

## ARTICLE OPEN



# Exposure to childhood maltreatment is associated with specific epigenetic patterns in sperm

Jetro J. Tuulari<sup>1,2,3,4</sup> <sup>✉</sup>, Matthieu Bourgery<sup>4,5</sup> , Jo Iversen<sup>6</sup> , Thomas Gade Koefoed<sup>6</sup>, Annukka Ahonen<sup>1</sup>, Ammar Ahmedani<sup>5</sup>, Eeva-Leena Kataja<sup>1,4</sup>, Linnea Karlsson<sup>1,4,7</sup> , Romain Barrès<sup>6,8,9</sup>, Hasse Karlsson<sup>1,2,4,9</sup> and Noora Kotaja<sup>5,9</sup>

© The Author(s) 2025

Childhood maltreatment exposure (CME) increases the risk of adverse long-term health consequences for the exposed individual. Animal studies suggest that CME may also influence the health and behaviour in the next generation offspring through CME-driven epigenetic changes in the germ line. Here we investigated the association between early life stress on the epigenome of sperm in humans with history of CME. We measured paternal CME using the Trauma and Distress Scale (TADS) questionnaire and mapped sperm-borne sncRNAs expression by small RNA sequencing (small RNA-seq) and DNA methylation (DNAm) in spermatozoa by reduced-representation bisulfite sequencing (RRBS-seq) in males from the FinnBrain Birth Cohort Study. The study design was a (nested) case-control study, high-TADS (TADS  $\geq 39$ ,  $n = 25$  for DNAm and  $n = 14$  for small RNA-seq) and low-TADS (TADS  $\leq 10$ ,  $n = 30$  for DNAm and  $n = 16$  for small RNA-seq). We identified 3 genomic regions with differential methylation between low and high-TADS and 68 tRNA-derived small RNAs (tsRNAs) and miRNAs with different levels in males with high CME (False discovery rate, FDR corrected  $p < 0.05$ ). Of potential interest, we identified differential expression of miRNA hsa-mir-34c-5p and differential methylation levels near the *CRTC1* and *GBX2* genes, which are documented to control brain development. Our results provide further evidence that early life stress influences the paternal germline epigenome and supports a possible effect in modulating the development of the central nervous system of the next generation.

*Molecular Psychiatry* (2025) 30:2635–2644; <https://doi.org/10.1038/s41380-024-02872-3>

## INTRODUCTION

Adverse childhood experiences (ACEs) include harms that affect children indirectly through their living environments (e.g., parental conflict, substance abuse, or mental illness) or directly (abuse and neglect). The direct harms are commonly described as childhood maltreatment exposure (CME). CME is highly prevalent, as shown by a recent systematic review and meta-analysis that reported a pooled prevalence of ca. 23% in Europe and the U.S. for adults who reported at least one ACE [1]. Worldwide, as many as 12% of adults report a history of childhood sexual abuse, 23% of childhood physical abuse, and 36% of emotional abuse [2]. ACEs have numerous adverse consequences for later health, via a range of hormonal, metabolic, and immunological pathways [3], especially for mental health outcomes [2, 4]. In addition to affecting health later in life [5], accumulating evidence indicates that paternal ACEs /CME may also affect the health of the next generation [6–10].

Animal studies of paternal inheritance induced by early life stress have shown that adverse psychological exposures change the epigenetic marks in sperm and the metabolic and behavioural phenotype of the offspring [11–15]. The commonly recognized

epigenetic marks are DNA methylation, histone modifications, and expression of small non-coding RNAs (sncRNA) [16–19]. When carried in gametes, these epigenetic marks have the potential to change the early embryonic developmental trajectory and affect offspring phenotype. Numerous studies have identified a link between paternal exposure to a plethora of physical environmental exposures like toxins, cigarette smoking, physical activity, and nutritional stress and changes in sperm epigenome or altered offspring phenotype [16, 20]. However, the association between early-life psychological stress and epigenetic changes in spermatozoa remains unclear.

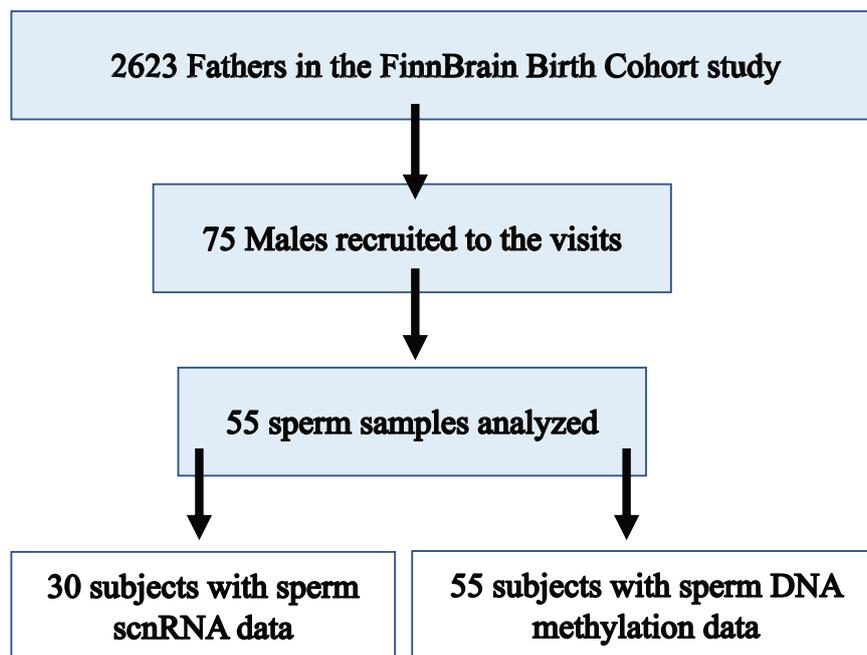
To our knowledge, three human studies have established a link between paternal early-life stress and changes in the sperm epigenome have been published to date. In the first, childhood Trauma Questionnaire (CTQ) and Conflict Tactics Scales (CTS) were used to quantify CME in 34 males (17 men exposed to high, 5 men to medium, and 12 men to no childhood abuse), an association was found between DNA methylation and CTQ/composite abuse score [8]. In this cohort, 12 DNA regions were found differentially methylated between individuals with different childhood abuse (high vs. low), including genes associated with neuronal function

<sup>1</sup>FinnBrain Birth Cohort Study, Turku Brain and Mind Center, Department of Clinical Medicine, University of Turku, Turku, Finland. <sup>2</sup>Department of Psychiatry, Turku University Hospital and University of Turku, Turku, Finland. <sup>3</sup>Turku Collegium for Science and Medicine, University of Turku, Turku, Finland. <sup>4</sup>Centre for Population Health Research, Turku University Hospital and University of Turku, Turku, Finland. <sup>5</sup>Institute of Biomedicine, Integrative Physiology and Pharmacology Unit, University of Turku, Turku, Finland. <sup>6</sup>Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. <sup>7</sup>Department of Paediatrics and Adolescent Medicine, Turku University Hospital and University of Turku, Turku, Finland. <sup>8</sup>Institut de Pharmacologie Moléculaire et Cellulaire, Université Côte d'Azur & Centre National pour la Recherche Scientifique (CNRS), Valbonne, France. <sup>9</sup>These authors jointly supervised this work: Romain Barrès, Hasse Karlsson, Noora Kotaja.

✉email: jetro.tuulari@utu.fi

Received: 20 December 2023 Revised: 25 November 2024 Accepted: 6 December 2024

Published online: 3 January 2025



**Fig. 1 Flowchart of the participant selection.** The participants were a subsample from FinnBrain Birth cohort study and the flowchart shows the participant selection that for the current study was based on low or high CME questionnaire scores.

(*MAPT*, *CLU*), fat cell regulation (*PRDM16*), and immune function (*SDK1*) [8]. In the second study, CME was quantified using an adverse childhood experiences (ACEs) screen; and comparing a group of individuals with the highest ACE score to individuals with the lowest score, a negative correlation was found between levels of multiple miRNAs of the miR-449/34 family and ACE scores [6]. In the third study (preprint at the time of writing), a longitudinal approach that used three different age groups, sperm miR-16 and miR-375 levels were higher in the serum of children exposed to paternal loss and maternal separation (ages 7–12 years) compared to controls [21]. This study also showed similar patterns in another cohort of 18–25-year-olds and, using the CTQ to quantify ACEs, report that the same miRNA species have lower expression in sperm of adult men exposed to higher CME at ages 21–50 years [21]. Interestingly, they also replicate prior findings [6] on lower levels of miR-34 but not miR-449.

