assess the possible involvement of sperm H3K4me3 in epigenetic transmission of paternal environmental exposure to DDT,

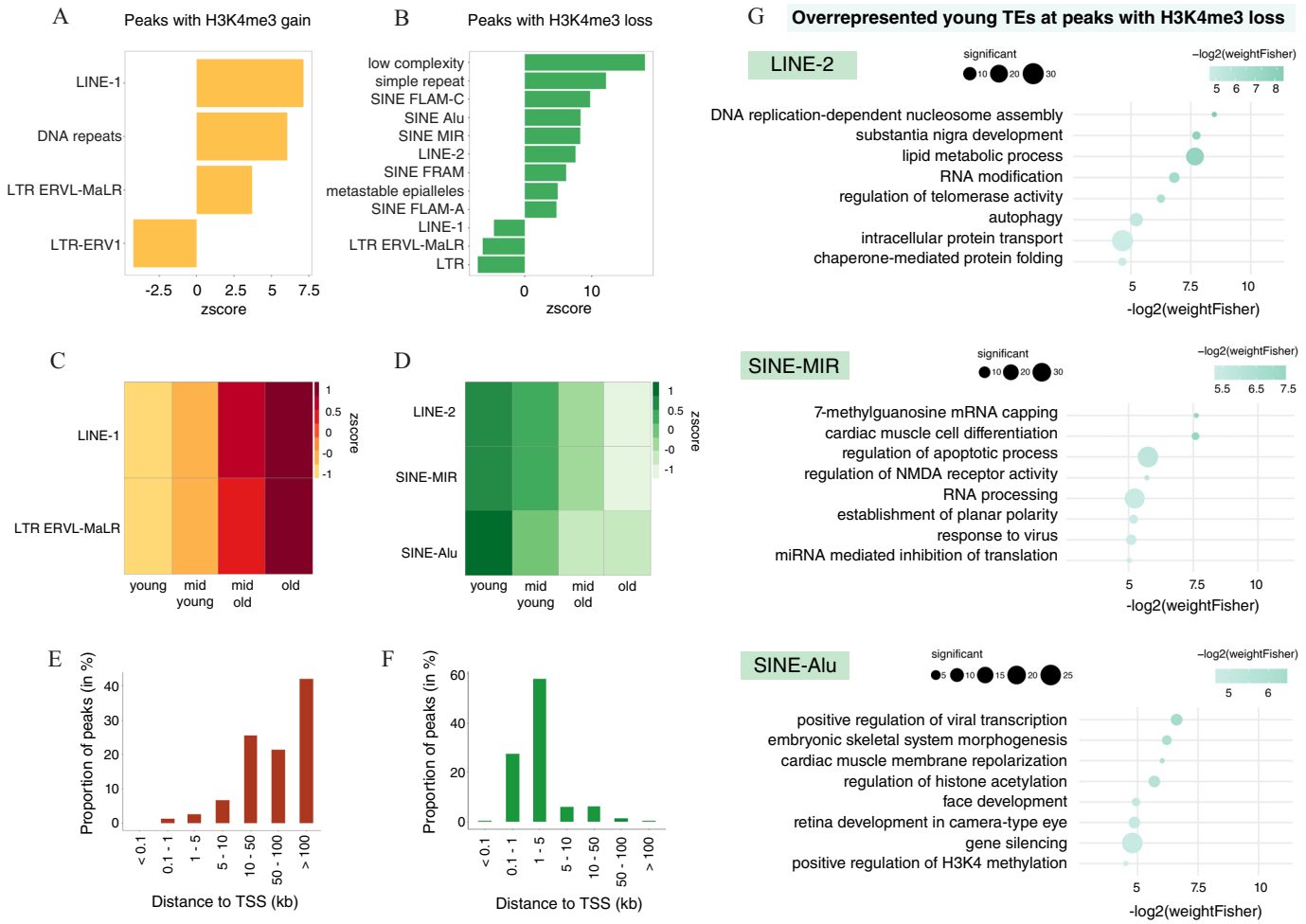

Figure 4. Overlap between peaks with H3K4me3 loss in South African sperm, TEs, and regions that retain H3K4me3 in the preimplantation embryo. (A) Enrichment for peaks with H3K4me3 gain in South African sperm at transposable element annotations (RepeatMasker hg19 library 20140131). Positive and negative enrichments were determined by *z*-scores using the Bioconductor package regioneR. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions (of the same size) resampled from sperm H3K4me3 peaks. (B) Enrichment for peaks with H3K4me3 loss in South African sperm at transposable element annotations (RepeatMasker hg19 library 20140131). Positive and negative enrichments were determined by *z*-scores using the Bioconductor package regioneR. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions (of the same size) resampled from sperm H3K4me3 peaks. (C) Distribution for peaks with H3K4me3 gain in South African sperm relative to age quarters of LINE-1 and LTR ERV-MaLR transposable element classes [significantly enriched for peaks with H3K4me3 gain (A)]. Age of TEs was determined by partitioning the class percentage divergence score in quarters (see Figure S4G,H). (D) Distribution for peaks with H3K4me3 loss in South African sperm relative to age quarters of LINE-2, SINE-MIR, and SINE-Alu transposable element classes [significantly enriched for peaks with H3K4me3 loss (B)]. Age of TEs was determined by partitioning the class percentage divergence score in quarters (see Figure S4I–K). (E) Genomic distribution of peaks with H3K4me3 gain in South African sperm that overlap an old (4th quarter) LINE-1 (139) or LTR ERVL-MaLR (74) transposable element, relative to the TSS. (F) Genomic distribution of peaks with H3K4me3 loss in South African sperm that overlap a young (1st quarter) LINE-2 (187), SINE-MIR (267), or SINE-Alu (507) transposable element, relative to the TSS. (G) Selected significant pathways from GO analysis on promoter peaks with H3K4me3 loss in sperm that overlap a young (1st quarter) LINE-2 (324 promoters), SINE-MIR (442 promoters), or SINE-Alu (809 promoters) TEs at (weighed Fisher $p < 0.05$). Size of dots corresponds to the number of genes from a significant pathway that overlap a peak with H3K4me3 loss. Color of the dots indicates $-\log_2$ (weightFisher) value of significant pathway. See Excel Tables S18–S20. (H) Enrichment for peaks with H3K4me3 gain (yellow) or loss (green) in South African sperm at tissue-specific putative enhancer annotations. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions (of the same size) resampled from sperm H3K4me3 peaks. Data are reported in Excel Table S19. (I) Sperm and preimplantation embryo (4-cell, 8-cell, ICM) H3K4me3 signal intensity heatmaps at ± 5 kb the center of the peaks with H3K4me3 gain in South African sperm (851 peaks). Data are reported in Excel Table S20. (J) Sperm and preimplantation embryo (4-cell, 8-cell, ICM) H3K4me3 signal intensity heatmaps at ± 5 kb the center of the peaks with H3K4me3 loss in South African sperm (1,014 peaks). (K–M) Scatter plots where the *x*-axis corresponds to the $\log_2$ (preimplantation embryo H3K4me3 promoter counts +1) and the *y*-axis corresponds to the $\log_2$ (sperm H3K4me3 promoter counts +1) at promoters with H3K4me3 loss in South African sperm that are expressed at the described stages of preimplantation embryo development (RPKM >1). Color of the scatter points correspond to the $\log_2$ preimplantation embryo RPKM gene expression +1. Dashed lines correspond to H3K4me3 promoter density cutoffs for preimplantation embryo (*x*-axis) or sperm (*y*-axis). Gray box denotes promoters with H3K4me3 in sperm and 4-cell (K), 8-cell (L), and ICM (M) preimplantation embryos. Note: GO, Gene Ontology; H3K4me3, histone H3 lysine 4 trimethylation; hESC, human embryonic stem cell; ICM, inner cell mass; NMDA, *N*-methyl-D-aspartate; NPC, neural progenitor cell; RPKM, reads per kilobase per million mapped reads; TEs, transposable elements; TSS, transcriptional start site.

we set out to identify whether sperm H3K4me3 peaks were predicted to persist in the early embryo. Regions with H3K4me3 enrichment in the preimplantation embryo were identified using an existing data set,[80] and we intersected sperm H3K4me3 peaks to preimplantation embryo H3K4me3 peaks (Figure S6A–F). We identified a high degree of overlap between sperm and preimplantation embryo H3K4me3 peaks. Indeed, 23,582 H3K4me3 peaks in sperm coincided with H3K4me3 peaks in the 4-cell embryo

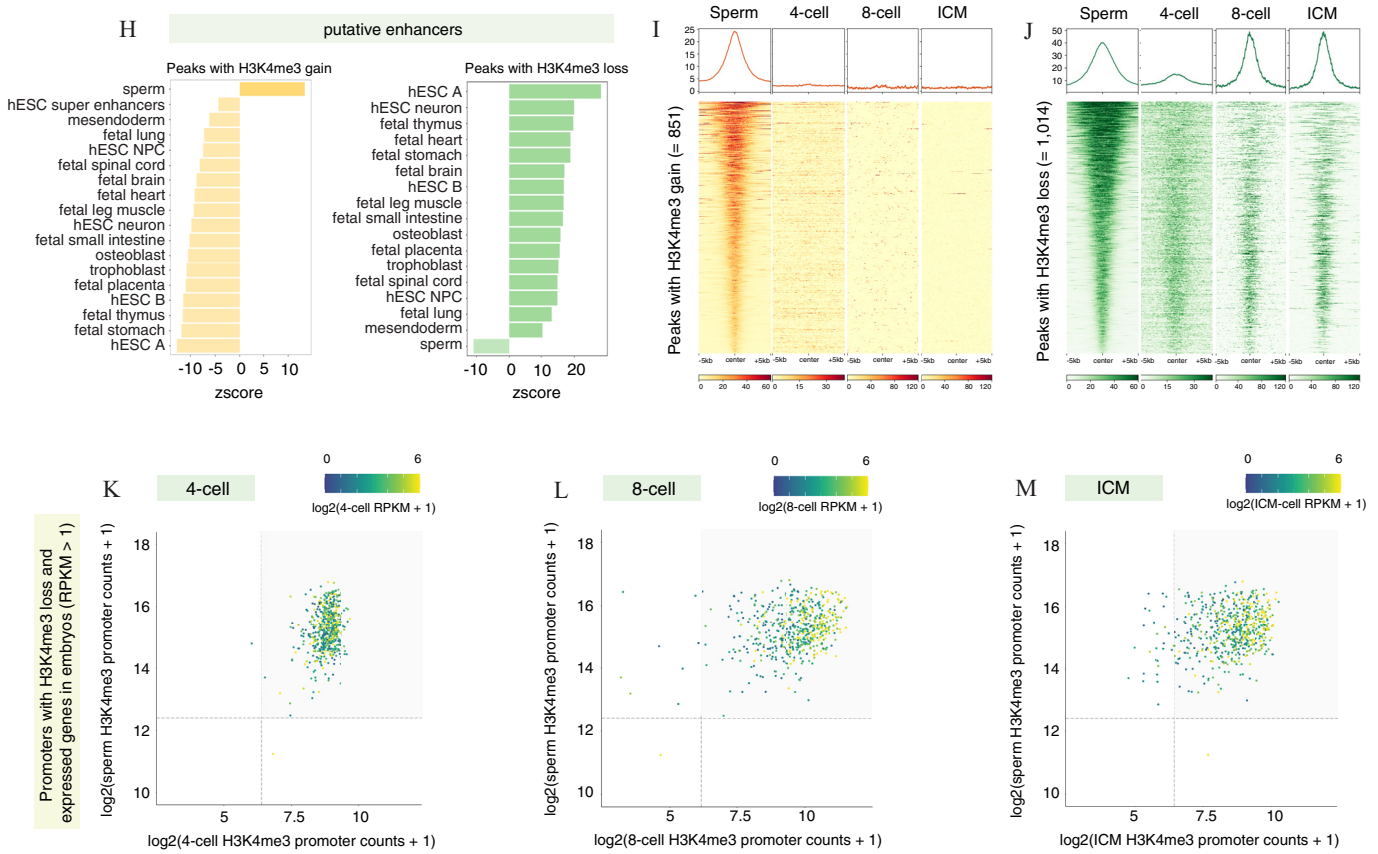

Figure 4. (*Continued.*)

