2021

Sperm DNA methylation mediates the association of male age on reproductive outcomes among couples undergoing infertility treatment

Oladele A. Oluwayiose

Haotian Wu

Hachem Saddiki

Brian W. Whitcomb

Laura B. Balzer

See next page for additional authors

Authors

Oladele A. Oluwayiose, Haotian Wu, Hachem Saddiki, Brian W. Whitcomb, Laura B. Balzer, Nicole Brandon, Alexander Suvorov, Rahil Tayyab, Cynthia K. Sites, Lisa Hill, Chelsea Marcho, and J. Richard Pilsner

scientific reports

OPEN Sperm DNA methylation mediates the association of male age on reproductive outcomes among couples undergoing infertility treatment

Oladele A. Oluwayiose[1,5], Haotian Wu[4,5], Hachem Saddiki[2], Brian W. Whitcomb[2], Laura B. Balzer[2], Nicole Brandon[1], Alexander Suvorov[1], Rahil Tayyab[3], Cynthia K. Sites[3], Lisa Hill[3], Chelsea Marcho[1] & J. Richard Pilsner[1✉]

Parental age at time of offspring conception is increasing in developed countries. Advanced male age is associated with decreased reproductive success and increased risk of adverse neurodevelopmental outcomes in offspring. Mechanisms for these male age effects remain unclear, but changes in sperm DNA methylation over time is one potential explanation. We assessed genome-wide methylation of sperm DNA from 47 semen samples collected from male participants of couples seeking infertility treatment. We report that higher male age was associated with lower likelihood of fertilization and live birth, and poor embryo development ($p < 0.05$). Furthermore, our multivariable linear models showed male age was associated with alterations in sperm methylation at 1698 CpGs and 1146 regions ($q < 0.05$), which were associated with > 750 genes enriched in embryonic development, behavior and neurodevelopment among others. High dimensional mediation analyses identified four genes (*DEFB126, TPI1P3, PLCH2* and *DLGAP2)* with age-related sperm differential methylation that accounted for 64% (95% CI 0.42–0.86%; $p < 0.05$) of the effect of male age on lower fertilization rate. Our findings from this modest IVF population provide evidence for sperm methylation as a mechanism of age-induced poor reproductive outcomes and identifies possible candidate genes for mediating these effects.

Over the past several decades, a trend toward delayed childbirth has led to increases in parental age at the time of conception, especially in developed countries[1]. This delayed parenthood is attributed to secular and socio-economic factors[2], reproductive technological advancement[3], and higher levels of post-graduate education[4]. To date, emphasis has been on the effects of advanced maternal age on adverse reproductive and offspring health[5–7]. However, new evidence suggests that, irrespective of maternal age, higher male age contributes to longer time-to-conception[8], poor pregnancy outcomes[8], lower odds of live birth in infertile cohort[9], and adverse health of the offspring at later life[10–12]. For instance, higher male age is associated with poor outcomes in offspring, including: reduced life span in mice[13]; increased susceptibility to early development of cancer[14]; chromosomal abnormalities[15]; and neurodevelopmental disorders such as schizophrenia[16] and autism[11,12,17,18].

Despite the host of adverse reproductive health outcomes associated with higher male age at conception, the precise mechanism remains unclear. Proposed mechanisms include reduced sperm quality[19], poor sperm DNA integrity[20], the accumulation of de novo mutations from the self-renewal and differentiation of spermatogonial stem cells[21], and adverse psychosocial environment[22]. Age-associated adverse reproductive outcomes, however,