In the current study, we measured CME using the Trauma and Distress Scale (TADS) questionnaire and quantified sperm sncRNA profiles and DNAm from males that were divided into 2 groups, a control group with low TADS scores ( $TADS \leq 10$ ) and a case group with high TADS scores ( $TADS \geq 39$ ). Following the primary analyses, we also performed replication analyses to prior work where possible. This was an exploratory study and no a priori hypotheses on the strength or direction of the possible associations between CME and sperm DNA methylation and sncRNA profiles.

## MATERIAL AND METHODS

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Hospital District of Southwest Finland (15.3.2011 §95, ETMK: 31/180/2011). We followed the Strengthening the Reporting of Observational Studies in Epidemiology (<https://www.strobe-statement.org/>) reporting guideline (case-control studies v4). All participants signed an informed consent for participating to the studies.

### Participants

Participants were recruited at gestational week (gwk) 12 from maternity clinics in Southwest Finland from 2011 to 2015 to take part in the FinnBrain

Birth Cohort Study (<http://www.finnbrain.fi>), which was established to prospectively investigate the effects of early life stress, including prenatal stress exposure, on child brain development and health. The cohort entails 3808 families and included full trios (mother, father, child) on approximately half of the families [22]. The fathers of the cohort were later recruited separately for these visits when their children were approximately 9 years of age.

The division into case and control groups was based on previously collected data. CME was assessed by questionnaires filled in by the fathers-to-be following the initial recruitment approximately at gwk 14 (2011–2015). There was no case - control matching during recruitment although the approach and analyses in the current study are based on case - control design. The new data were collected during visits that took place between 02/2019 – 07/2021. The recruitment rate for the visits has been 59.6%, which is typical for our cohort study. Until the time of writing, this has been a cross-sectional study.

Altogether 75 males participated in the current study. The final sample size was determined by the available resources and comparison to prior studies with similar design. We processed sperm samples from 58 individuals with TADS scores ranging between 0 and 78. 55 samples were analyzed by RRBS-seq to identify DNA methylation patterns, and 30 were analyzed by small RNA-seq (27 samples overlapping). The flowchart of participant selection is presented in Fig. 1. The personnel carrying out the sperm epigenome analyses were blind to the group allocation.

The bias assessment is as follows: there was no systematic monitoring for selection bias during recruitment as the primary recruitment for the cohort entailed a population-representative sample and was fairly uniform on e.g., for participants' age range; we strived to quantify the influence of potential confounding factors by including multivariate statistical tests; questionnaire-based measurements are prone to measurement error and TADS questionnaire to recall bias, but these limitations apply to all similar studies; finally we performed the processing of semen samples with utmost care and where possible assured their quality to reduce measurement errors.

### The demographics and questionnaire data

Data were mostly obtained over the prenatal period and data collection was continued postnatally. Information about the parents' CME was collected using the Trauma and Distress Scale (TADS) at gwk 14 [23]. The TADS comprises five core domains: emotional neglect, emotional abuse, physical neglect, physical abuse, and sexual abuse. In this study, we calculated the cumulative exposure to early life stress events of the infants' fathers and mothers by the age of 18 years (direct sum scores). Of note, the

TADS scores have two possible derived values so that one can use direct sum scores or sums of factor scores [23]. Here we used the direct TADS sum scores as the primary variable to perform case-control divisions and statistical testing, but we have used the factor sum scores in our prior article [24]. The direct and factor sum scores were highly correlated (in our sample Spearman's  $\rho = 0.954$ ,  $p < 0.001$ ).

Depressive symptoms were assessed by implementing the Edinburgh Postnatal Depression Scale (EPDS) [25]. This 10-item questionnaire is scaled from 0 to 30 points with a bigger score denoting increased symptom severity. Anxiety symptoms were quantified with the anxiety subscale of Symptom Checklist 90 (SCL-90) [26]. The SCL-90 anxiety subscale consists of 10 items with a total score range of 0 to 40 with larger scores denoting increased symptom severity. Alcohol use was quantified with a structured questionnaire [27], and the derived variable we chose was mean alcohol use per week (1 unit = 12 g of alcohol). Smoking was quantified with a yes/no binary variable. BMI was obtained from height and weight measurements performed during the research visits that are described below.

### The study visits

Participants that matched the case-control criteria (control group TADS  $\leq 10$ ; and case group TADS  $\geq 39$ ) were invited to a separate study visit. During the visit fathers filled in questionnaires for smoking and alcohol use habits, depressive symptoms (EPDS), and anxiety symptoms (SCL-90). They also gave additional biological samples (not reported here). Weight and height were measured for defining body mass index (BMI), and waist circumference was measured following standard procedures. Sperm samples were collected before or at the visit.

### The sperm sample collection

Participants were allowed two options for delivering the semen samples following 2–7-day abstinence from ejaculation: either collecting the sample at home and delivering it to the research site at the start of the visit or giving the sample during the visit. The semen was collected by masturbation. Obtained samples were incubated at  $+37^\circ\text{C}$  for 5–30 min for liquefaction, and spermatozoa were purified by centrifuging through 50% Puresperm (Nidacon) solution at  $400 \times g$  for 15 min. The sperm pellet was subsequently washed with mild somatic cell lysis buffer (0.01% SDS, 0.005% Triton X-100) to eliminate remaining somatic cell contamination. The purity of the samples was assessed by light microscopic analysis. The total number of spermatozoa before purification ranged between 40 and 800 million (only one high-TADS sample had less than 40 million spermatozoa), and after purification, all samples were pure with only very minor somatic cell contamination (Supplementary Fig. 1A). Sperm purification was conducted promptly within the same day and purified samples were frozen for storage.

### Small RNA sequencing

**Sample preparation.** Total RNA was extracted from 14 high-TADS (TADS  $\geq 39$ ) and 16 low-TADS (TADS  $\leq 10$ ) samples (10 million spermatozoa per sample) by TRIzol LS (Invitrogen) containing tris(2-carboxyethyl) phosphine (TCEP, Sigma-Aldrich) as a reducing agent to enhance sperm nucleus lysis and precipitated with isopropanol in the presence of 2  $\mu\text{l}$  of GlycoBlue (15 mg/ml, Invitrogen). After DNaseI treatment (Sigma-Aldrich) without heat inactivation, RNA was re-extracted with TRIzol LS to remove DNaseI. The quality of the RNA sample was analyzed by Bioanalyzer (Agilent RNA 6000 Pico Kit). Bioanalyzer analysis validated the lack of somatic cell contamination, as demonstrated by the absence of ribosomal RNA peaks and low RIN values [2, 3] (Supplementary Fig. 1B). The RNA yield from 10 million spermatozoa was 10–200 ng. Libraries for small RNA-seq were prepared using NEB Next<sup>®</sup> Multiplex Small RNA kit (New England Biolabs), and the libraries were sequenced by NovaSeq 6000 system (Illumina).

**SncRNA-Seq processing.** The quality of the reads was assessed using FastQC (v0.11.9) (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Adapters and low quality reads were trimmed off using cutadapt (v3.5) [28, 29] was used to map the 15–45 nucleotides long reads first to the human genome (hg38), then subsequently to ribosomal RNAs (rRNAdb), miRNAs (miRBase v22), YRNA mapping and transfer RNAs (tRNA) (GtRNAdb v2.0) using software default settings. Sequences mapped to tRNAs were annotated as 5' end-, 3' end-, or 3' CCA end-derived according to their location on the parental tRNA sequence. All sperm samples showed typical size distribution of averaged sncRNA

reads, including a peak at 21–23 nt for miRNAs, a peak at 31–32 nt for YRNAs, a peak at 31–32 nt for tRNA-derived small RNAs (tsRNAs), and a peak at 31–33 nt for PIWI-interacting RNAs (piRNAs) (Supplementary Fig. 1C, D). All sequences mapping to rRNAs, miRNAs, and tRNAs were extracted from SPORTS1.1 as a text output file. For piRNA analysis, reads were mapped to piRNA clusters [30] in the Human genome (UCSC: Hg38) using HISAT2 (v2.1.0) [31], and assigned and counted using featureCounts (v2.0.0) [32–34] against reference gtf files. Altogether, we identified unfiltered normalized reads mapping to a total of 838 miRNAs, 266 tsRNAs, 6195 piRNA genomic clusters, 8 rRNA genes (12S-rRNA, 16S-rRNA, 18S-rRNA, 28S-rRNA, 45S-rRNA, 5.8S-rRNA, 5S-rRNA, other-rRNA) and 4 YRNA genes (YRNA1, 3, 4 and 5). Raw read counts were filtered to a minimum of 10 total counts across all 35 samples and normalized.