(sperm-to-4-cell persistent peaks), and 7,004 of these overlapping peaks further overlapped H3K4me3 peaks in the 8-cell embryo through to the ICM (sperm-to-ICM persistent peaks; Figure S6G–I and Excel Tables S21–S23). We next determined whether peaks with *p*,*p*′-DDE–associated deH3K4me3 retained H3K4me3 in embryos at the 4-cell, 8-cell, and blastocyst stage (Figure 4I,J, Figure S6J–M). Interestingly, peaks gaining H3K4me3 in the sperm of DDT-exposed men showed a lack of overlap with embryonic H3K4me3 (Figure 4I, Figure S6J,K). In contrast, peaks losing H3K4me3 in sperm were highly enriched in H3K4me3 across the studied stages of preimplantation embryo development (Fisher test $p < 0.0001$; Figure 4J; Figure S6L,M). Revealing a potential cooperativity between DNAme and H3K4me3 in epigenetic transmission of environmental exposures to DDT, peaks with H3K4me3 loss were also highly enriched for intermediate- and high-DNAme sperm-to-zygote persistent regions (z-score >0, $p < 0.0001$, $n = 10{,}000$ permutations; Figure S6N).

Finally, we established whether sperm deH3K4me3 peaks that coincide with the retained H3K4me3 in the embryo could potentially impact embryonic gene expression. Promoters bearing H3K4me3 in both sperm and preimplantation embryos were associated with gene expression at the respective stage of embryogenesis (Fisher test $p < 0.0001$; see the "Materials and Methods" section; Figure S7A–J). Interestingly, most promoters with H3K4me3 loss in sperm retained H3K4me3 enrichment in the preimplantation embryo at genes that are expressed in 4-cell, 8-cell, and blastocyst embryos (Fisher test $p < 0.0001$; Figure 4K–M). To assess the functionality of the associated genes, we performed a GO analysis on promoters that retained H3K4me3 in sperm and embryos and overlapped them to genes that were expressed at the corresponding embryonic developmental stage (Figure S7K–M and Excel Tables S24–S26). Significant GO terms included pathways that were critical for development, such as placenta vascularization, mRNA processing, chromatin modification, and signaling (Figure S7K–M and Excel Tables S24–S26). Genes of interest included the *BRD1* (chromatin interacting protein), *MAP2K1* (linked to placenta vascularization), and the paternally expressed imprinted gene *PEG3*.

Regions with p,p′-DDE–Associated Differential DNAme and Overlap with deH3K4me3 Peaks in Sperm

We recently identified regions in sperm bearing both DNAme and H3K4me3 that are involved in important developmental processes.[27] We consequently wanted to determine if mDMCs and deH3K4me3 co-occurred at the same genomic loci. From the 830,188 MCC-seq target regions, 157,753 intersected H3K4me3 peaks (Figure 5A). Conversely, 34,741 H3K4me3 peaks overlapped MCC-seq target regions (Figure 5A), which can be explained by the preferential distribution of H3K4me3 peaks in CpG enriched regions that are well covered in the capture design (Chan et al.[59]). Interestingly, 1,744 mDMCs overlapped at least one H3K4me3 peak and 106 mDMCs overlapped a deH3K4me3 peak in South African sperm (Figure 5A). mDMCs that gained DNAme but lost deH3K4me3 (88 mDMCs) showed the highest overlap (Figure 5B). DNAme gain mDMCs and H3K4me3 loss regions may be the most biologically relevant in the embryo because, based on our findings, they *a*) may conserve DNAme and/or H3K4me3 in the embryo, and *b*) were more functionally relevant for development (Figures 2 and 4). Overlapping DNAme gain mDMCs and H3K4me3 loss regions were more likely found at <5 kb to the TSS at CpG shores compared with the other differential regions without both H3K4me3 loss and DNAme gain (Figure 5C,D; z-score >0; $p < 0.0001$; $n = 10{,}000$

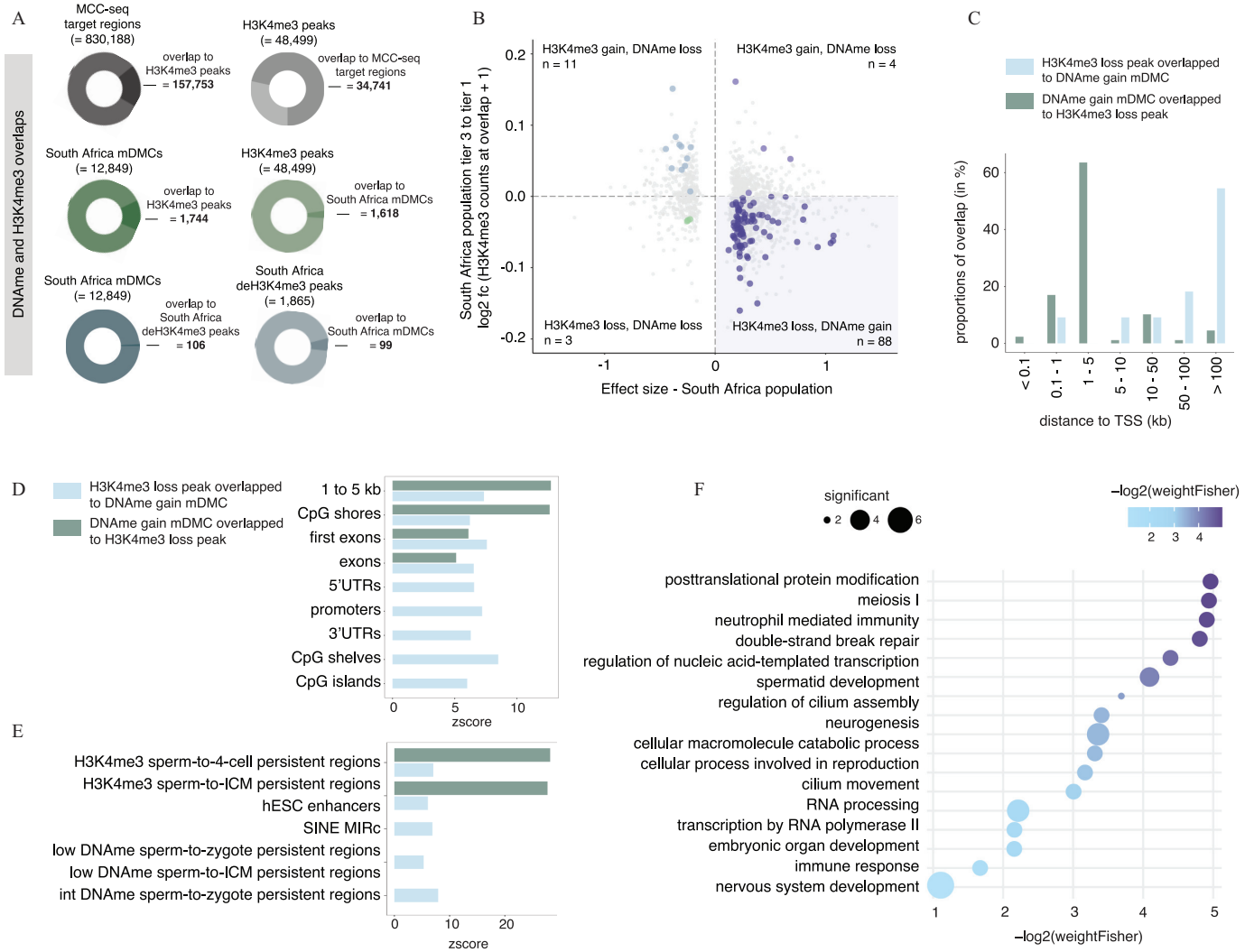

Figure 5. Intersection between differentially enriched H3K4me3 (deH3K4me3) peaks and merged differentially methylated CpGs (mDMCs) in sperm of South African men with different *p*,*p*′-DDE serum levels. (A) Proportion of MCC-seq target regions that overlap a H3K4me3 peak in sperm (overlap = 157,753; dark gray donut plot), proportion of H3K4me3 peaks in sperm that overlap a MCC-seq target region (overlap = 34,741; light gray donut plot), proportion of South Africa mDMCs that overlap a H3K4me3 peak in sperm (overlap = 1,744; dark green donut plot), proportion of H3K4me3 peaks in sperm that intersect a South Africa mDMC (overlap = 1,618; light green), proportion of South Africa mDMCs that overlap a deH3K4me3 peak in sperm (overlap = 106; dark blue), proportion of South Africa deH3K4me3 peaks in sperm that intersect an mDMC in sperm (overlap = 99; light blue). (B) Scatter plot corresponding to South African population ter3 to ter1 $\log_2$-fold change (H3K4me3 counts at overlapping mDMCs +1) relative to South Africa beta value at the mDMCs. Gray dots correspond to mDMCs that do not overlap with a deH3K4me3 peak. mDMCs intersecting a deH3K4me3 peak are denoted by color. H3K4me3 loss + DNAme gain overlap was the predominant overlap ($n = 88$ of 106). (C) Distribution of peaks with H3K4me3 loss overlapped to DNAme gain mDMC (light blue) and DNAme gain mDMC overlapped to H3K4me3 loss region (dark green) relative to the TSS. (D) Enrichment for peaks with H3K4me3 loss overlapped to DNAme gain mDMC (light blue) and DNAme gain mDMC overlapped to H3K4me3 loss region (dark green) at genic and CpG annotations. Positive enrichments are determined by *z*-scores using the Bioconductor package regioneR. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions resampled from deH3K4me3 peaks (light blue) and mDMCs (dark green), respectively. (E) Enrichment for peaks with H3K4me3 loss overlapped to DNAme gain mDMC (blue) and DNAme gain mDMC overlapped to H3K4me3 loss region (green) at TE annotations and characterized DNAme/H3K4me3 persistent regions (see Figure S2 and S5). Positive enrichments are determined by *z*-scores using the Bioconductor package regioneR. For all annotations displayed, $p < 0.0001$ and $n = 10{,}000$ permutations of random regions resampled from deH3K4me3 peaks (light blue) and mDMCs (dark green), respectively. (F) Selected significant pathways from GO analysis on peak with H3K4me3 loss that overlaps a DNAme gain mDMC at promoters and 1–5 kb genic space. Size of dots corresponds to the number of significant genes in the pathway. Color and position of dots correspond to $-\log_2$ (weightFisher). See Excel Table S27. (G) Representative IGV tracks of peak with H3K4me3 loss and DNAme gain in South African sperm at an LTR-ERV1 and in the *TRAPPC9* genic space. Purple shaded box corresponds to the mDMC. Tracks below are a zoom at the mDMC. Intermediate CpGs are indicated in purple. Note: ChiP-Seq, chromatin immunoprecipitation targeting histone H3K4me3 followed by sequencing; DNAme, DNA methylation; GO, Gene Ontology; H3K4me3, histone H3 lysine 4 trimethylation; hESC, human embryonic stem cell; ICM, inner cell mass; IGV, Integrative Genome Viewer; int, intermediate; MCC-seq, MethylC-Capture-sequencing; *p*,*p*′-DDE, *p*,*p*′-dichlorodiphenyldichloroethylene; TE, transposable element; ter, tertile; TSS, transcriptional start site; UTR, untranslated region.