[1]Department of Environmental Health Sciences, School of Public Health and Health Sciences, University of Massachusetts Amherst, 173A Goessmann, 686 North Pleasant Street, Amherst, MA 01003, USA. [2]Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts Amherst, 715 North Pleasant Street, Amherst, MA, USA. [3]Division of Reproductive Endocrinology and Infertility, Baystate Medical Center, 759 Chestnut Street, Springfield, MA, USA. [4]Present address: Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, 722 W 168th St, New York, NY 10032, USA. [5]These authors contributed equally: Oladele A. Oluwayiose and Haotian Wu. ✉email: rpilsner@umass.edu
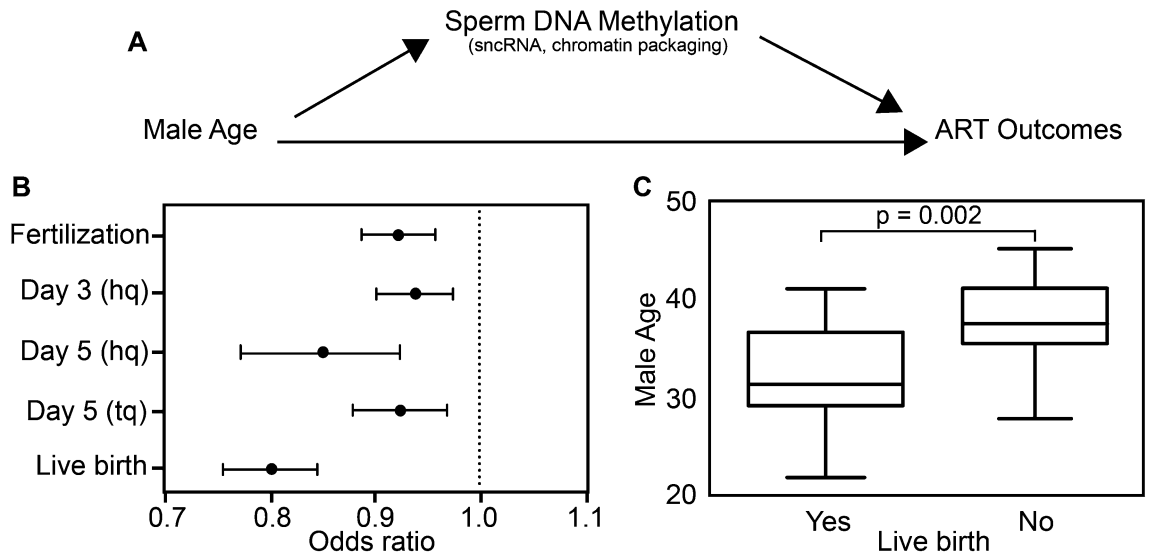
Figure 1. Male age and assisted reproductive technology (ART) outcomes. (**A**) Flow chart of our a priori hypothesis that sperm DNA methylation mediates the effect of male age on ART outcomes. Other potential epigenetic mechanisms not included in this study are listed below sperm DNA methylation including small noncoding RNA and chromatin packaging. (**B**) Odds ratios between male age and ART outcomes adjusting for male BMI, infertility, smoking and female age. (**C**) Unadjusted box plots of male age for couples with and without successful live births (t-test: − 3.48 p-value = 0.002). hq: high quality embryo; tq: transfer quality embryo.

are likely multifactorial, and as such, additional biological pathways are important to identify to inform avenues of translational research to improve reproductive health.

Age-associated changes in sperm epigenetics is another potential mechanism that may explain the association of higher male age with adverse reproductive outcomes, but limited data exist evaluating this hypothesis[23]. The epigenetic landscape of cells is determined by the interaction of DNA methylation, chromatin structure governed by histone modifications, and non-coding RNAs. Of particular relevance are the recent findings in mice and human indicating that the sperm epigenetic landscape in adult males can be modified by environmental conditions such as low protein diet[24], cannabis use[25], physical activity[26], stress[27] and environmental chemicals such as phthalates[28]. Several studies have reported age associated changes in DNA methylation in somatic tissues[29–31] with recent advances enabling prediction of adverse health outcomes including lifespan[32–34]. However, the impact of male age on methylation profiles of male germ cells is limited[35,36] and most recently, it has been shown that age-associated sperm methylation may not persist across generations[37]. However, to date, no study has examined age-associated sperm methylation profiles with respect to reproductive outcomes. Research in clinical populations of couples seeking treatment for infertility allows direct observation of reproductive events including fertilization, embryo development, and pregnancy outcomes. However, such observations are largely occult among the general population. Hence, the objective of this study was to examine the role of sperm DNA methylation as a potential mechanism by which male age affects assisted reproductive technology (ART) outcomes among couples seeking infertility treatment (Fig. 1A).