### Reduced Representation Bisulfite Sequencing (RRBS)

**Sample preparation.** Sperm DNA was extracted from ~10 million spermatozoa/sample by lysing cells in RLT+ buffer (Qiagen) supplied with 1% Beta-mercaptoethanol for 15 min, and then passing the lysate through AllPrep Mini Spin Columns (Qiagen) and proceeding with manufacturer's instructions. The quantification was done by measuring the DNA concentration on a Qubit fluorometer using the Qubit dsDNA High Sensitivity Assay Kit (Thermo Fischer), and was further confirmed by comparing to the quantities measured on a Nanodrop 2000 Spectrophotometer (Thermo Fischer). The quality was further assessed by looking at the 260/280 ratio on the Nanodrop, and samples were only accepted with a ratio exceeding 1.8. The reduced-representation bisulfite sequencing (RRBS-seq) was performed by the Single-Cell Omics platform at the Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen. RRBS libraries were constructed from 100 ng genomic DNA using the Ovation<sup>®</sup> RRBS Methyl-Seq library preparation kit (Tecan) according to the manufacturer's instructions. Final libraries were quantified by Qubit (Thermo Fisher) and quality checked on Bioanalyzer (Agilent). Pooled libraries were subjected to either a 101-bp single-end sequencing on a NovaSeq 6000 platform (Illumina) or a 76-bp single-end sequencing on a NextSeq 500 platform (Illumina). A total of 5.5 billion reads were generated.

**RRBS processing.** FASTQ files were first generated using bcl2fastq (v. 2.20.0) and subsequently trimmed using Trim Galore! (v. 0.6.4) as described in The Analysis Guide for NuGen Ovation RRBS Methyl-Seq (<https://github.com/nugentechnologies/NuMetRRBS>). Similarly, diversity adapter sequences were removed using the supplied trimRRBSdiversityAdaptCustomers.py script. Methylation coverage was extracted using Bismark by aligning reads to the GRCh38 assembly. Here, the deduplicate\_bismark step was performed using the --barcode option to deduplicate reads based on UMIs, and the --ignore 3 parameter was used during methylation extraction to disregard restriction enzyme sites.

### Statistical analyses

In line with prior work, we used between-group comparisons to identify features associated with CME in the sperm epigenome, comparing those with low to those with high CME. These discovery analyses were followed by multivariate analyses that included potential confounders. Third, we performed replication analyses by Dickson et al. [6]. The multivariate and replication analyses were performed with JASP 0.16.3 (<https://jasp-stats.org/>). Most variables reported in Table 1 had non-normal distributions and as this applied to the majority of sperm epigenetic variables as well, we thus used non-parametric correlation statistics in multivariate and replication analyses.

### Discovery analyses (high vs. low CME group comparisons)

Differential expression analysis of sncRNAs was performed using DESeq2 (v1.36.0) which is based on a negative binomial generalized linear model [35]. RsRNA and YRNAs were analyzed using a Wilcoxon-rank exact test. All 30 samples were split into 2 groups, low-TADS (TADS  $\leq 10$ ,  $n = 16$ ) and high-TADS (TADS  $\geq 39$ ,  $n = 14$ ). Low-TADS group was defined as the control group. Because of the small sample size, we opted to carry out the discovery analyses as a simple between groups comparison but provide multivariate analyses for all identified sncRNAs as outlined below.

For DNA methylation analysis, the methylation coverage files were analysed using the DMRichR R package (v. 1.7.8). In brief, all 55 sperm DNA samples were split into 2 groups, low TADS (TADS  $< 10$ ,  $n = 30$ ) and high TADS (TADS  $> 40$ ,  $n = 25$ ). DMRs were subsequently detected using the

**Table 1.** Descriptive statistics of demographics across the whole sample ( $N = 55$ ).

Continuous variables	Whole sample $N = 55$		Low CME $N = 26$		High CME $N = 29$		High/Low CME group difference	
	Mean	SD	Mean	SD	Mean	SD	Mann-Whitney U test	
							$p$	Rank-Biserial Correlation
Age (years)	38.764	5.910	36.828	4.699	40.583	6.593	0.054	-0.310
Body mass index (kg/m <sup>2</sup> )	26.197	4.057	25.869	3.912	26.790	4.357	0.589	-0.089
Waist circumference (cm)	92.095	11.184	92.369	11.824	91.917	11.003	0.754	0.052
EPDS score	3.505	3.317	2.303	2.436	4.750	3.391	0.005	-0.451
SCL-90 score	3.455	3.452	2.276	2.814	4.917	3.741	0.002	-0.493
Average daily alcohol consumption	0.483	0.531	0.631	0.645	0.315	0.295	0.046	0.319
TADS factor sum	13.321	14.350	1.034	1.017	28.167	6.631	<0.001	-1.000
TADS direct sum score	24.528	22.655	4.690	2.480	48.500	7.553	<0.001	-1.000
Semen sample volume (ml)	3.182	1.490	3.410	1.435	3.000	1.573	0.227	0.195
Sperm concentration 10exp6/ml	113.724	84.667	109.800	78.234	107.304	85.355	0.837	0.034
Purified sperm concentration 10exp6/ml	167.213	153.247	177.193	163.072	152.346	147.700	0.642	0.076
<b>Categorical variables (frequencies)</b>		<b>N</b>	<b>N</b>		<b>N</b>			
Educational level	low	16	4		12		0.273	1.201 <sup>a</sup>
	mid	17	11		6			
	high	22	11		11			
Relationship status	single	2	1		1		0.744	0.107 <sup>a</sup>
	married	35	15		20			
	cohabitation	15	7		8			
	separated	3	3		0			
							0.321	0.134
Smoking	Yes	19	10		9			
	No	36	16		20			

<sup>a</sup>Kruskall Wallis test statistic

DMRichR::DM.R-command. Specifically, the TADS group was supplied as the test-covariate, and the following variables were included for adjustment of possible confounding effects: BMI, age, smoking, daily alcohol consumption, summed SCL-90-R scores (assessing current psychological symptoms) summed EPDS-scores (assessing current depressive symptoms), semen volume and purified semen concentration.

All statistical tests were based on two-sided  $p$  values. Multiple comparison correction for both scRNA and DNAm analyses was carried out using Benjamini and Hochberg procedure yielding false discovery rate corrected  $p$  values [36].

#### Multivariate analyses for sncRNAs with covariates

We performed three different multivariate analyses with partial correlations testing for the association between low vs high CME: (A) controlling for semen sample volume and sperm concentration; (B) controlling for health characteristics including age, BMI, smoking (yes/no), average alcohol use per day, as well as depressive and anxiety symptoms at the time of the sperm sample collection; (C) controlling for covariates in A + B.

#### Replication analyses of prior work (miRNAs hsa-miR-34c-5p and hsa-miR-449a)

We also performed replication analyses of Dickson et al. [6] for the sncRNA data by exploring the associations between miRNA that their work implicated important and ACE exposure. For this, we looked at between-group differences (low vs. high ACE exposure; reported in Supplementary Fig. 4) and Spearman correlations between CME (TADS scores) and miRNA expression levels for hsa-miR-34c-5p and hsa-miR-449a. We chose to use the TADS factor scores here as the variable had a "more normal" distribution (correlation to TADS sum score Spearman  $\rho = 0.954$ ). We

extended these analyses also to partial correlation models where we controlled for age and BMI at measurement, and smoking (yes/no).