permutations). These overlapping regions were also more likely enriched at H3K4me3 sperm-to-4-cell and sperm-to-ICM persistent regions, SINE-MIRc, hESC enhancers, as well as low- and intermediate-DNAme persistent regions than other differential regions with only one change in either DNAme or H3K4me3 (Figure 5E; *z*-score >0; $p < 0.0001$; $n = 10{,}000$ permutations). GO analysis of peaks with a loss of H3K4me3 and a gain in DNAme revealed enrichment for genes involved in reproduction, spermatogenesis, embryo development, and neurodevelopment (Figure 5F; Excel Table S27). Representative visualizations of dose–response differences in enrichment for DNAme and H3K4me3 in overlapping regions are shown for HERV/LTR43

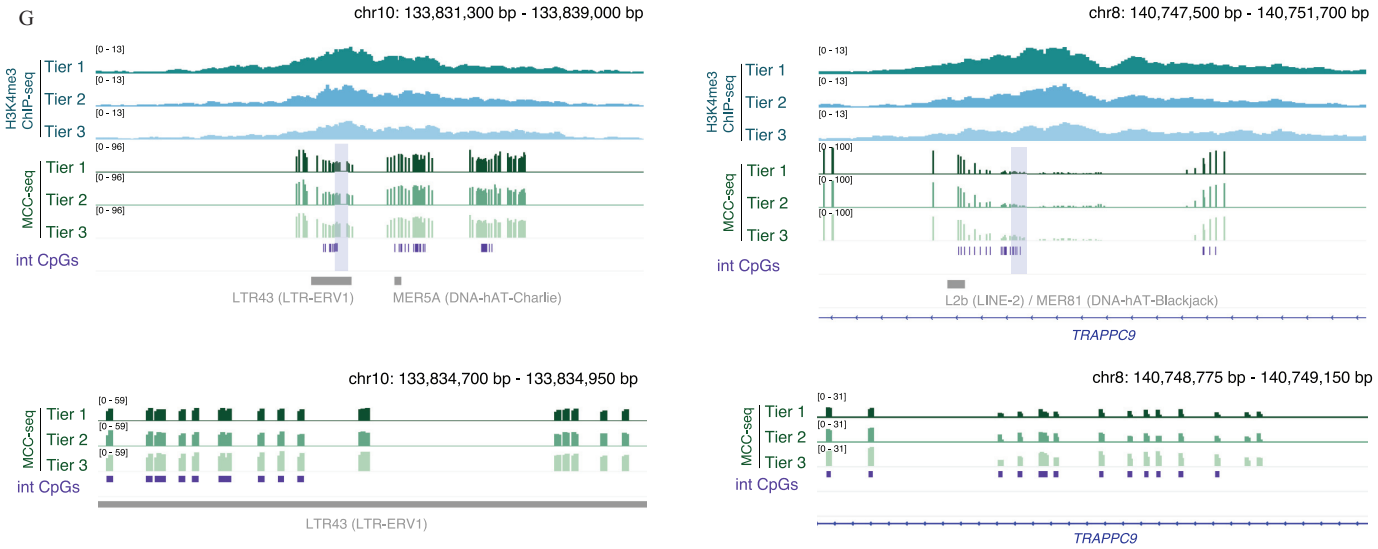

Figure 5. (*Continued.*)

and the gene *TRAPPC9* (Figure 5G), which has been implicated in cognitive disability and microcephaly.[116] Interestingly, a search for regulatory elements in HERV/LTR43,[117] revealed that it is bound by numerous transcription factors implicated in development and disease. These included PAX5 (implicated in the developing central nervous system and testis), FOS (regulator of cell proliferation and differentiation), MYC (proto-oncogene), NFKB1 (inflammatory regulator), and NFE2 (blood vessel and bone development).

Discussion

DDT has been used worldwide since the 1940s as an organochlorine insecticide for agriculture crops and control of vector-borne diseases.[118–120] Although DDT was banned in the 1970s owing to concerns for wildlife and the chemical's bioaccumulation and biomagnification in the environment, DDT still carries sanctioned uses for malarial control.[121] Alarmingly, large epidemiological studies have implicated DDT and/or its metabolite *p*,*p*′-DDE in cancer,[122] diabetes,[123] early pregnancy losses,[124] shorter gestational period and lower birth weight,[8] urogenital birth defects,[125] poor semen parameters,[12] and neurodevelopmental disorders.[7,11] There is now an abundance of evidence from experimental models that paternal exposures can negatively impact offspring phenotypes.[126] In rodent models of paternal epigenetic inheritance DNAme, chromatin and noncoding RNA have been implicated in response to environmental challenges, including diet,[17,28,127–129] trauma,[130] and toxicants.[13,41,131] In humans, the intergenerational consequences of maternal exposure to environmental factors, such as toxicants, are well studied. In contrast, few studies have examined the paternal modes of transmission that can impact child health, and the ones that have were predominantly epidemiology based.[132] Our overview and interpretation of the main findings from this study are presented in (Figure 6).

Limitations in the present study include the relatively small sample size. Unlike samples acquired through most clinical studies, these samples were highly precious and unique, coming from remote areas of the world from small indigenous populations. The size and regions of the populations limit the metadata that can be publicly released owing to risk of participant identification (for GDPR adherence). Because the Greenlandic cohort samples were from the INUENDO biobank, it was not possible to coordinate factors with the prospective South African study. As such, the populations differed in the timing of collection, in mean age (25 vs. 31 years of age in South African and Greenlandic men, respectively) and the site of *p*,*p*′-DDE measurement. Moreover, the Greenlandic Inuit men were proven fertile, whereas the fertility status of the South African men was unknown. Semen analysis has been performed on both populations in prior studies. There was no association with *p*,*p*′-DDE and sperm concentration or morphology, whereas higher doses were associated with reduced motility in samples from Greenlandic Inuit men.[133] Likely reflecting higher exposures in the South African population, in a previous study we identified increased incidences of teratozoospermia, asthenozoospermia, and lower motility.[12] In this study, to best match the fertility status of the Greenlandic Inuit men, we selected South African samples with normal sperm counts and all but one sample had normal motility. It was not deemed possible for sociocultural reasons to request participants to provide information on fertility status. Altered fertility status was associated with changes in DNAme, yet the identification of CpGs that were consistently changed in men with infertility linked to poor embryo quality was inconclusive.[134] Conceivably some of the DNAme differences detected between the two populations may be linked to fertility status. However, given that there is no clear infertile DNAme signature, we cannot, for example, remove a designated set of CpGs from our analysis. Consideration must also be given to the impact of unidentified environmental confounders, such as exposures to other EDC compounds, and an inability to discriminate between acute vs. long-term exposures. In addition, we do not know how stable the sperm epigenome is and whether *p*,*p*′-DDE–associated alterations are temporary or permanent. Large-scale intergenerational studies that include sperm epigenome analyses are needed to better understand the connections between paternal exposures, the sperm epigenome, and child development. However, owing to cultural and societal differences, such studies are inherently challenging in the context of these population and are more suited to a clinical setting.

Notwithstanding the population differences mentioned above and the range in *p*,*p*′-DDE body burdens, there was a high degree of consistency of *p*,*p*′-DDE-associated DNAme alterations. Although we cannot fully rule out genetic influences, selection bias, or residual confounding factors, the overlapping DNAme changes confirmed in two independent populations of different selection/confounder structure, clearly indicates that this overlap was unlikely to be driven by bias. In both populations, mDMCs localized to genes and regulatory elements that reflect the