Results

Male age and assisted reproductive outcomes. Outcomes of ART treatment considered in these analyses include fertilization of retrieved oocytes; high quality embryo status at multiple time points; and whether the treatment resulted in a live birth. Characteristics of the couple participants (n = 47) are described in Table 1. Male age ranged from 21 to 45 years with the majority of the participants (53.2%) between 30 and 40 years of age. Similarly, female partners' age ranged from 24 to 42 years and were mostly (74%) between 30 and 40 years of age. In multivariable logistic regression models adjusting for male BMI, male infertility, male smoking and female age, each 1 year higher male age was associated with lower likelihood of fertilization (Fig. 1B. OR = 0.92, 95% confidence interval (CI) 0.89, 0.96); day 3 high quality embryos (OR = 0.94; 95% CI 0.90, 0.98), day 5 high quality embryos (OR = 0.85; 95% CI 0.77, 0.93), and day 5 transfer quality embryos (OR = 0.93; 95% CI 0.88, 0.97). Additionally, each 1 year higher male age was associated with lower likelihood of live birth (OR = 0.80; 95% CI 0.76, 0.85); likewise average male age was higher among couples without successful live birth compared to those with successful live birth (Fig. 1C. 35.1 ± 4.0 and 32.4 ± 4.4 years, respectively; p < 0.05). As a sensitivity analysis, models were also run stratified on infertility status to evaluate whether observed male age associations differed between fertile and infertile strata. Comparable estimates were observed, and results did not appreciably change the overall interpretation of our findings (Supplementary Table S1).

Individual characteristics	Male (n (%))	Female (n (%))
Age (years)		
< 26	2 (4.2)	1 (2.1)
26–30	7 (14.9)	8 (17.0)
31–35	8 (17.0)	16 (34.0)
36–40	17 (36.2)	19 (40.4)
> 40	13 (27.7)	3 (6.4)
BMI (kg/m²)[a]		
< 25	10 (21.3)	17 (36.2)
25–30	19 (40.4)	11 (23.4)
30+	17 (36.2)	19 (40.4)
Race[b]		
White (non-hispanic)	36 (76.6)	44 (93.6)
Others	4 (8.5)	2 (4.3)
Current smoking[c]		
Yes	4 (8.5)	0 (0%)
No	43 (91.5)	47 (100%)
Diagnosed infertility		
Male factor only	5 (10.6)	
Female factor only	20 (42.5)	
Both	6 (12.7)	
Unidentifiable	16 (34.0)	
Semen parameters	**Median**	**n (%) < WHO reference**
Normal morphology (%)	5.5	16 (34.0)
Total motility (%)	58	8 (17.0)
Sperm count (10⁶)	91	2 (4.3)
Semen volume (mL)	3	9 (19.2)
Sperm concentration (10⁶/mL)	47.5	12 (25.5)
Cycle-specific characteristics	**(Mean ± SD or n (%))**	
Oocyte retrieved	15.1 ± 8.0	
Fertilized embryos	9.0 ± 6.6	
High quality embryos		
Day 3	5.1 ± 4.1	
Day 5	0.9 ± 1.4	
Live birth	15 (31.9)	

Table 1. Characteristics of SEEDS participants (n = 47). [a]Missing male BMI data (n = 1). [b]Missing race data (Male, n = 7 and female, n = 1). [c]Males based on urinary cotinine and females based on self-report.