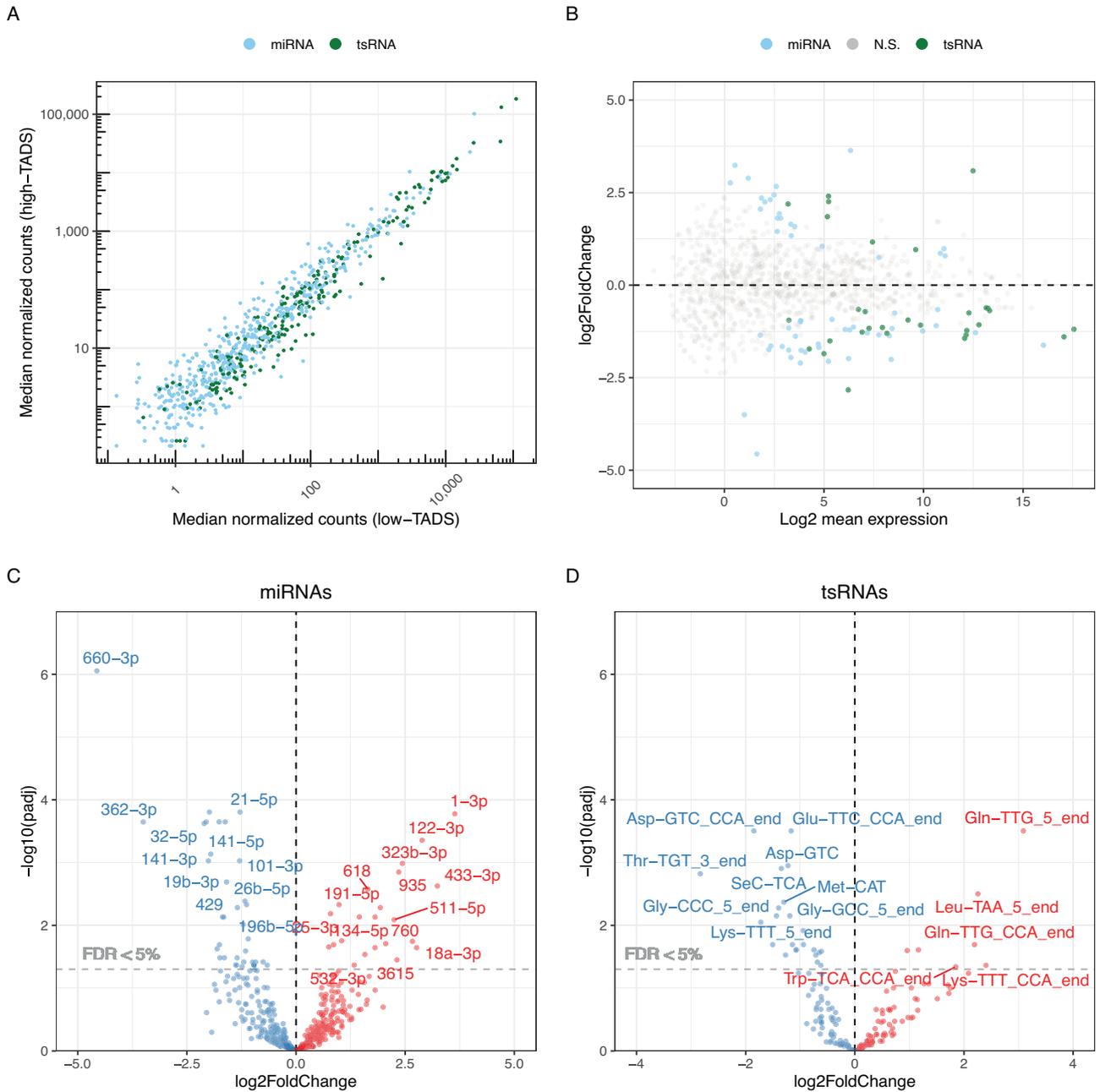
#### RESULTS

We analyzed the association between CME, as quantified by TADS scores, and two epigenetic features in sperm: the abundance of sncRNAs and levels of DNA methylation. The study enrolment is presented as a flowchart (Fig. 1), demographics are reported in Table 1, and correlations across the sperm characteristics and demographics are available in Supplementary Table 1.

#### Childhood maltreatment exposure is associated with modified sperm sncRNA profile

The analysis of small RNA-seq data from 14 high-TADS and 16 low-TADS sperm samples showed that the abundance of five analysed classes of sncRNAs was generally similar between high-TADS and low-TADS groups (Fig. 2A). The vast majority of sncRNA reads originated from rRNAs and YRNAs (Supplementary Fig. 1). The reads derived from YRNA and rRNA genes were similarly distributed in high- and low-TADS sperm samples, with some differences in the relative number of reads derived from 12S-rRNA, 16S-rRNA, 5.8-rRNA, RNY1 and RNY4 between high- and low-TADS samples (Supplementary Fig. 1E).

Differential expression analysis of individual miRNAs, tsRNAs, and piRNA clusters showed significant differences in high-TADS vs. low-TADS sperm (Fig. 2). The expression of a total of 29 miRNAs, 15 tsRNAs,

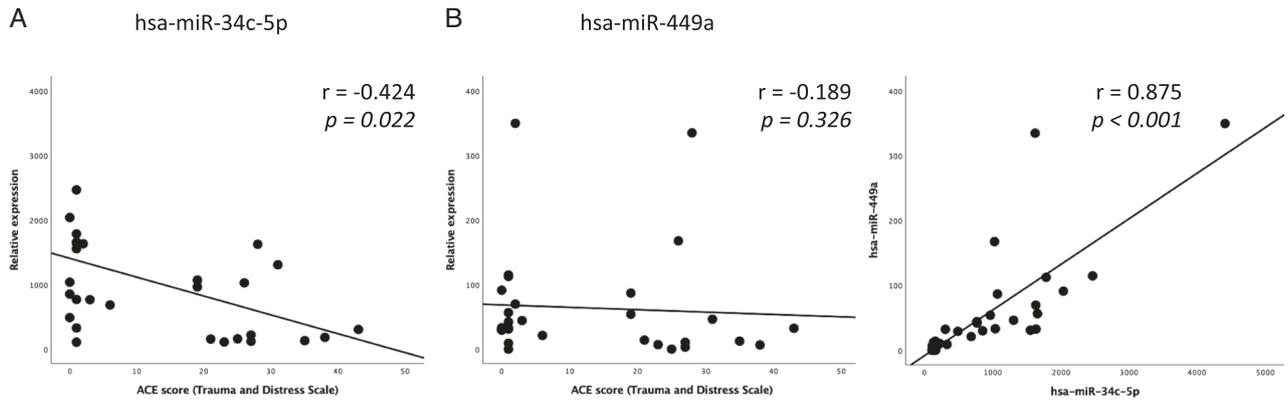


**Fig. 2 Differential levels of sperm sncRNAs in high-TADS vs low-TADS sperm samples.** **A** Scatter plot shows sncRNA expression between high- and low-TADS score groups. One dot represents the median of normalized counts of all samples belonging to either high- or low-TADS group. **B** MA plot displays the relationship between the log2 fold change in high-TADS vs low-TADS samples and log2 mean expression for all miRNAs and tsRNAs. **C** Volcano plots show differential expression of miRNAs in high- vs low-TADS samples. **D** Volcano plots show differential expression of tsRNAs in high- vs low-TADS samples. Blue and red colours refer to downregulated and upregulated sncRNAs respectively.

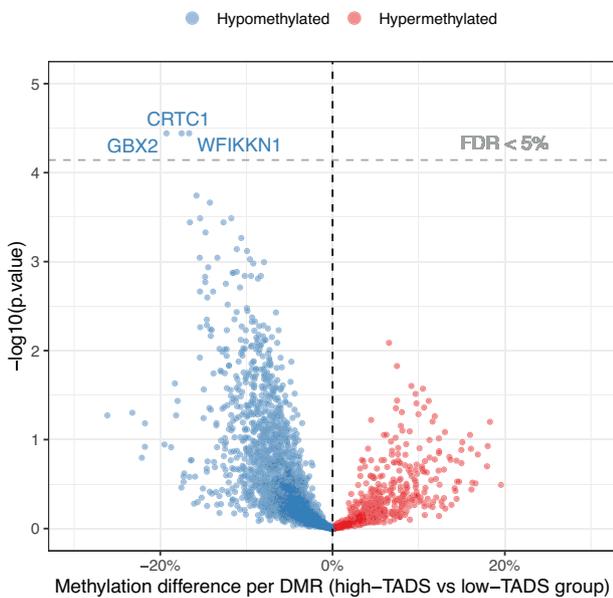
and 3 piRNA clusters were lower in high-TADS groups sperm compared to low-TADS sperm ( $\log_2FC < -1.0$  and  $P_{adj} < 0.05$ ), while correspondingly 18 miRNAs, 6 tsRNAs, and 1 piRNA cluster were higher ( $\log_2FC > 1.0$  and  $P_{adj} < 0.05$ ) (Fig. 2B–D, Supplementary Tables 2–4). The expression levels of three tsRNAs and five miRNAs showing the most significant changes ( $\log_2FC > 2$ , and  $P_{adj} < 0.01$ ) and the number of normalized counts among all samples above the threshold “baseMean > 10” are plotted in Supplementary Fig. 2. Interestingly, hsa-miR-34c-5p, which was earlier shown to be lower in sperm samples of individuals with high ACE scores [6, 21], was also lower in our discovery analysis (Supplementary Fig. 2H).