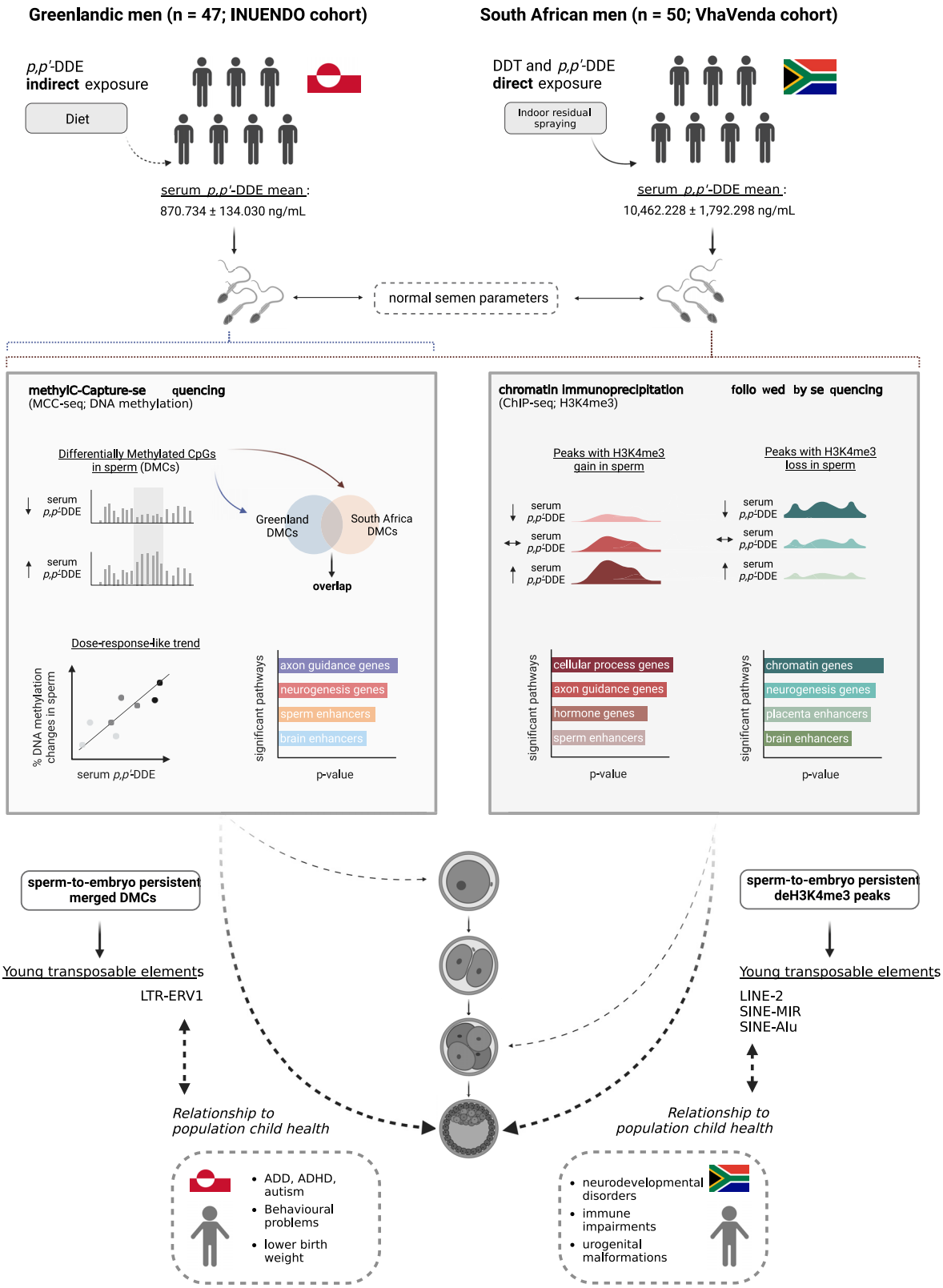

Figure 6. Overview and author interpretation of the main findings; a summary of the consequences of *p,p′*-DDE exposure on the sperm epigenome of Greenlandic Inuit and South African VhaVenda men and the health implications for the next generation. We assessed the sperm epigenome (DNAme and H3K4me3) of two geographically diverse populations, Greenlandic Inuit and South African men, in relationship to *p,p′*-DDE serum levels. Dose–response trends between sperm DNAme/H3K4me3 and *p,p′*-DDE serum levels were observed within both populations. Regions with altered epigenetic marks in sperm were related to neurodevelopment and fertility, co-localized to young TEs, and overlapped regions that were predicted to resist epigenetic reprogramming in the preimplantation embryo. Note: ADD, attention deficit disorder; ADHD, attention deficit/hyperactivity disorder; ChIP-seq, chromatin immunoprecipitation targeting histone H3K4me3 followed by sequencing; DDT, dichlorodiphenyltrichloroethane; DMCs, differentially methylated CpGs; DNAme, DNA methylation; H3K4me3, histone H3 lysine 4 trimethylation; *p,p′*-DDE, *p,p′*-dichlorodiphenyldichloroethylene; TEs, transposable elements.

increased susceptibility to disease and developmental disorders associated with EDC and DDT exposures, including increased birth defects and neurodevelopmental impairments.[14,55,135,136] Strikingly, 17.5% of regions with altered DNAme in Greenlandic sperm were also altered in South African sperm. The nonoverlapping mDMCs could, in part, reflect the different routes of exposure between the two populations whereby Greenlandic men are predominantly exposed to *p*,*p*′-DDE through diet and South African men are mostly exposed to DDT through indoor residual spraying. We also observed a striking association between measured *p*,*p*′-DDE serum levels and the magnitude of DNAme and H3K4me3 changes in Greenlandic and South African sperm. Because we cannot infer a causal dose–response relationship between serum *p*,*p*′-DDE and sperm DNAme/H3K4me3 in sperm owing to confounding factors and the nature of the study (cross-sectional), we coin these associations as dose–response-like trends. *p*,*p*′-DDE may influence epigenetic marks in sperm, and subsequently paternal epigenetic inheritance, via its endocrine disruptor effects on androgenic signaling. Androgen regulates genes in one-carbon metabolism, including glycine *N*-methyltransferase, cystathionine β-synthase, and ornithine decarboxylase, and in this way can impact methylation of DNA and histones.[137] Associations between DDT and one-carbon metabolism have previously been drawn in epidemiological studies. Indeed, in a cross-sectional analysis of Chinese Women, DDT isomers were found to be inversely correlated to serum folate levels,[138] suggesting that high levels of DDT may exert a negative impact on folate and cysteine levels. In another study of 291 women in China, maternal preconception folic acid supplementation and vitamin B sufficiency protected against the adverse effects of DDT exposure.[139] In addition, folic acid supplementation in rat models reduced transgenerational micro-RNA perturbations and offspring birth defects associated with paternal DDT-containing POP exposure.[45] Although South African men from this cohort had folate and homocysteine serum levels above the WHO safe cutoff level, DDT exposure could lead to perturbations of their folate and methionine cycles that are reflected at the level of the sperm epigenome.

In both populations, DNAme was preferentially altered in the intergenic space, at intermediate-DNAme CpGs, where there were more gains than losses of DNAme. The intergenic space is enriched for TEs, including retroviruses, which are controlled by DNA and histone methylation.[140,141] TEs function as promoters, transcription factor binding sites, and enhancers implicated in pluripotency.[142] Enhancer–promoter interactions are facilitated by the CCCTC-binding factor (CTCF) binding sites in TEs, which, in turn, influence chromatin looping.[143] Factors such as aging, infection, or hormones have been postulated to reactivate human endogenous retroviruses (ERVs).[144] Perhaps reflecting the endocrine disrupting activity of DDT metabolites was the specific sensitivity associated with DDT exposure to the DNAme changes observed in the long terminal repeat ERVs (LTR-ERV1, LTR-ERVK, LTR-ERVL). Expression of ERV families is temporally and spatially regulated during human embryogenesis, and they function in gene regulation as alternate promoters and enhancers.[145,146] Likewise, deH3K4me3 was overrepresented at LINE-2, SINE-MIR, and SINE-Alu, whereas LTR-ERV1 overlapped both deH3K4me3 and mDMCs. These epigenetic changes at TEs may, in turn, serve in the epigenetic inheritance of DDT-associated phenotypes. TEs known to be resistant to epigenetic reprogramming in human preimplantation embryos include ERVs, SINEs, and LINEs.[79,147] In line with our findings, exposing male rats to the endocrine disruptor vinclozolin *in utero*, led to transgenerational DNAme changes in their sperm, with the most pronounced DNAme differences occurring at promoters that contain ERVs.[148] All factors taken together, this study bolsters the evidence that in humans, as in plants, invertebrates, and rodents,[149] TEs are attractive genome mediators of environmentally influenced nongenetic inheritance. What remains unknown for all species is how epigenetic marks at TEs escape embryonic reprogramming.

A limitation of epigenetic inheritance studies is the difficulty to go beyond association of an epigenetic alteration from an exposure and a change of function in the offspring. This has been particularly challenging in human studies, where it is difficult to access tissues in the next generation that are matched to a father, as was the case in this study. Here, we addressed this limitation by predicting whether regions in sperm bearing altered DNAme and H3K4me3 in association with DDT and *p*,*p*′-DDE serum levels persisted in the embryo. To do so, we mined preexisting RRBS[79] and low-input H3K4me3 ChIP-seq[80] data sets from human preimplantation embryos (zygote, 2-cell, 4-cell, 8-cell, and ICM for the RRBS from donor couples who had undergone *in vitro* fertilization (IVF) treatments; 4-cell, 8-cell, ICM for the H3K4me3 ChIP-seq from donor couples who had undergone IVF or intracytoplasmic sperm injection treatments), followed by *in silico* analyses overlapping mDMCs and deH3K4me3 peaks to regions that retained DNAme or H3K4me3 in the preimplantation embryo. This is a plausible route for paternally mediated epigenetic inheritance that leads to effects in offspring and is supported by exposure studies in animal models.[17,41,129,131,150,151] Of relevance are studies in rats of paternal DDT exposures that were associated with transgenerational alterations in sperm DNAme profiles, obesity, and developmental abnormalities.[41,150] However, in humans, the connection between paternal exposures, the sperm epigenome, and health in the next generation is underexplored. In addition, our DNAme experiments on Greenlandic and South African sperm, as well as our *in silico* preimplantation embryo analysis,[79] focused on 5-methyl-cytosine (5mC) and omitted hydroxylated 5-methyl-cytosine (5hmC) from our analyses. 5hmC results from the oxidation of 5mC by ten-eleven translocation (TET) enzymes and is found at low levels in ESCs and preimplantation embryos.[152–154] This epigenetic modification is associated with poised chromatin marks found at developmental genes and is linked to epigenetic reprogramming in the paternal pronucleus of the mouse zygote.[153,155] The MCC-seq (sperm[59]) and RRBS (preimplantation embryo[79]) technologies do not allow for determination of 5hmC. Nonetheless, as the 5mC state is predominant over 5hmC it is likely the more informative state.

A consistent finding for *p*,*p*′-DDE -associated mDMCs and deH3K4me3 peaks in sperm was their occurrence at genes or regulatory elements related to neurodevelopment and neurofunction. The health connections between paternal DDT and *p*,*p*′-DDE exposures and neurodevelopment are unknown. The incidence of neurodevelopmental comorbidities associated with maternal serum *p*,*p*′-DDE in Greenlandic children of the INUENDO cohort have been reported. Prenatal and postnatal exposures via the mother to *p*,*p*′-DDE and PCB-153 was associated with an increased incidence of hyperactivity in children.[156] Interestingly, maternal exposures to EDCs have been associated with reductions in cognitive function, such as delayed language acquisition. For example, a Finnish birth cohort study of >1 million mother–child pairs found that the risk of childhood autism was significantly increased with serum *p*,*p*′-DDE levels that were in the highest 75th percentile.[136] There are few large-scale epidemiological studies suggesting that exposures to DDT promotes cancer progression, notably cancers of the liver and testis and non-Hodgkin lymphoma.[157] We did not find clear epigenomic connections between DDT or *p*,*p*′-DDE and exposure to cancer in the sperm of Greenlandic or South African men. However, we did identify an enrichment of pathways involved in immunity at

genes overlapping mDMCs or deH3K4me3 peaks in Greenlandic and South African sperm. Interestingly, in the South African VhaVenda cohort, high rates of allergies in children have been reported,[97] along with immunosuppressed response to vaccines.[96] Finally, we also found an association between mDMCs and deH3K4me3 in sperm and insulin secretion pathways. In line with population epidemiological data, Greenlandic children exhibit elevated rates of obesity.[99] In addition, higher body mass in VhaVenda girls has been associated with elevated maternal serum DDT levels.[158] As in most birth cohorts, the paternal contributions were not studied, yet conceivably paternal *p*,*p*′-DDE exposure may also have contributed to the findings of increased complex disease risk. Our inability to track childhood health in relation to paternal exposures in the studied populations prevented us from making these essential connections.

Despite a multitude of documented adverse health effects attributed to DDT exposure, its effectiveness as an antimalarial control method has led to its continued use in 14 countries.[159] The impacts of DDT on human health are not restricted to regions of use given that DDT has been shown to be transported over long distances in the environment by weather patterns and ocean currents.[160] Owing to climate change, the long-range contamination, environmental concentration, persistence, and bioaccumulation of DDT and its persistent transformation product, *p*,*p*′-DDE, are predicted to worsen.[161] Human exposures to DDT are consequently expected to increase in regions distant from its use.[161] Taking this into consideration with the implication that DDT exposure can alter the sperm epigenome and may be implicated in epigenetic inheritance of disease, it is essential to consider alternative methods for malarial control. In addition, our findings highlight the need for future regulatory decision-making to be based on chemical risk assessments that incorporate experimental toxicology and epigenetic profiling of sperm.