Male age and sperm methylation. In order to evaluate the extent to which the observed associations of higher male age with poorer ART outcomes might be mediated by sperm epigenetics, we next considered relations of male age with sperm DNA methylation of individual CpG sites and regionally. After adjusting for male BMI, male infertility and male smoking, male age was associated with 1698 individual sperm CpG sites (q < 0.05). The majority of CpGs (91%; n = 1546) showed an increase in methylation with respect to male age, whereby for every 5-year increase in male age, an increase in methylation at these CpGs ranged from 0.2 to 11.7% (mean ± SD: 2.8% ± 1.8%). Of the CpG sites associated with male age, 19 CpGs remained significant after Bonferroni correction (Fig. 2A,B; Supplementary Table S2).

Analysis was also conducted to identify differentially methylated regions (DMRs) associated with male age. In multivariable linear regression analyses adjusted for covariates as described previously, male age was associated with 1146 DMRs (q < 0.05), of which roughly 10% (n = 156) were statistically significant at the Bonferroni correction level (Fig. 2C,D; Supplementary Table S3). Unlike the overwhelming hypermethylation observed with individual CpGs, age-associated sperm DMRs represented a mix of hypomethylation (43%) and hypermethylation (57%) regions; however, the differences in percent methylation of DMRs per 5 years of age were similar to those of individual CpGs.

Genomic features of age-associated sperm DMRs. To better understand the localization of the age-associated sperm DMRs, we next examined DMRs based on their genomic features. Age-associated sperm DMRs were depleted in regions previously identified with retained nucleosomes reported by Donkin et al.[38] compared to all clusters (Fig. 3A: 12.9% vs. 18.3%; p = 4.7 × 10⁻⁶). For CpG features, compared to all clusters, age-associated DMRs were enriched in CpG shores (45.5% vs. 40.5%, p < 0.002) and depleted in CpG islands (13.2%