**CME associations are generally robust to potential confounders in multivariate analyses**

We identified the following potentially relevant covariates and confounders, based on prior literature, that were also measured in the current study (Table 1, Supplementary Table 5): sperm sample volume, sperm concentration, age, BMI, smoking, alcohol use, depressive symptoms, and anxiety symptoms. The multivariate statistical analyses showed that the associations that were implicated in the discovery analyses were very robust to include covariates. Most of the associations remained statistically significant after controlling for all potential confounders (Supplementary Tables 6 and 7).



**Fig. 3 The associations between TADS scores and miRNA expression levels that were identified in prior work.** **A** We found a negative association between ACE score and hsa-miR-34c-5p, **B** but no associations for hsa-miR-449a. The scatter plots correspond to prior report by Dickson et al. [6], Figure 2.



**Fig. 4 Differential levels of CpG methylation in high-TADS vs low-TADS sperm samples.** Each dot represents one candidate genomic region tested for differential methylation. The three regions that showed statistically significant group differences after correction for multiple testing (FDR-corrected  $p$  values  $< 0.05$ ) are annotated according to the nearest gene.

**Expression of hsa-miR-34c-5p is robustly associated with CME across studies**

Some associations between TADS scores and the expression of different miRNA in sperm were previously identified [6]. In line with prior work, our discovery analysis showed that the relative expression levels of hsa-miR-34c-5p was lower in high early life stress exposure group. A separate correlation-based replication analysis returned a negative association between ACE score and hsa-miR-34c-5p with a similar effect size, but no associations for hsa-miR-449a levels (Fig. 3A and B; replicating Fig. 2 plots of Dickson et al.). Like shown previously, the expression of the hsa-miR-34c-5p and hsa-miR-34c-5p (Fig. 3C). Partial correlation analyses for the depicted associations and controlled for age and BMI at measurement, and smoking (yes/no) identified associations between hsa-miR-34c-5p and TADS score ( $r = -0.528, p = 0.001$ ), but not between hsa-miR-449a and TADS score ( $r = -0.185, p = 0.339$ ). Of note, the association of hsa-miR-34c-5p vs. TADS score was slightly weaker when controlling for all

covariates ( $r = -0.420, p = 0.073$ ) (Supplementary Table 6). We next performed corresponding between-group comparisons (Supplementary Fig. 3; replicating plots in Dickson et al. Figure 1). We replicated the lower expression levels in ACE exposed group for hsa-miR-34c-5p ( $W = 140, p = 0.017$ , rank biserial correlation [rbc] = 0.538) and that there were no differences in expression levels of hsa-miR-152-3p and hsa-miR-375-3p. We replicated the differences between low/high CME groups in the discovery analyses and replication analyses as well as the negative correlation between continuous CME scores and hsa-miR-34c-5p. Our analyses thus show that sperm expression of hsa-miR-34c-5p is robustly negatively associated with CME.

**Childhood maltreatment exposure is associated with modified sperm CpG methylation**

To investigate if early life stress experience is associated with particular sperm DNA methylation signatures, we measured DNA methylation levels across CpG-rich regions using Reduced Representation Bisulfite Sequencing (RRBS). We identified 3 differentially methylated regions (DMRs) between the high and the low TADS group (false discovery rate,  $FDR < 0.05$ , Fig. 4, Supplementary Fig. 4, and Supplementary Table 8). These DMRs are located in the 3' UTR of the CREB Regulated Transcription Coactivator 1 gene (*CRTC1*), the promoter WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 1 gene (*WFIKKN1*), and in the distal intergenic region nearest the Gastrulation Brain Homeobox 2 (*GBX2*) gene.

Using pyrosequencing, we validated our RRBS results on a selection of genomics regions identified by RRBS; three regions with FDR below 0.2 and two with FDR above 0.7 (Supplementary Table 9). Consistent with the RRBS results, we detected lower DNA methylation across the statistically significant DMR ( $FDR < 0.05$ ) in the high TADS group (Supplementary File 1). Thus, CME is associated with distinct DNA methylation signatures in sperm.

**DISCUSSION**

Here, we report that CME is associated with specific epigenetic signatures in sperm. We identified differential expression of numerous sncRNAs and 3 genomic regions with hypomethylation in the high TADS score group. We found an interesting overlap with previous reports for miRNAs, particularly miR-34c-5p, although most of our results were distinct from prior reports. Specifically, we discovered that the expression of miRNA hsa-miR-34c-5p and the methylation levels near the *CRTC1* and *GBX2* genes are different in individuals with high TADS scores. Given the documented role of these genes on brain development, we speculate that remodelling of the sperm epigenome by psychological stress during childhood

may affect the developmental programming of the embryo post fertilization and affect the function of the central nervous system in the next generation offspring.

### Questionnaires of early life stressors

Early life adversity and stress are usually assessed retrospectively, and such measures are inherently prone to recall bias and they are unable to capture exposures for the youngest ages, e.g., under the age of three years. It is clear, however, that these questionnaires capture useful information about the cumulative stressors that have been present during childhood. Attesting to this ACEs have been associated with multiple adverse health outcomes [4, 37, 38].

We measured early life stress with TADS questionnaire that captures five dimensions of neglect and abuse CME that are also part of the most widely adopted CTQ. The other frequently adopted option is to use questionnaires that quantify ACEs. For instance, the ACE Study Questionnaire includes yes/no answers to 10 questions involving participants' experiences at home until the age of 18. Five of the questions probe CME: physical abuse, verbal abuse, sexual abuse, physical neglect, and emotional neglect which are also the main features of CTQ and TADS questionnaires. The other five questions probe adversity of family members (that has likely a negative influence on the exposed individual): a caretaker with alcoholism or alcohol abuse, experiencing domestic violence, caretaker incarceration, a family member diagnosed with a mental illness, and the loss of a caretaker through divorce, death, or abandonment [6]. In summary, there are many options for quantifying ACEs and CME [39], and while most of them probe the five main types of childhood maltreatment, there is very little knowledge on similarities and differences between the questionnaires [39].

### Stable epigenetic marks in sperm

The stability of the sperm epigenome is not well known. Indeed, human studies frequently use cross-sectional data, which creates obvious limitations for quantifying measurement error (this limitation applies to the current study as well). Roberts et al. [8] report an intra-class correlation coefficient (ICC) between replicate sperm samples taken ca. 3 months apart for the DNAm profiles that were associated with abuse exposure so that ICC values were higher than 0.7 for 90% of implicated sites. Dickson et al. did not collect replicate samples [6]. The field would benefit from larger-scale studies that describe the "normative" sncRNA and DNAm profiles in the sperm and describe, which components are stable and which ones are more dynamic across months and longer-term e.g., over several years. Commonly used measures of sperm cell epigenetics, small RNAs and DNA methylation (DNAm) patterns, are modifiable through lifestyle factors, health, and environmental exposures [16]. Correspondingly, if ACEs cause epigenetic programming in germ line cells that are relevant to intergenerational inheritance, most of them should be relatively stable following the exposure and they would not increase or decrease in time, which is in contrast to many other studied exposures such as cigarette smoking, exercise, acute stress, diet, and obesity.

### Possible links from sperm epigenome to offspring brain development

Although epigenetic mechanisms following fertilization related to DNAm and sncRNAs are likely essential in typical development and may be closely intertwined [15, 16], the most intriguing and robust evidence supports sncRNAs in conveying epigenetic inheritance [15, 40, 41]. Rodent studies have shown transgenerational inheritance of paternal ACEs via changes in miRNA profiles [12–14, 42, 43]. Similar effects have been reproduced without paternal exposure by injecting the implicated miRNA into zygotes [6, 14, 43].