Acknowledgments

Author contributions: A. Lismer: methodology, data curation, investigation, formal analysis, visualization, and writing—original draft, editing/reviewing; X. Shao: software, methodology, formal analysis, and writing—editing/reviewing; M.C. Dumargne: methodology, investigation, and formal analysis; C. Lafleur: methodology and investigation; R. Lambrot: methodology and investigation; D. Chan: formal analysis, visualization, and writing—editing/reviewing; G. Toft: data curation and writing—editing/reviewing; J.P. Bonde: data curation and writing—editing/reviewing; A.J. MacFarlane: methodology, investigation, funding acquisition, and writing—editing/reviewing; M. Bornman: methodology, investigation, data curation, and writing—editing/reviewing; N. Aneck-Hahn: methodology, investigation, data curation, and writing—editing/reviewing, project administration; S. Patrick: methodology, investigation, data curation, and writing—editing/reviewing, project administration; J.M. Bailey: conceptualization and funding acquisition; C. de Jager: conceptualization, methodology, funding acquisition, and writing—editing/reviewing; J.M. Trasler: supervision, conceptualization, methodology, funding acquisition, and writing—editing/reviewing; V. Dumeaux: software, formal analysis, visualization, and writing—editing/reviewing; S. Kimmins: supervision, conceptualization, methodology, data curation, funding acquisition, and writing—original draft, editing/reviewing, project administration.

We thank Olusola Sotunde and Hope Weiler for their assistance with the study, as well as the South African Medical Research Council for funding (Self Initiated Research grant), the fieldworkers, and participants from the different villages in the Vhembe District, Limpopo Province, South Africa. We also thank Genome Quebec for the construction of the whole-genome bisulfite sequencing (WGBS) libraries. We thank Guillaume Bourque from the McGill Genome Centre for his support for the MethylC-Capture-sequencing (MCC-Seq) project and Marie-Michelle Simon for preparing the MCC-Seq libraries. Epigenomic sequencing was performed at the McGill Genome Center under the guidance of Tony Kwan. We appreciate advice from Yann Joly (McGill, Research Director of the Centre of Genomics and Policy) regarding regulatory ethics related to genomic data access. This research was enabled in part by support provided by Calcul Quebec and the Digital Research Alliance of Canada.

S. Kimmins was awarded funding for this study from the Canadian Institutes of Health Research (CIHR) grants (Developmental Origins of Health and Disease team grant 358654 and Operating grant 350129) and J. Bailey was awarded funding for this study from the CIHR (CIHR Boys and Men's Health team grant TE1-138294). J.M. Trasler is a Distinguished James McGill Professor. J.M. Trasler's funding for this study is provided by the CIHR Foundation grant (FDN-148425). Other funding includes CEE-151618 to the McGill Epigenomics Mapping Centre, which is part of the Canadian Epigenetics Environment and Health Research Consortium Network.

Compliance with General Data Protection Regulations in Denmark permits access to the Greenlandic Inuit MCC-seq data upon request via material transfer agreement only. South African phenotypic and epigenomic data will be made available by request via material transfer agreement only.

References

1. Lallas PL. 2001. The Stockholm convention on persistent organic pollutants. Am j Int Law 95(3):692–708, https://doi.org/10.2307/2668517.
2. Turusov V, Rakitsky V, Tomatis L. 2002. Dichlorodiphenyltrichloroethane (DDT): ubiquity, persistence, and risks. Environ Health Perspect 110(2):125–128, PMID: 11836138, https://doi.org/10.1289/ehp.02110125.
3. van den Berg H, Manuweera G, Konradsen F. 2017. Global trends in the production and use of DDT for control of malaria and other vector-borne diseases. Malar J 16(1):401, PMID: 28982359, https://doi.org/10.1186/s12936-017-2050-2.
4. WHO (World Health Organization). 2006. *Indoor Residual Spraying: Use of Indoor Residual Spraying for Scaling up Global Malaria Control and Elimination: WHO Position Statement*. Geneva, Switzerland: World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/69386/WHO_HTM_MAL_2006.1112_eng.pdf?sequence=1 [accessed 1 August 2021].
5. Kelce WR, Stone CR, Laws SC, Gray LE, Kemppainen JA, Wilson EM. 1995. Persistent DDT metabolite p,p′-DDE is a potent androgen receptor antagonist. Nature 375(6532):581–585, PMID: 7791873, https://doi.org/10.1038/375581a0.
6. Bitman J, Cecil HC, Harris SJ, Fries GF. 1968. Estrogenic activity of *o,p′*;-DDT in the mammalian uterus and avian oviduct. Science 162(3851):371–372, PMID: 5677532, https://doi.org/10.1126/science.162.3851.371.
7. Eskenazi B, Chevrier J, Rosas LG, Anderson HA, Bornman MS, Bouwman H, et al. 2009. The Pine River statement: human health consequences of DDT use. Environ Health Perspect 117(9):1359–1367, PMID: 19750098, https://doi.org/10.1289/ehp.11748.
8. Longnecker MP, Klebanoff MA, Zhou H, Brock JW. 2001. Association between maternal serum concentration of the DDT metabolite DDE and preterm and small-for-gestational-age babies at birth. Lancet 358(9276):110–114, PMID: 11463412, https://doi.org/10.1016/S0140-6736(01)05329-6.
9. Bornman R, de Jager C, Worku Z, Farias P, Reif S. 2010. DDT and urogenital malformations in newborn boys in a malarial area. BJU Int 106(3):405–411, PMID: 19849691, https://doi.org/10.1111/j.1464-410X.2009.09003.x.
10. Bouwman H, Kylin H. 2009. Malaria control insecticide residues in breast milk: the need to consider infant health risks. Environ Health Perspect 117(10):1477–1480, PMID: 20019894, https://doi.org/10.1289/ehp.0900605.
11. Cartier C, Muckle G, Jacobson SW, Jacobson JL, Dewailly E, Ayotte P, et al. 2014. Prenatal and 5-year *p,p′*-DDE exposures are associated with altered sensory processing in school-aged children in Nunavik: a visual evoked potential study. Neurotoxicology 44:8–16, PMID: 24812027, https://doi.org/10.1016/j.neuro.2014.04.009.
12. Aneck-Hahn NH, Schulenburg GW, Bornman MS, Farias P, de Jager C. 2007. Impaired semen quality associated with environmental DDT exposure in young

men living in a malaria area in the Limpopo Province, South Africa. J Androl 28(3):423–434, PMID: 17192596, https://doi.org/10.2164/jandrol.106.001701.
13. Skinner MK, Ben Maamar M, Sadler-Riggleman I, Beck D, Nilsson E, McBirney M, et al. 2018. Alterations in sperm DNA methylation, non-coding RNA and histone retention associate with DDT-induced epigenetic transgenerational inheritance of disease. Epigenetics Chromatin 11(1):8, PMID: 29482626, https://doi.org/10.1186/s13072-018-0178-0.
14. Van Oostdam J, Donaldson SG, Feeley M, Arnold D, Ayotte P, Bondy G, et al. 2005. Human health implications of environmental contaminants in Arctic Canada: a review. Sci Total Environ 351–352:165–246, PMID: 16297438, https://doi.org/10.1016/j.scitotenv.2005.03.034.
15. Weihe P, Debes F, Halling J, Petersen MS, Muckle G, Odland JØ, et al. 2016. Health effects associated with measured levels of contaminants in the Arctic. Int J Circumpolar Health 75:33805, PMID: 27974137, https://doi.org/10.3402/ijch.v75.33805.
16. Lismer A, Siklenka K, Lafleur C, Dumeaux V, Kimmins S. 2020. Sperm histone H3 lysine 4 trimethylation is altered in a genetic mouse model of transgenerational epigenetic inheritance. Nucleic Acids Res 48(20):11380–11393, PMID: 33068438, https://doi.org/10.1093/nar/gkaa712.
17. Lismer A, Dumeaux V, Lafleur C, Lambrot R, Brind'Amour J, Lorincz MC, et al. 2021. Histone H3 lysine 4 trimethylation in sperm is transmitted to the embryo and associated with diet-induced phenotypes in the offspring. Dev Cell 56(5):671–686.e6, PMID: 33596408, https://doi.org/10.1016/j.devcel.2021.01.014.
18. Siklenka K, Erkek S, Godmann M, Lambrot R, McGraw S, Lafleur C, et al. 2015. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science 350(6261):aab2006, PMID: 26449473, https://doi.org/10.1126/science.aab2006.
19. Kimmins S, Sassone-Corsi P. 2005. Chromatin remodelling and epigenetic features of germ cells. Nature 434(7033):583–589, PMID: 15800613, https://doi.org/10.1038/nature03368.
20. Lambrot R, Siklenka K, Lafleur C, Kimmins S. 2019. The genomic distribution of histone H3K4me2 in spermatogonia is highly conserved in sperm. Biol Reprod 100(6):1661–1672, PMID: 30951591, https://doi.org/10.1093/biolre/ioz055.
21. Maezawa S, Yukawa M, Alavattam KG, Barski A, Namekawa SH. 2018. Dynamic reorganization of open chromatin underlies diverse transcriptomes during spermatogenesis. Nucleic Acids Res 46(2):593–608, PMID: 29126117, https://doi.org/10.1093/nar/gkx1052.
22. Larose H, Shami AN, Abbott H, Manske G, Lei L, Hammoud SS. 2019. Gametogenesis: a journey from inception to conception. Curr Top Dev Biol 132:257–310, PMID: 30797511, https://doi.org/10.1016/bs.ctdb.2018.12.006.
23. Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, et al. 2013. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. Nat Struct Mol Biol 20(7):868–875, PMID: 23770822, https://doi.org/10.1038/nsmb.2599.
24. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. 2009. Distinctive chromatin in human sperm packages genes for embryo development. Nature 460(7254):473–478, PMID: 19525931, https://doi.org/10.1038/nature08162.
25. Lesch BJ, Silber SJ, McCarrey JR, Page DC. 2016. Parallel evolution of male germline epigenetic poising and somatic development in animals. Nat Genet 48(8):888–894, PMID: 27294618, https://doi.org/10.1038/ng.3591.
26. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, et al. 2010. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 17(6):679–687, PMID: 20473313, https://doi.org/10.1038/nsmb.1821.
27. Lambrot R, Chan D, Shao X, Aarabi M, Kwan T, Bourque G, et al. 2021. Whole-genome sequencing of H3K4me3 and DNA methylation in human sperm reveals regions of overlap linked to fertility and development. Cell Rep 36(3):109418, PMID: 34289352, https://doi.org/10.1016/j.celrep.2021.109418.
28. Pepin AS, Lafleur C, Lambrot R, Dumeaux V, Kimmins S. 2022. Sperm histone H3 lysine 4 tri-methylation serves as a metabolic sensor of paternal obesity and is associated with the inheritance of metabolic dysfunction. Mol Metab 59:101463, PMID: 35183795, https://doi.org/10.1016/j.molmet.2022.101463.
29. Lesch BJ, Tothova Z, Morgan EA, Liao Z, Bronson RT, Ebert BL, et al. 2019. Intergenerational epigenetic inheritance of cancer susceptibility in mammals. Elife 8:e39380, PMID: 30963999, https://doi.org/10.7554/eLife.39380.
30. Stringer JM, Forster SC, Qu Z, Prokopuk L, O'Bryan MK, Gardner DK, et al. 2018. Reduced PRC2 function alters male germline epigenetic programming and paternal inheritance. BMC Biol 16(1):104, PMID: 30236109, https://doi.org/10.1186/s12915-018-0569-5.
31. Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME. 2010. Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. Nat Cell Biol 12(9):853–862, PMID: 20676102, https://doi.org/10.1038/ncb2089.
32. Aoshima K, Inoue E, Sawa H, Okada Y. 2015. Paternal H3K4 methylation is required for minor zygotic gene activation and early mouse embryonic development. EMBO Rep 16(7):803–812, PMID: 25925669, https://doi.org/10.15252/embr.201439700.
33. Ishiuchi T, Abe S, Inoue K, Yeung WKA, Miki Y, Ogura A, et al. 2021. Reprogramming of the histone H3.3 landscape in the early mouse embryo. Nat Struct Mol Biol 28(1):38–49, PMID: 33169018, https://doi.org/10.1038/s41594-020-00521-1.
34. Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, et al. 2011. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell 146(6):1029–1041, PMID: 21925323, https://doi.org/10.1016/j.cell.2011.08.016.
35. Bestor TH. 1998. Cytosine methylation and the unequal developmental potentials of the oocyte and sperm genomes. Am J Hum Genet 62(6):1269–1273, PMID: 9585619, https://doi.org/10.1086/301891.
36. Bourc'his D, Bestor TH. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 431(7004):96–99, PMID: 15318244, https://doi.org/10.1038/nature02886.
37. Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99(3):247–257, PMID: 10555141, https://doi.org/10.1016/s0092-8674(00)81656-6.
38. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. Nature 448(7154):714–717, PMID: 17687327, https://doi.org/10.1038/nature05987.
39. Harutyunyan AS, Krug B, Chen H, Papillon-Cavanagh S, Zeinieh M, De Jay N, et al. 2019. H3K27M induces defective chromatin spread of PRC2-mediated repressive H3K27me2/me3 and is essential for glioma tumorigenesis. Nat Commun 10(1):1262, PMID: 30890717, https://doi.org/10.1038/s41467-019-09140-x.
40. Michealraj KA, Kumar SA, Kim LJY, Cavalli FMG, Przelicki D, Wojcik JB, et al. 2020. Metabolic regulation of the epigenome drives lethal infantile ependymoma. Cell 181(6):1329–1345.e24, PMID: 32445698, https://doi.org/10.1016/j.cell.2020.04.047.
41. Maurice C, Dalvai M, Lambrot R, Deschênes A, Scott-Boyer MP, McGraw S, et al. 2021. Early-life exposure to environmental contaminants perturbs the sperm epigenome and induces negative pregnancy outcomes for three generations via the paternal lineage. Epigenomes 5(2):10, PMID: 34968297, https://doi.org/10.3390/epigenomes5020010.
42. Lambrot R, Xu C, Saint-Phar S, Chountalos G, Cohen T, Paquet M, et al. 2013. Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. Nat Commun 4:2889, PMID: 24326934, https://doi.org/10.1038/ncomms3889.
43. Radford EJ, Ito M, Shi H, Corish JA, Yamazawa K, Isganaitis E, et al. 2014. In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. Science 345(6198):1255903, PMID: 25011554, https://doi.org/10.1126/science.1255903.
44. King SE, McBirney M, Beck D, Sadler-Riggleman I, Nilsson E, Skinner MK. 2019. Sperm epimutation biomarkers of obesity and pathologies following DDT induced epigenetic transgenerational inheritance of disease. Environ Epigenet 5(2):dvz008, PMID: 31186947, https://doi.org/10.1093/eep/dvz008.
45. Herst PM, Dalvai M, Lessard M, Charest PL, Navarro P, Joly-Beauparlant C, et al. 2019. Folic acid supplementation reduces multigenerational sperm miRNA perturbation induced by *in utero* environmental contaminant exposure. Environ Epigenet 5(4):dvz024, PMID: 31853372, https://doi.org/10.1093/eep/dvz024.
46. Lessard M, Herst PM, Charest PL, Navarro P, Joly-Beauparlant C, Droit A, et al. 2019. Prenatal exposure to environmentally-relevant contaminants perturbs male reproductive parameters across multiple generations that are partially protected by folic acid supplementation. Sci Rep 9(1):13829, PMID: 31554827, https://doi.org/10.1038/s41598-019-50060-z.
47. Consales C, Toft G, Leter G, Bonde JPE, Uccelli R, Pacchierotti F, et al. 2016. Exposure to persistent organic pollutants and sperm DNA methylation changes in Arctic and European populations. Environ Mol Mutagen 57(3):200–209, PMID: 26801515, https://doi.org/10.1002/em.21994.
48. Muir DCG, Wagemann R, Hargrave BT, Thomas DJ, Peakall DB, Norstrom RJ. 1992. Arctic marine ecosystem contamination. Sci Total Environ 122(1–2):75–134, PMID: 1514106, https://doi.org/10.1016/0048-9697(92)90246-o.
49. Brown TM, Macdonald RW, Muir DCG, Letcher RJ. 2018. The distribution and trends of persistent organic pollutants and mercury in marine mammals from Canada's Eastern Arctic. Sci Total Environ 618:500–517, PMID: 29145101, https://doi.org/10.1016/j.scitotenv.2017.11.052.
50. Spanò M, Toft G, Hagmar L, Eleuteri P, Rescia M, Rignell-Hydbom A, et al. 2005. Exposure to PCB and *p,p′*-DDE in European and Inuit populations: impact on human sperm chromatin integrity. Hum Reprod 20(12):3488–3499, PMID: 16223788, https://doi.org/10.1093/humrep/dei297.
51. Christensen LH, Høyer BB, Pedersen HS, Zinchuk A, Jönsson BAG, Lindh C, et al. 2016. Prenatal smoking exposure, measured as maternal serum cotinine, and children's motor developmental milestones and motor function: a follow-up study. Neurotoxicology 53:236–245, PMID: 26899398, https://doi.org/10.1016/j.neuro.2016.02.007.
52. Richthoff J, Rylander L, Jönsson BA, Akesson H, Hagmar L, Nilsson-Ehle P, et al. 2003. Serum levels of 2,2′,4,4′,5,5′-hexachlorobiphenyl (CB-153) in

relation to markers of reproductive function in young males from the general Swedish population. Environ Health Perspect 111(4):409–413, PMID: 12676591, https://doi.org/10.1289/ehp.5767.
53. Jönsson BAG, Rylander L, Lindh C, Rignell-Hydbom A, Giwercman A, Toft G, et al. 2005. Inter-population variations in concentrations, determinants of and correlations between 2,2′,4,4′,5,5′-hexachlorobiphenyl (CB-153) and 1,1-dichloro-2,2-bis (*p*-chlorophenyl)-ethylene (p,p′-DDE): a cross-sectional study of 3161 men and women from Inuit and European populations. Environ Health 4:27, PMID: 16283941, https://doi.org/10.1186/1476-069X-4-27.
54. Rignell-Hydbom A, Rylander L, Giwercman A, Jönsson BA, Nilsson-Ehle P, Hagmar L. 2004. Exposure to CB-153 and p,p′-DDE and male reproductive function. Hum Reprod 19(9):2066–2075, PMID: 15284211, https://doi.org/10.1093/humrep/deh362.
55. de Jager C, Aneck-Hahn NH, Bornman MS, Farias P, Leter G, Eleuteri P, et al. 2009. Sperm chromatin integrity in DDT-exposed young men living in a malaria area in the Limpopo Province, South Africa. Hum Reprod 24(10):2429–2438, PMID: 19608568, https://doi.org/10.1093/humrep/dep249.
56. WHO. 1999. *WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction*. Cambridge, UK: Cambridge University Press.
57. Evenson DP, Larson KL, Jost LK. 2002. Sperm chromatin structure assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques. J Androl 23(1):25–43, PMID: 11780920, https://doi.org/10.1002/j.1939-4640.2002.tb02599.x.
58. Tillaut H, Monfort C, Giton F, Warembourg C, Rouget F, Cordier S, et al. 2023. Persistent organic pollutant exposure and thyroid function among 12-year-old children. Neuroendocrinology 113(12):1232–1247, PMID: 36502793, https://doi.org/10.1159/000528631.
59. Chan D, Shao X, Dumargne MC, Aarabi M, Simon MM, Kwan T, et al. 2019. Customized MethylC-Capture sequencing to evaluate variation in the human sperm DNA methylome representative of altered folate metabolism. Environ Health Perspect 127(8):087002, PMID: 31393794, https://doi.org/10.1289/EHP4812.
60. Bourgey M, Dali R, Eveleigh R, Chen KC, Letourneau L, Fillon J, et al. 2019. GenPipes: an open-source framework for distributed and scalable genomic analyses. Gigascience 8(6):giz037, PMID: 31185495, https://doi.org/10.1093/gigascience/giz037.
61. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120, PMID: 24695404, https://doi.org/10.1093/bioinformatics/btu170.
62. Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics 27(11):1571–1572, PMID: 21493656, https://doi.org/10.1093/bioinformatics/btr167.
63. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359, PMID: 22388286, https://doi.org/10.1038/nmeth.1923.
64. Amemiya HM, Kundaje A, Boyle AP. 2019. The ENCODE blacklist: identification of problematic regions of the genome. Sci Rep 9(1):9354, PMID: 31249361, https://doi.org/10.1038/s41598-019-45839-z.
65. Liu Y, Siegmund KD, Laird PW, Berman BP. 2012. Bis-SNP: combined DNA methylation and SNP calling for Bisulfite-seq data. Genome Biol 13(7):R61, PMID: 22784381, https://doi.org/10.1186/gb-2012-13-7-r61.
66. Storey JD, Bass A, Dabney A, Robinson D. 2022. qvalue: Q-value estimation for false discovery rate control. http://github.com/jdstorey/qvalue [accessed 3 August 2022].
67. Shabalin AA. 2012. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics 28(10):1353–1358, PMID: 22492648, https://doi.org/10.1093/bioinformatics/bts163.
68. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57(1):289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
69. Hisano M, Erkek S, Dessus-Babus S, Ramos L, Stadler MB, Peters AHFM. 2013. Genome-wide chromatin analysis in mature mouse and human spermatozoa. Nat Protoc 8(12):2449–2470, PMID: 24232248, https://doi.org/10.1038/nprot.2013.145.
70. Krueger F. 2015. TrimGalore. a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. https://github.com/FelixKrueger/TrimGalore [accessed 3 August 2023].
71. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078–2079, PMID: 19505943, https://doi.org/10.1093/bioinformatics/btp352.
72. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42(Web Server issue):W187–W191, PMID: 24799436, https://doi.org/10.1093/nar/gku365.
73. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1):139–140, PMID: 19910308, https://doi.org/10.1093/bioinformatics/btp616.
74. Gao T, Qian J. 2020. EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. Nucleic Acids Res 48(D1):D58–D64, PMID: 31740966, https://doi.org/10.1093/nar/gkz980.
75. Lanciano S, Cristofari G. 2020. Measuring and interpreting transposable element expression. Nat Rev Genet 21(12):721–736, PMID: 32576954, https://doi.org/10.1038/s41576-020-0251-y.
76. Alexa A, Rahnenführer J, Lengauer T. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22(13):1600–1607, PMID: 16606683, https://doi.org/10.1093/bioinformatics/btl140.
77. Cavalcante RG, Sartor MA. 2017. annotatr: genomic regions in context. Bioinformatics 33(15):2381–2383, PMID: 28369316, https://doi.org/10.1093/bioinformatics/btx183.
78. Gel B, Díez-Villanueva A, Serra E, Buschbeck M, Peinado MA, Malinverni R. 2016. regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics 32(2):289–291, PMID: 26424858, https://doi.org/10.1093/bioinformatics/btv562.
79. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, et al. 2014. The DNA methylation landscape of human early embryos. Nature 511(7511):606–610, PMID: 25079557, https://doi.org/10.1038/nature13544.
80. Xia W, Xu J, Yu G, Yao G, Xu K, Ma X, et al. 2019. Resetting histone modifications during human parental-to-zygotic transition. Science 365(6451):353–360, PMID: 31273069, https://doi.org/10.1126/science.aaw5118.
81. Lun ATL, Smyth GK. 2016. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res 44(5):e45, PMID: 26578583, https://doi.org/10.1093/nar/gkv1191.
82. Wickham H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer Publishing Co.
83. Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14(2):178–192, PMID: 22517427, https://doi.org/10.1093/bib/bbs017.
84. Bonde JP, Toft G, Rylander L, Rignell-Hydbom A, Giwercman A, Spano M, et al. 2008. Fertility and markers of male reproductive function in Inuit and European populations spanning large contrasts in blood levels of persistent organochlorines. Environ Health Perspect 116(3):269–277, PMID: 18335090, https://doi.org/10.1289/ehp.10700.
85. Toft G, Axmon A, Giwercman A, Thulstrup AM, Rignell-Hydbom A, Pedersen HS, et al. 2005. Fertility in four regions spanning large contrasts in serum levels of widespread persistent organochlorines: a cross-sectional study. Environ Health 4:26, PMID: 16280075, https://doi.org/10.1186/1476-069X-4-26.
86. Niedzwiecki MM, Hall MN, Liu X, Oka J, Harper KN, Slavkovich V, et al. 2013. A dose–response study of arsenic exposure and global methylation of peripheral blood mononuclear cell DNA in Bangladeshi adults. Environ Health Perspect 121(11–12):1306–1312, PMID: 24013868, https://doi.org/10.1289/ehp.1206421.
87. Zhang Y, Schöttker B, Florath I, Stock C, Butterbach K, Holleczek B, et al. 2016. Smoking-associated DNA methylation biomarkers and their predictive value for all-cause and cardiovascular mortality. Environ Health Perspect 124(1):67–74, PMID: 26017925, https://doi.org/10.1289/ehp.1409020.
88. Nishikura K. 2010. Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79:321–349, PMID: 20192758, https://doi.org/10.1146/annurev-biochem-060208-105251.
89. Paz N, Levanon EY, Amariglio N, Heimberger AB, Ram Z, Constantini S, et al. 2007. Altered adenosine-to-inosine RNA editing in human cancer. Genome Res 17(11):1586–1595, PMID: 17908822, https://doi.org/10.1101/gr.6493107.
90. Akbarian S, Smith MA, Jones EG. 1995. Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia. Brain Res 699(2):297–304, PMID: 8616634, https://doi.org/10.1016/0006-8993(95)00922-d.
91. Cao M, Shao X, Chan P, Cheung W, Kwan T, Pastinen T, et al. 2020. High-resolution analyses of human sperm dynamic methylome reveal thousands of novel age-related epigenetic alterations. Clin Epigenetics 12(1):192, PMID: 33317634, https://doi.org/10.1186/s13148-020-00988-1.
92. Reed ML, Leff SE. 1994. Maternal imprinting of human *SNRPN*, a gene deleted in Prader–Willi syndrome. Nat Genet 6(2):163–167, PMID: 7512861, https://doi.org/10.1038/ng0294-163.
93. Feinberg JI, Bakulski KM, Jaffe AE, Tryggvadottir R, Brown SC, Goldman LR, et al. 2015. Paternal sperm DNA methylation associated with early signs of autism risk in an autism-enriched cohort. Int J Epidemiol 44(4):1199–1210, PMID: 25878217, https://doi.org/10.1093/ije/dyv028.
94. Cheroni C, Caporale N, Testa G. 2020. Autism spectrum disorder at the crossroad between genes and environment: contributions, convergences, and interactions in ASD developmental pathophysiology. Mol Autism 11(1):69, PMID: 32912338, https://doi.org/10.1186/s13229-020-00370-1.
95. WHO, United Nations Environment Program, Inter-Organization Programme for the Sound Management of Chemicals, Åke B, Heindel JJ, Jobling S, et al. 2013. *State of the Science of Endocrine Disrupting Chemicals 2012: Summary*