Prior work implicated that ACEs were negatively associated with the abundance of miRNAs 449/34 in sperm [6, 21]. Importantly,

these miRNAs are unexpressed in oocytes but transmitted to them upon fertilization [6, 44], and are key regulators of brain development [45–48], including foetal human brain development [49, 50], and possibly also later in development [51]. Our analyses provided partial replication for prior work by implicating a negative association between ACEs and miR-34 (Fig. 4, Supplementary Fig. 3). As a novel and interesting finding, we identified 21 tsRNAs which had differential abundance in high- vs low-CME participants. In addition to miRNAs, tsRNAs have also been linked to epigenetic inheritance and could be a biomarker for CME in line with miRNAs [52].

Prior studies and the current study identified sperm epigenetic features that could potentially have effects on offspring brain development, which ties in with our recent neuroimaging studies that link paternal CME with offspring neonate brain structure [24, 53, 54]. Several studies, including from members of our group, have identified that lifestyle factors remodel DNA methylation near genes controlling the development of the brain in human sperm [55, 56], supporting that genomic regions involved in brain development are hotspots of epigenetic variation in response to environmental stress.

Our analyses identified three differentially methylated regions near *CRTC1*, *WFIKK1*, and *GBX2*. Expression of *CRTC1* in the hippocampus has been implicated in mood-related disorders such as depression [57, 58]. Increased methylation in the 3' UTR region might relate to increased gene expression of *CRTC1* [59]. Expression of *CRTC1* may be decreased in the brain, as methylation in the *CRTC1* DMR in spermatozoa of individuals with high-TADS has the potential to programme *CRTC1* expression throughout embryogenesis [59]. *GBX2* regulates brain development and is linked with neural crest differentiation, dopaminergic neurogenesis, and the formation of the neural plate [60, 61]. Whether the DMR near *GBX2* regulate brain development and neuronal function requires further investigation. *WFIKK1* regulates the effects of transforming growth factor beta (TGFβ) especially in muscle tissue [62].

The reduction in methylation percentage between the high TADS and the control group was in the range of 19–21% for all the statistically significant DMRs. These magnitudes implicate that in one-single ejaculate, a substantial number of spermatozoa are affected by epigenetic reprogramming, representing a potentially high effect size. Thus, in the light of the role of gametic epigenetic signature in embryonic development, remodelling of the sperm epigenome in high TADS individuals may carry biological relevance. While intergenerational effects carried by the sperm epigenome in humans have not been definitively demonstrated, an altered gametic epigenome after CME may influence the development of the central nervous system and modulate the behaviour of the next generation offspring.

### Sperm epigenome as a biomarker of childhood ACEs/CME

Epigenetics is a nascent field with typically small sample sizes that are related to high costs and the need for special infrastructure. This is not too different from the early stages of fields such as genetics and neuroimaging. Within this context, collecting very large data sets is not always feasible. Future studies would benefit from including replicate sperm samples to at least part of the participants to quantify the stable/repeatable elements in both early life stress exposure and control groups. This could potentially be used to identify the most relevant measures of interest for later analyses and greatly decrease the need for multiple comparisons correction. Early life stressors are challenging to quantify reliably, and it may well be that the effects of the exposures are different for different ages, for instance before and after puberty. Prospective cohorts that have detailed information over childhood and measures of early life stressors, including ACEs and CME, could share light on the matter by including sperm data collection e.g., during early adulthood.

Roberts et al. were able to estimate a parsimonious epigenetic marker for childhood abuse using an elastic net model (penalized regression), which identified three DNAm probes that predicted high vs. no childhood abuse in 71% of participants [8]. Such findings are very promising and may lead to parsimonious predictive models in the future. Future mechanistic studies should advance our understanding of how DNAm are affected by the underlying genome, and how sncRNAs and DNAm interact in causal pathways following fertilization. Dickson et al. combined data from human and animal models and were able to show that the effects of their CSI stress paradigm implicate similar sncRNA profiles in mice and that these effects are transmitted to embryos and can thus have inter/transgenerational effects [6]. Studies that focus on RNA are much better able to measure such epigenetic signatures since DNA methylation undergoes erasure and reestablishment following fertilization, which makes studying the immediate post-fertilization effects challenging [16].

### Practical implications

Intergenerational transmission of well-being, health and disease is an important research topic with many implications for health care and societies. It has been postulated that a key component of ACEs, CME is the single most important preventable risk factor for future mental health [1, 4, 37, 38]. CME has also been shown to have effects on health outcomes even when genetic confounding is taken into account [63]. Total annual costs attributable to ACEs were estimated to be US\$581 billion in Europe and \$748 billion in North America [1, 37]. Over 75% of these costs arose in individuals with two or more ACEs. Elucidating the mechanisms of intergenerational epigenetic inheritance in humans should be of high priority. Finding interventions to intergenerational effects that pass from one generation to the other could potentially spare future generations from the exposures of their ancestors.

While we fully acknowledge the lack of definitive evidence in humans, and the inherent limitations of our study (detailed below), our results may have revealed a mechanism by which early life psychological stress may programme gametes and affect the next generation offspring, through a process referred to as epigenetic inheritance [64]. Since most of the epigenetic signatures that we detected in the high-TADS group were robust to covariates relating to mental health and lifestyle, it is not unlikely that these distinct epigenetic signatures were the result of early life stress specifically, and that these signatures are relatively stable across the male lifespan. We profiled the sperm epigenome across a population of sperm cells of single ejaculates, so that the reported differences reflect the average for each participant. In the hypothetical scenario where a man with CME history would conceive a child, the sperm which would fertilize the egg would be more likely to carry epigenetic features of CME response, such as lower methylation of *CRTC1*, *WFIKKN1*, and *GBX2*, and lower amount of hsa-miR-34c-5p, than individuals without history of CME, provided these epigenetic signatures are not related to lowered fertility [65]. While fertilizing with a spermatozoa carrying altered signatures could have implications to offspring development and phenotype later in life, we acknowledge that epigenetic inheritance per se was not addressed in this study, and that there are many factors that limit the possibilities for demonstrating epigenetic inheritance in humans [66]. Even if CME-induced epigenetic inheritance were to be definitely identified in humans, whether the traits that are passed on are beneficial or deleterious to the offspring across the life course would need long explorations [67].

### Strengths and limitations

We used the largest sample size to date for identifying epigenetic marks relates to CME, but the sample size is still modest and larger sample sizes are needed in the future. Measurement error and test-retest reliability of sperm epigenome were not quantified,

which could be done with repeated measurements of sperm epigenome. CME measures were obtained retrospectively and are prone to recall bias, but CME measures are widely used and provide the only tangible way of assessing childhood experiences. All participants were Scandinavian/Caucasian, which makes the source population homogenous but necessitates the inclusion of more ethnically diverse populations in the future. One key strength of our study is that samples were not confounded with epigenetic signatures of infertility. Sperm was indeed collected from men with proven fertility that is, after a pregnancy was achieved. Samples from participants seeking fertility treatment or with unknown fertility status are risking that CME causes fertility issues, associated with altered sncRNA and DNAm profiles.

### CONCLUSIONS

In the current study, multiple sncRNAs were found associated with CME. For the majority of sncRNAs differentially expressed in the high compared to low TADS score group, these associations were robust after testing for semen quality and health-related factors as confounders. We discovered 3 regions of the genome carrying distinct DNA methylation profiles in sperm from individuals with history of childhood psychological stress, notably at proximity of genes regulating the development of the central nervous system. Collectively, our findings support that earlier stress exposures, even from as far as childhood, can leave epigenetic imprints in sperm cells that is are stable across the life course. Our work adds to the accumulating number of studies showing that environmental factors at the broad sense influences the epigenome for human sperm, and further opens for determining the consequence of environmentally acquired epigenetic signatures in sperm on the health of the offspring.

### DATA AVAILABILITY

The Finnish law and ethical permissions do not allow the sharing of the data used in this study. Data sharing is possible via formal agreements and interested investigators are requested to contact FinnBrain study administration; (<https://sites.utu.fi/finnbrain/en/contact/>).

### CODE AVAILABILITY

The preprocessing and statistical models were performed with openly available tools, and we report the details in the methods section.