for Decision-Makers. https://apps.who.int/iris/handle/10665/78102 [accessed 3 August 2023].
96. Chevrier J, Di Lenardo T, Pillet S, Bornman R, Obida M, Ward B. 2019. Prenatal exposure to DDT and pyrethroids and humoral response to vaccines among South African children from an area sprayed for malaria control: the VHEMBE study. Environ Epidemiol 3:68, https://doi.org/10.1097/01.EE9.0000606408.64332.71.
97. Huq F, Obida M, Bornman R, Di Lenardo T, Chevrier J. 2020. Associations between prenatal exposure to DDT and DDE and allergy symptoms and diagnoses in the Venda Health Examination of Mothers, Babies and their Environment (VHEMBE), South Africa. Environ Res 185:109366, PMID: 32299029, https://doi.org/10.1016/j.envres.2020.109366.
98. Smit LAM, Lenters V, Høyer BB, Lindh CH, Pedersen HS, Liermontova I, et al. 2015. Prenatal exposure to environmental chemical contaminants and asthma and eczema in school-age children. Allergy 70(6):653–660, PMID: 25753462, https://doi.org/10.1111/all.12605.
99. Kok Grouleff M, Wielsøe M, Berthelsen D, Mulvad G, Isidor S, Long M, et al. 2021. Anthropometric measures and blood pressure of Greenlandic preschool children. Int J Circumpolar Health 80(1):1954382, PMID: 34291707, https://doi.org/10.1080/22423982.2021.1954382.
100. Greenberg MVC, Bourc'his D. 2019. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol 20(10):590–607, PMID: 31399642, https://doi.org/10.1038/s41580-019-0159-6.
101. Wang L, Zhang J, Duan J, Gao X, Zhu W, Lu X, et al. 2014. Programming and inheritance of parental DNA methylomes in mammals. Cell 157(4):979–991, PMID: 24813617, https://doi.org/10.1016/j.cell.2014.04.017.
102. Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, et al. 2018. Single-cell DNA methylome sequencing of human preimplantation embryos. Nat Genet 50(1):12–19, PMID: 29255258, https://doi.org/10.1038/s41588-017-0007-6.
103. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. 2012. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484(7394):339–344, PMID: 22456710, https://doi.org/10.1038/nature10960.
104. Reik W, Collick A, Norris ML, Barton SC, Surani MA. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. Nature 328(6127):248–251, PMID: 3600805, https://doi.org/10.1038/328248a0.
105. Barakat TS, Halbritter F, Zhang M, Rendeiro AF, Perenthaler E, Bock C, et al. 2018. Functional dissection of the enhancer repertoire in human embryonic stem cells. Cell Stem Cell 23(2):276–288.e8, PMID: 30033119, https://doi.org/10.1016/j.stem.2018.06.014.
106. Tang WWC, Dietmann S, Irie N, Leitch HG, Floros VI, Bradshaw CR, et al. 2015. A unique gene regulatory network resets the human germline epigenome for development. Cell 161(6):1453–1467, PMID: 26046444, https://doi.org/10.1016/j.cell.2015.04.053.
107. Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, Kutter C, et al. 2012. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell 148(1–2):335–348, PMID: 22244452, https://doi.org/10.1016/j.cell.2011.11.058.
108. Ding R, Jin Y, Liu X, Ye H, Zhu Z, Zhang Y, et al. 2017. Dose- and time- effect responses of DNA methylation and histone H3K9 acetylation changes induced by traffic-related air pollution. Sci Rep 7:43737, PMID: 28256616, https://doi.org/10.1038/srep43737.
109. Gamallat Y, Fang X, Mai H, Liu X, Li H, Zhou P, et al. 2021. Bi-allelic mutation in *Fsip1* impairs acrosome vesicle formation and attenuates flagellogenesis in mice. Redox Biol 43:101969, PMID: 33901807, https://doi.org/10.1016/j.redox.2021.101969.
110. Severinsen JE, Bjarkam CR, Kiaer-Larsen S, Olsen IM, Nielsen MM, Blechingberg J, et al. 2006. Evidence implicating *BRD1* with brain development and susceptibility to both schizophrenia and bipolar affective disorder. Mol Psychiatry 11(12):1126–1138, PMID: 16924267, https://doi.org/10.1038/sj.mp.4001885.
111. Chuong EB, Rumi MAK, Soares MJ, Baker JC. 2013. Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat Genet 45(3):325–329, PMID: 23396136, https://doi.org/10.1038/ng.2553.
112. Jacques PÉ, Jeyakani J, Bourque G. 2013. The majority of primate-specific regulatory sequences are derived from transposable elements. PLoS Genet 9(5):e1003504, PMID: 23675311, https://doi.org/10.1371/journal.pgen.1003504.
113. Lynch VJ, Nnamani MC, Kapusta A, Brayer K, Plaza SL, Mazur EC, et al. 2015. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. Cell Rep 10(4):551–561, PMID: 25640180, https://doi.org/10.1016/j.celrep.2014.12.052.
114. Lynch VJ, Leclerc RD, May G, Wagner GP. 2011. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. Nat Genet 43(11):1154–1159, PMID: 21946353, https://doi.org/10.1038/ng.917.
115. Veselovska L, Smallwood SA, Saadeh H, Stewart KR, Krueger F, Maupetit-Méhouas S, et al. 2015. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. Genome Biol 16:209, PMID: 26408185, https://doi.org/10.1186/s13059-015-0769-z.
116. Mochida GH, Mahajnah M, Hill AD, Basel-Vanagaite L, Gleason D, Hill RS, et al. 2009. A truncating mutation of *TRAPPC9* is associated with autosomal-recessive intellectual disability and postnatal microcephaly. Am J Hum Genet 85(6):897–902, PMID: 20004763, https://doi.org/10.1016/j.ajhg.2009.10.027.
117. Ito J, Sugimoto R, Nakaoka H, Yamada S, Kimura T, Hayano T, et al. 2017. Systematic identification and characterization of regulatory elements derived from human endogenous retroviruses. PLoS Genet 13(7):e1006883, PMID: 28700586, https://doi.org/10.1371/journal.pgen.1006883.
118. Allen N, Walker RL, Fife LC, Chisholm RD, Koblitsky L, Bullock JF, et al. 1954. Persistence of BHC, DDT, and toxaphene in soil and the tolerances of certain crops to their residues. Technical Bulletin 156658. Washington, DC: U.S. Department of Agriculture, Economic Research Service. https://doi.org/10.22004/ag.econ.156658.
119. Muñoz-Arnanz J, Jiménez B. 2011. New DDT inputs after 30 years of prohibition in Spain. A case study in agricultural soils from south-western Spain. Environ Pollut 159(12):3640–3646, PMID: 21864957, https://doi.org/10.1016/j.envpol.2011.07.027.
120. Vieira ED, Torres JP, Malm O. 2001. DDT environmental persistence from its use in a vector control program: a case study. Environ Res 86(2):174–182, PMID: 11437464, https://doi.org/10.1006/enrs.2001.4258.
121. van den Berg H. 2009. Global status of DDT and its alternatives for use in vector control to prevent disease. Environ Health Perspect 117(11):1656–1663, PMID: 20049114, https://doi.org/10.1289/ehp.0900785.
122. Cocco P. 2002. On the rumors about the silent spring: review of the scientific evidence linking occupational and environmental pesticide exposure to endocrine disruption health effects. Cad Saude Publica 18(2):379–402, PMID: 11923880, https://doi.org/10.1590/s0102-311x2002000200003.
123. Son HK, Kim SA, Kang JH, Chang YS, Park SK, Lee SK, et al. 2010. Strong associations between low-dose organochlorine pesticides and type 2 diabetes in Korea. Environ Int 36(5):410–414, PMID: 20381150, https://doi.org/10.1016/j.envint.2010.02.012.
124. Venners SA, Korrick S, Xu X, Chen C, Guang W, Huang A, et al. 2005. Preconception serum DDT and pregnancy loss: a prospective study using a biomarker of pregnancy. Am J Epidemiol 162(8):709–716, PMID: 16120699, https://doi.org/10.1093/aje/kwi275.
125. Bornman MS, Chevrier J, Rauch S, Crause M, Obida M, Sathyanarayana S, et al. 2016. Dichlorodiphenyltrichloroethane exposure and anogenital distance in the Venda Health Examination of Mothers, Babies and their Environment (VHEMBE) birth cohort study, South Africa. Andrology 4(4):608–615, PMID: 27457477, https://doi.org/10.1111/andr.12235.
126. Skvortsova K, Iovino N, Bogdanović O. 2018. Functions and mechanisms of epigenetic inheritance in animals. Nat Rev Mol Cell Biol 19(12):774–790, PMID: 30425324, https://doi.org/10.1038/s41580-018-0074-2.
127. Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, et al. 2016. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science 351(6271):391–396, PMID: 26721685, https://doi.org/10.1126/science.aad6780.
128. Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, et al. 2016. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. Science 351(6271):397–400, PMID: 26721680, https://doi.org/10.1126/science.aad7977.
129. Ly L, Chan D, Aarabi M, Landry M, Behan NA, MacFarlane AJ, et al. 2017. Intergenerational impact of paternal lifetime exposures to both folic acid deficiency and supplementation on reproductive outcomes and imprinted gene methylation. Mol Hum Reprod 23(7):461–477, PMID: 28535307, https://doi.org/10.1093/molehr/gax029.
130. Gapp K, van Steenwyk G, Germain PL, Matsushima W, Rudolph KLM, Manuella F, et al. 2020. Alterations in sperm long RNA contribute to the epigenetic inheritance of the effects of postnatal trauma. Mol Psychiatry 25(9):2162–2174, PMID: 30374190, https://doi.org/10.1038/s41380-018-0271-6.
131. Anway MD, Skinner MK. 2006. Epigenetic transgenerational actions of endocrine disruptors. Endocrinology 147(6 suppl):S43–S49, PMID: 16690803, https://doi.org/10.1210/en.2005-1058.
132. Soubry A, Hoyo C, Jirtle RL, Murphy SK. 2014. A paternal environmental legacy: evidence for epigenetic inheritance through the male germ line. Bioessays 36(4):359–371, PMID: 24431278, https://doi.org/10.1002/bies.201300113.
133. Toft G, Rignell-Hydbom A, Tyrkiel E, Shvets M, Giwercman A, Lindh CH, et al. 2006. Semen quality and exposure to persistent organochlorine pollutants. Epidemiology 17(4):450–458, PMID: 16755259, https://doi.org/10.1097/01.ede.0000221769.41028.d2.
134. Aston KI, Uren PJ, Jenkins TG, Horsager A, Cairns BR, Smith AD, et al. 2015. Aberrant sperm DNA methylation predicts male fertility status and embryo quality. Fertil Steril 104(6):1388–1397.e5, PMID: 26361204, https://doi.org/10.1016/j.fertnstert.2015.08.019.
135. Caporale N, Leemans M, Birgersson L, Germain PL, Cheroni C, Borbély G, et al. 2022. From cohorts to molecules: adverse impacts of endocrine disrupting