### REFERENCES

- Hughes K, Ford K, Bellis MA, Glendinning F, Harrison E, Passmore J. Health and financial costs of adverse childhood experiences in 28 European countries: a systematic review and meta-analysis. *Lancet Public Health*. 2021;6:e848–57.
- Stoltenborgh M, Bakermans-Kranenburg MJ, Alink LRA, IJzendoorn MH. The prevalence of child maltreatment across the globe: review of a series of meta-analyses. *Child Abuse Rev*. 2015;24:37–50.
- Soares S, Rocha V, Kelly-Irving M, Stringhini S, Fraga S. Adverse childhood events and health biomarkers: a systematic review. *Front Public Health*. 2021;9:649825.
- Sara G, Lappin J. Childhood trauma: psychiatry's greatest public health challenge? *Lancet Public Health*. 2017;2:e300–1.
- Waehrer GM, Miller TR, Marques SCS, Oh DL, Harris NB. Disease burden of adverse childhood experiences across 14 states. *Plos One*. 2020;15:e0226134.
- Dickson DA, Paulus JK, Mensah V, Lem J, Saavedra-Rodriguez L, Gentry A, et al. Reduced levels of miRNAs 449 and 34 in sperm of mice and men exposed to early life stress. *Transl Psychiat*. 2018;8:101.
- Moog NK, Entringer S, Rasmussen JM, Styner M, Gilmore JH, Kathmann N, et al. Intergenerational effect of maternal exposure to childhood maltreatment on newborn brain anatomy. *Biol Psychiat*. 2018;83:120–7.
- Roberts AL, Gladish N, Gatev E, Jones MJ, Chen Y, MacIsaac JL, et al. Exposure to childhood abuse is associated with human sperm DNA methylation. *Transl Psychiat*. 2018;8:194.
- Scorza P, Duarte CS, Hipwell AE, Posner J, Ortin A, Canino G, et al. Research Review: Intergenerational transmission of disadvantage: epigenetics and parents' childhoods as the first exposure. *J Child Psychol Psych*. 2019;60:119–32.

10. Yehuda R, Lehrner A. Intergenerational transmission of trauma effects: putative role of epigenetic mechanisms. *World Psychiatry*. 2018;17:243–57.
11. Bohacek J, Farinelli M, Mirante O, Steiner G, Gapp K, Coiret G, et al. Pathological brain plasticity and cognition in the offspring of males subjected to postnatal traumatic stress. *Mol Psychiatr*. 2015;20:621–31.
12. Bohacek J, Rassoulzadegan M. Sperm RNA: Quo vadis? *Semin Cell Dev Biol*. 2019;97:123–30.
13. Gapp K, van Steenwyk G, Germain PL, Matsushima W, Rudolph KLM, Manuella F, et al. Alterations in sperm long RNA contribute to the epigenetic inheritance of the effects of postnatal trauma. *Mol Psychiatr*. 2020;25:2162–74.
14. Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados P, et al. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat Neurosci*. 2014;17:667–9.
15. Kretschmer M, Gapp K. Deciphering the RNA universe in sperm in its role as a vertical information carrier. *Environ Epigenetics*. 2022;8:dvac011.
16. Ghai M, Kader F. A review on epigenetic inheritance of experiences in humans. *Biochem Genet*. 2022;60:1107–40.
17. Mamar MB, Beck D, Nilsson E, McCarrey JR, Skinner MK. Developmental alterations in DNA methylation during gametogenesis from primordial germ cells to sperm. *Iscience*. 2022;25:103786.
18. Nestler EJ. Transgenerational epigenetic contributions to stress responses: fact or fiction? *Plos Biol*. 2016;14:e1002426.
19. Wang Y, Liu H, Sun Z. Lamarck rises from his grave: parental environment-induced epigenetic inheritance in model organisms and humans. *Biol Rev*. 2017;92:2084–111.
20. Senaldi L, Smith-Raska M. Evidence for germline non-genetic inheritance of human phenotypes and diseases. *Clin Epigenetics*. 2020;12:136.
21. Jawaid A, Kunzi M, Mansoor M, Khan ZY, Abid A, Taha M, et al. Distinct microRNA signature in human serum and germline after childhood trauma. *Medrxiv*. 2020;2020.08.11.20168393.
22. Karlsson L, Tolvanen M, Scheinin NM, Uusitupa HM, Korja R, Ekholm E, et al. Cohort Profile: The FinnBrain Birth Cohort Study (FinnBrain). *Int J Epidemiol*. 2018;47:15–16j.
23. Salokangas RKR, Schultze-Lutter F, Patterson P, von Reventlow HG, Heinimaa M, From T, et al. Psychometric properties of the Trauma and Distress Scale, TADS, in an adult community sample in Finland. *Eur J Psychotraumatol*. 2016;7:30062.
24. Karlsson H, Merisaari H, Karlsson L, Scheinin NM, Parkkola R, Saunavaara J, et al. Association of cumulative paternal early life stress with white matter maturation in newborns. *JAMA Netw Open*. 2020;3:e2024832.
25. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. *Brit J Psychiatr*. 1987;150:782–6.
26. Holi MM, Sarmallahti PR, Aalberg VA. A Finnish validation study of the SCL-90. *Acta Psychiatr Scand*. 1998;97:42–6.
27. Danielsson O, Nano J, Pakkala K, Rospleszcz S, Lehtimäki T, Schlett CL, et al. Validity of fatty liver disease indices in the presence of alcohol consumption. *Scand J Gastroenterol*. 2022;57:1349–60.
28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet J*. 2011;17:10–2.
29. Shi J, Ko EA, Sanders KM, Chen Q, Zhou T. SPORTS1.0: A tool for annotating and profiling non-coding RNAs optimized for rRNA- and tRNA-derived small RNAs. *Genom Proteom Bioinform*. 2018;16:144–51.
30. Ha H, Song J, Wang S, Kapusta A, Feschotte C, Chen KC, et al. A comprehensive analysis of piRNAs from adult human testis and their relationship with genes and mobile elements. *Bmc Genomics*. 2014;15:545.
31. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37:907–15.
32. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–30.
33. Bock C, Halbritter F, Carmona FJ, Tierling S, Datlinger P, Assenov Y, et al. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. *Nat Biotechnol*. 2016;34:726–37.
34. Pajares MJ, Palanca-Ballester C, Urtasun R, Aleman-Cosme E, Lahoz A, Sandoval J. Methods for analysis of specific DNA methylation status. *Methods*. 2021;187:3–12.
35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B (Methodol)*. 1995;57:289–300.
37. Bellis MA, Hughes K, Ford K, Rodriguez GR, Sethi D, Passmore J. Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. *Lancet Public Health*. 2019;4:e517–28.
38. Hughes K, Bellis MA, Hardcastle KA, Sethi D, Butchart A, Mikton C, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *Lancet Public Health*. 2017;2:e356–66.
39. Thabrew H, Sylva S, de, Romans S. Evaluating childhood adversity. *Adv Psychosom Med*. 2011;32:35–57.
40. Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA. Delivering spermatozoan RNA to the oocyte. *Nature*. 2004;429:154–154.
41. Sandler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, et al. Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic Acids Res*. 2013;41:4104–17.
42. Luo J, Tan X, Li H, Ding X. sncRNAs in Epididymosomes: The contribution to embryonic development and offspring health. *Int J Mol Sci*. 2022;23:10851.
43. Rodgers AB, Morgan CP, Leu NA, Bale TL. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *Proc Natl Acad Sci*. 2015;112:13699–704.
44. Yuan S, Schuster A, Tang C, Yu T, Ortogero N, Bao J, et al. Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. *Development*. 2015;143:635–47.
45. Grad M, Nir A, Levy G, Trangle SS, Shapira G, Shomron N, et al. Altered white matter and microRNA expression in a murine model related to Williams Syndrome suggests that miR-34b/c Affects Brain Development via Ptpu and Dcx Modulation. *Cells*. 2022;11:158.
46. Jauhari A, Singh T, Singh P, Parmar D, Yadav S. Regulation of miR-34 family in neuronal development. *Mol Neurobiol*. 2018;55:936–45.
47. Mollinari C, Racaniello M, Berry A, Pieri M, Stefano MC, de, Cardinale A, et al. miR-34a regulates cell proliferation, morphology and function of newborn neurons resulting in improved behavioural outcomes. *Cell Death Dis*. 2015;6:e1622–e1622.
48. Wu J, Bao J, Kim M, Yuan S, Tang C, Zheng H, et al. Two miRNA clusters, miR-34b/c and miR-449, are essential for normal brain development, motile ciliogenesis, and spermatogenesis. *Proc Natl Acad Sci*. 2014;111:E2851–7.
49. Rao VTS, Ludwin SK, Fuh SC, Sawaya R, Moore CS, Ho MK, et al. MicroRNA expression patterns in human astrocytes in relation to anatomical location and age. *J Neuropathol Neurol*. 2016;75:156–66.
50. Venø MT, Venø ST, Rehberg K, Asperen JV, van, Clausen BH, Holm IE, et al. Cortical morphogenesis during embryonic development is regulated by miR-34c and miR-204. *Front Mol Neurosci*. 2017;10:31.
51. Morgunova A, Flores C. MicroRNA regulation of prefrontal cortex development and psychiatric risk in adolescence. *Semin Cell Dev Biol*. 2021;118:83–91.
52. Park J, Ahn SH, Shin MG, Kim HK, Chang S. tRNA-derived small RNAs: Novel epigenetic regulators. *Cancers*. 2020;12:2773.
53. Tuulari JJ, Kataja EL, Karlsson L, Karlsson H. Associations of cumulative paternal and maternal childhood maltreatment exposure with neonate brain anatomy. *bioRxiv*. 2022;2022.10.16.512276.
54. Tuulari JJ, Pulli EP, Kataja EL, Perasto L, Lewis JD, Karlsson L, et al. Parental childhood maltreatment associates with offspring left amygdala volume at early infancy. *Biorxiv*. 2023;2023.02.23.529799.
55. Donkin I, Verstehe S, Ingerslev LR, Qian K, Mechta M, Nordkap L, et al. Obesity and bariatric surgery drive epigenetic variation of Spermatozoa in humans. *Cell Metab*. 2016;23:369–78.
56. Ingerslev LR, Donkin I, Fabre O, Verstehe S, Mechta M, Pattamaprapanont P, et al. Endurance training remodels sperm-borne small RNA expression and methylation at neurological gene hotspots. *Clin Epigenetics*. 2018;10:12.
57. Cherix A, Poitry-Yamate C, Lanz B, Zanoletti O, Grosse J, Sandi C, et al. Deletion of Crtc1 leads to hippocampal neuroenergetic impairments associated with depressive-like behavior. *Mol Psychiatry*. 2022;27:4485–501.
58. Li D, Liao Q, Tao Y, Ni S, Wang C, Xu D, et al. Downregulation of CRTC1 Is Involved in CUMS-Induced Depression-Like Behavior in the Hippocampus and Its RNA Sequencing Analysis. *Mol Neurobiol*. 2022;59:4405–18.
59. Luo R, Bai C, Yang L, Zheng Z, Su G, Gao G, et al. DNA methylation subpatterns at distinct regulatory regions in human early embryos. *Open Biology*. 2018;8:180131.
60. Méndez-Maldonado K, Vega-López GA, Aybar MJ, Velasco I. Neurogenesis from neural crest cells: molecular mechanisms in the formation of cranial nerves and Ganglia. *Front Cell Dev Biol*. 2020 [cited 2024 Sep 30];8. Available from: <https://www.frontiersin.org/journals/cell-and-developmental-biology/articles/10.3389/fcell.2020.00635/full>
61. Maimaitili M, Chen M, Febraro F, Ucuncu E, Kelly R, Niclis JC, et al. Enhanced production of mesencephalic dopaminergic neurons from lineage-restricted human undifferentiated stem cells. *Nat Commun*. 2023;14:7871.
62. Monestier O, Blanquet V. WFIKKN1 and WFIKKN2: “Companion” proteins regulating TGFβ activity. *Cytokine Growth Factor Rev*. 2016;32:75–84.
63. Baldwin JR, Sallis HM, Schoeler T, Taylor MJ, Kwong ASF, Tielbeek JJ, et al. A genetically informed Registered Report on adverse childhood experiences and mental health. *Nat Hum Behav*. 2023;7:269–90.

64. Breton CV, Landon R, Kahn LG, Enlow MB, Peterson AK, Bastain T, et al. Exploring the evidence for epigenetic regulation of environmental influences on child health across generations. *Commun Biol.* 2021;4:1–15.
65. Hosseini M, Khalafiyani A, Zare M, Karimzadeh H, Bahrami B, Hammami B, et al. Sperm epigenetics and male infertility: unraveling the molecular puzzle. *Human Genomics.* 2024;18:57.
66. Horsthemke B. A critical view on transgenerational epigenetic inheritance in humans. *Nat Commun.* 2018;9:2973.
67. Duempelmann L, Skribbe M, Bühler M. Small RNAs in the transgenerational inheritance of epigenetic information. *Trends Genetics.* 2020;36:203–14.

## ACKNOWLEDGEMENTS

We thank all men who participated to the study visits. In addition, we also thank our research nurses for their valuable contribution to the data collection. We want to thank Elina Louramo, Anna Puisto, Tiina Peromaa and Johanna Järvi for technical assistance in organizing the semen collection and performing the sperm purification, and Ashmeet Jolly for helping in reference management and technical formatting of the manuscript.

## AUTHOR CONTRIBUTIONS

Jetro J. Tuulari lead the analyses and the manuscript writing, participated to the conceptualization of the study, performed statistical analyses. Matthieu Bourgery preprocessed the sncRNA data, had a key role in data visualization, performed statistical analyses with major input to the discovery analyses. Jetro J. Tuulari and Matthieu Bourgery contributed equally and share the first authorship. Jo Iversen preprocessed the DNAm data for primary analyses and carried out the pysequencing validation. Thomas Gade Koefoed preprocessed the DNAm data, contributed to data visualization, performed statistical analyses with major input to the DNAm discovery analyses. Jo Iversen, Thomas Gade Koefoed contributed equally and share the second authorship. Annukka Ahonen prepared and helped in implementation of the data collection, Ammar Ahmedani optimized the sperm RNA extraction protocol and prepared the sperm RNA samples for sequencing. Eeva-Leena Kataja provided expertise on interpretation of the CME data and results. Linnea Karlsson conceptualized the study and build the infrastructure of the FinnBrain Birth Cohort. Romain Barrès provided the infrastructure for the DNAm analyses, supervised the DNAm analyses. Hasse Karlsson conceptualized the study and build the infrastructure of the FinnBrain Birth Cohort. Noora Kotaja conceptualized the study and build the infrastructure for the sperm sample storage and analyses and supervised the sncRNA analyses. All authors participated in writing and critically revising the manuscript and accepted it in its final form.

## FUNDING

Jetro J. Tuulari was supported by the State Research Grant (ERVA), Sigrid Juselius Foundation, Finnish Medical Foundation, Signe ja Ane Gyllenberg Foundation, and Emil Aaltonen Foundation. Linnea Karlsson was supported by the Academy of Finland (#325292). Romain Barrès and Jo Iversen were supported by a Challenge Programme Grant from the Novo Nordisk Foundation (NNF18OC0033754) to the Gametic Epigenetics Consortium against Obesity (GECKO). The Novo Nordisk Foundation

Center for Basic Metabolic Research is an independent research center at the University of Copenhagen, partially funded by an unrestricted donation from the Novo Nordisk Foundation (NNF18CC0034900). This work was supported by the French Government (National Research Agency, ANR) through the "Investments for the Future" programmes LABEX SIGNALIFE ANR-11-LABX-0028-01 and IDEX UCAJedi ANR-15-IDEX-01. Thomas Gade Koefoed and the Single-Cell Omics Platform is funded by the Novo Nordisk Foundation. The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent research center at the University of Copenhagen, partially funded by an unrestricted donation from the Novo Nordisk Foundation (NNF18CC0034900). Eeva-Leena Kataja was funded by Signe ja Ane Gyllenberg Foundation, Turku University Foundation, Academy of Finland (308252). Noora Kotaja was supported by Academy of Finland, Sigrid Jusélius Foundation, Novo Nordisk Foundation, Jalmari and Rauha Ahokas Foundation, and Jane and Atos Erkkö Foundation. Open Access funding provided by University of Turku (including Turku University Central Hospital).

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-024-02872-3>.

**Correspondence** and requests for materials should be addressed to Jetro J. Tuulari.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2025