mixtures. Science 375(6582):eabe8244, PMID: 35175820, https://doi.org/10.1126/science.abe8244.
136. Brown AS, Cheslack-Postava K, Rantakokko P, Kiviranta H, Hinkka-Yli-Salomäki S, McKeague IW, et al. 2018. Association of maternal insecticide levels with autism in offspring from a national birth cohort. Am J Psychiatry 175(11):1094–1101, PMID: 30111184, https://doi.org/10.1176/appi.ajp.2018.17101129.
137. Corbin JM, Ruiz-Echevarría MJ. 2016. One-carbon metabolism in prostate cancer: the role of androgen signaling. Int J Mol Sci 17(8):1208, PMID: 27472325, https://doi.org/10.3390/ijms17081208.
138. Arguelles LM, Liu X, Venners SA, Ronnenberg AG, Li Z, Yang F, et al. 2009. Serum folate and DDT isomers and metabolites are inversely associated in Chinese women: a cross-sectional analysis. J Am Coll Nutr 28(4):380–387, PMID: 20368376, https://doi.org/10.1080/07315724.2009.10718100.
139. Ouyang F, Longnecker MP, Venners SA, Johnson S, Korrick S, Zhang J, et al. 2014. Preconception serum 1,1,1-trichloro-2,2,bis(*p*-chlorophenyl)ethane and B-vitamin status: independent and joint effects on women's reproductive outcomes. Am J Clin Nutr 100(6):1470–1478, PMID: 25411282, https://doi.org/10.3945/ajcn.114.088377.
140. Liu S, Brind'Amour J, Karimi MM, Shirane K, Bogutz A, Lefebvre L, et al. 2014. *Setdb1* is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. Genes Dev 28(18):2041–2055, PMID: 25228647, https://doi.org/10.1101/gad.244848.114.
141. Goodier JL. 2016. Restricting retrotransposons: a review. Mob DNA 7:16, PMID: 27525044, https://doi.org/10.1186/s13100-016-0070-z.
142. Senft AD, Macfarlan TS. 2021. Transposable elements shape the evolution of mammalian development. Nat Rev Genet 22(11):691–711, PMID: 34354263, https://doi.org/10.1038/s41576-021-00385-1.
143. Diehl AG, Ouyang N, Boyle AP. 2020. Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes. Nat Commun 11(1):1796, PMID: 32286261, https://doi.org/10.1038/s41467-020-15520-5.
144. Küry P, Nath A, Créange A, Dolei A, Marche P, Gold J, et al. 2018. Human endogenous retroviruses in neurological diseases. Trends Mol Med 24(4):379–394, PMID: 29551251, https://doi.org/10.1016/j.molmed.2018.02.007.
145. Rodriguez-Terrones D, Torres-Padilla ME. 2018. Nimble and ready to mingle: transposon outbursts of early development. Trends Genet 34(10):806–820, PMID: 30057183, https://doi.org/10.1016/j.tig.2018.06.006.
146. Göke J, Lu X, Chan YS, Ng HH, Ly LH, Sachs F, et al. 2015. Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. Cell Stem Cell 16(2):135–141, PMID: 25658370, https://doi.org/10.1016/j.stem.2015.01.005.
147. Zeng L, Wang M, Zhou J, Wang X, Zhang Y, Su P. 2022. A hypothesis: retrotransposons as a relay of epigenetic marks in intergenerational epigenetic inheritance. Gene 817:146229, PMID: 35063571, https://doi.org/10.1016/j.gene.2022.146229.
148. Guerrero-Bosagna C, Settles M, Lucker B, Skinner MK. 2010. Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. PLoS One 5(9):13100, PMID: 20927350, https://doi.org/10.1371/journal.pone.0013100.
149. Heard E, Martienssen RA. 2014. Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157(1):95–109, PMID: 24679529, https://doi.org/10.1016/j.cell.2014.02.045.
150. Skinner MK, Manikkam M, Tracey R, Guerrero-Bosagna C, Haque M, Nilsson EE. 2013. Ancestral dichlorodiphenyltrichloroethane (DDT) exposure promotes epigenetic transgenerational inheritance of obesity. BMC Med 11:228, PMID: 24228800, https://doi.org/10.1186/1741-7015-11-228.
151. Oluwayiose OA, Marcho C, Wu H, Houle E, Krawetz SA, Suvorov A, et al. 2021. Paternal preconception phthalate exposure alters sperm methylome and embryonic programming. Environ Int 155:106693, PMID: 34120004, https://doi.org/10.1016/j.envint.2021.106693.
152. Guo JU, Su Y, Zhong C, Ming GL, Song H. 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell 145(3):423–434, PMID: 21496894, https://doi.org/10.1016/j.cell.2011.03.022.
153. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, et al. 2011. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473(7347):394–397, PMID: 21552279, https://doi.org/10.1038/nature10102.
154. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 2011. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun 2:241, PMID: 21407207, https://doi.org/10.1038/ncomms1240.
155. Bocker MT, Tuorto F, Raddatz G, Musch T, Yang FC, Xu M, et al. 2012. Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian *HOXA* cluster. Nat Commun 3:818, PMID: 22569366, https://doi.org/10.1038/ncomms1826.
156. Rosenquist HAH, Høyer BB, Julvez J, Sunyer J, Pedersen HS, Lenters V, et al. 2017. Prenatal and postnatal PCB-153 and *p,p′*-DDE exposures and behavior scores at 5–9 years of age among children in Greenland and Ukraine. Environ Health Perspect 125(10):107002, PMID: 28974479, https://doi.org/10.1289/EHP553.
157. IARC (International Agency for Research on Cancer). 2018. *DDT, Lindane, and 2,4-D*. Lyon, France: IARC.
158. Coker E, Chevrier J, Rauch S, Bradman A, Obida M, Crause M, et al. 2018. Association between prenatal exposure to multiple insecticides and child body weight and body composition in the VHEMBE South African birth cohort. Environ Int 113:122–132, PMID: 29421401, https://doi.org/10.1016/j.envint.2018.01.016.
159. van den Berg H. 2010. DDT and malaria prevention: van den Berg responds. Environ Health Perspect 118(1):A15–A16, https://doi.org/10.1289/ehp.0901276R.
160. Simonich SL, Hites RA. 1995. Global distribution of persistent organochlorine compounds. Science 269(5232):1851–1854, PMID: 7569923, https://doi.org/10.1126/science.7569923.
161. Teran T, Lamon L, Marcomini A. 2012. Climate change effects on POPs' environmental behaviour: a scientific perspective for future regulatory actions. Atmos Pollut Res 3(4):466–476, https://doi.org/10.5094/APR.2012.054.