3
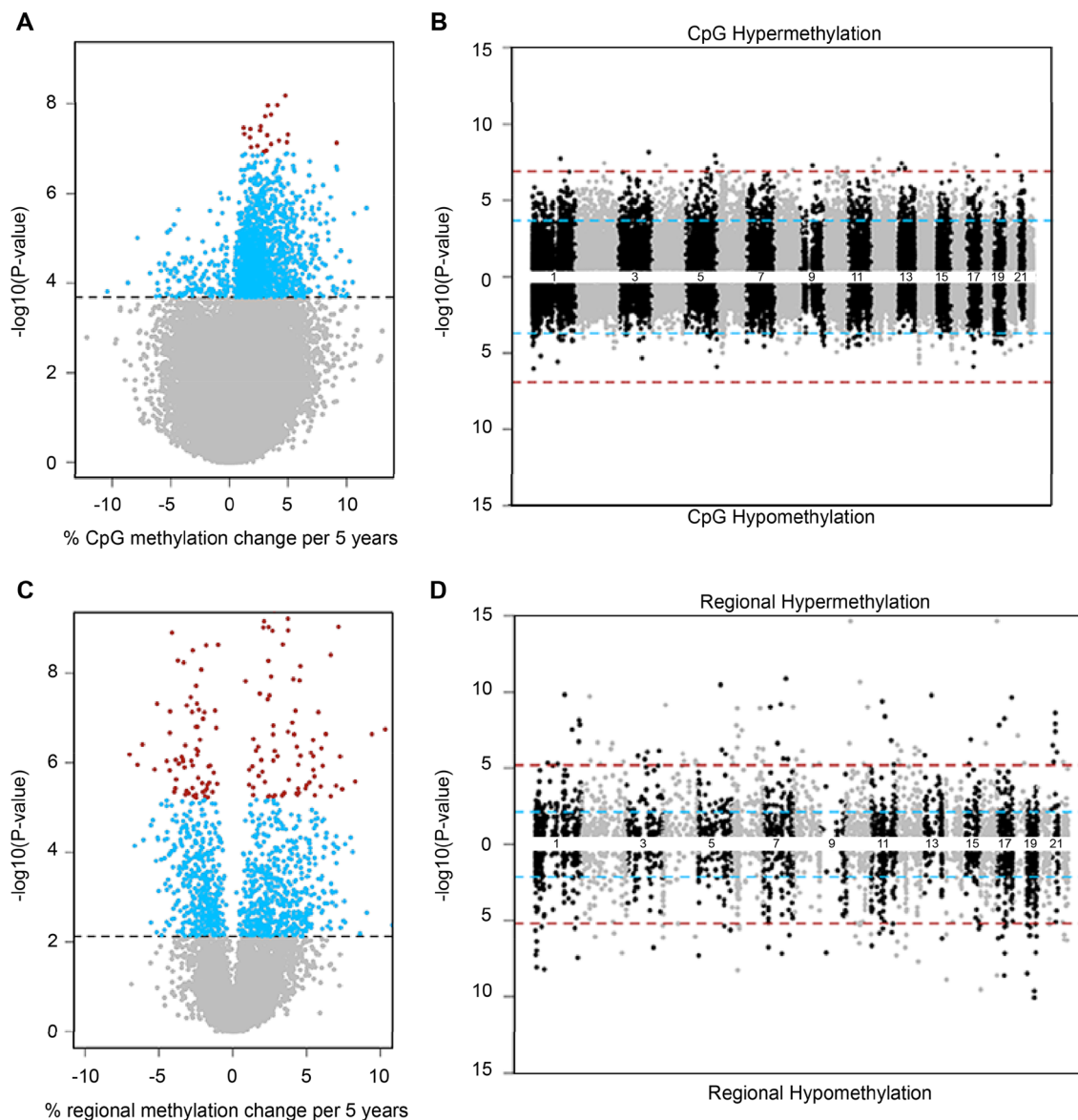
Figure 2. Volcano and Miami plots of age-associated sperm methylation at individual CpGs and regionally. Volcano plots show methylation change per 5 years of age adjusting for male BMI, infertility and smoking at (**A**) individual CpGs and (**C**) regionally where grey, blue and red-colored data points representing non-significant, FDR significant, and Bonferroni significant, respectively. Miami plots show the chromosomal distribution of age-associated sperm methylation at (**B**) individual CpGs and (**D**) regionally where blue and red dashed lines represent FDR and Bonferroni cut-off levels, respectively.

vs. 23.0%, p = 2.8×10^{-15}). For genic regions, age-associated DMRs were enriched in promoter regions compared to all clusters (Fig. 3A: 25.5% vs. 20.2%, p = 5.0×10^{-5}). Of the 148 transcription factors binding sites examined, three were enriched (*EZH2*, *ZNF263*, and *TAF7*) while *SUZ12* was depleted for age-associated DMRs compared to all clusters (Fig. 3B and Supplementary Table S4). Finally, the loss or gain of methylation in sperm DMRs was approximately equally distributed (40–60%) in most genomic features. However, pronounced hypermethylation of sperm DMRs was observed in open sea (88%) and intergenic (73%) regions, whereas shelves (66%) and shores (76%) were predominately hypomethylated (Fig. 3C). Hypomethylation of sperm DMRs was also pronounced among transcription factor bindings sites of *EZH2* (84%), *ZNF263* (74%) and *SUZ12* (92%) (Fig. 3D).

Pathway analyses of age-associated sperm DMRs. To determine the potential biological significance of age-associated sperm DMRs, we conducted pathway analyses by restricting analyses to DMRs located within 1500 bp distance from TSS of genes. Of the 1146 sperm DMRs identified, 783 unique genes remained for pathway analysis (Supplementary Table S5). Top enriched pathways were largely those involving early-life development, such as muscle structure development and embryonic organ morphogenesis, as well as neurodevelopmental pathways including spinal cord and forebrain development, neuron differentiation and regionalization, and behavior (Fig. 4A). Likewise, when we restricted our analyses to the 101 unique genes associated with the 156
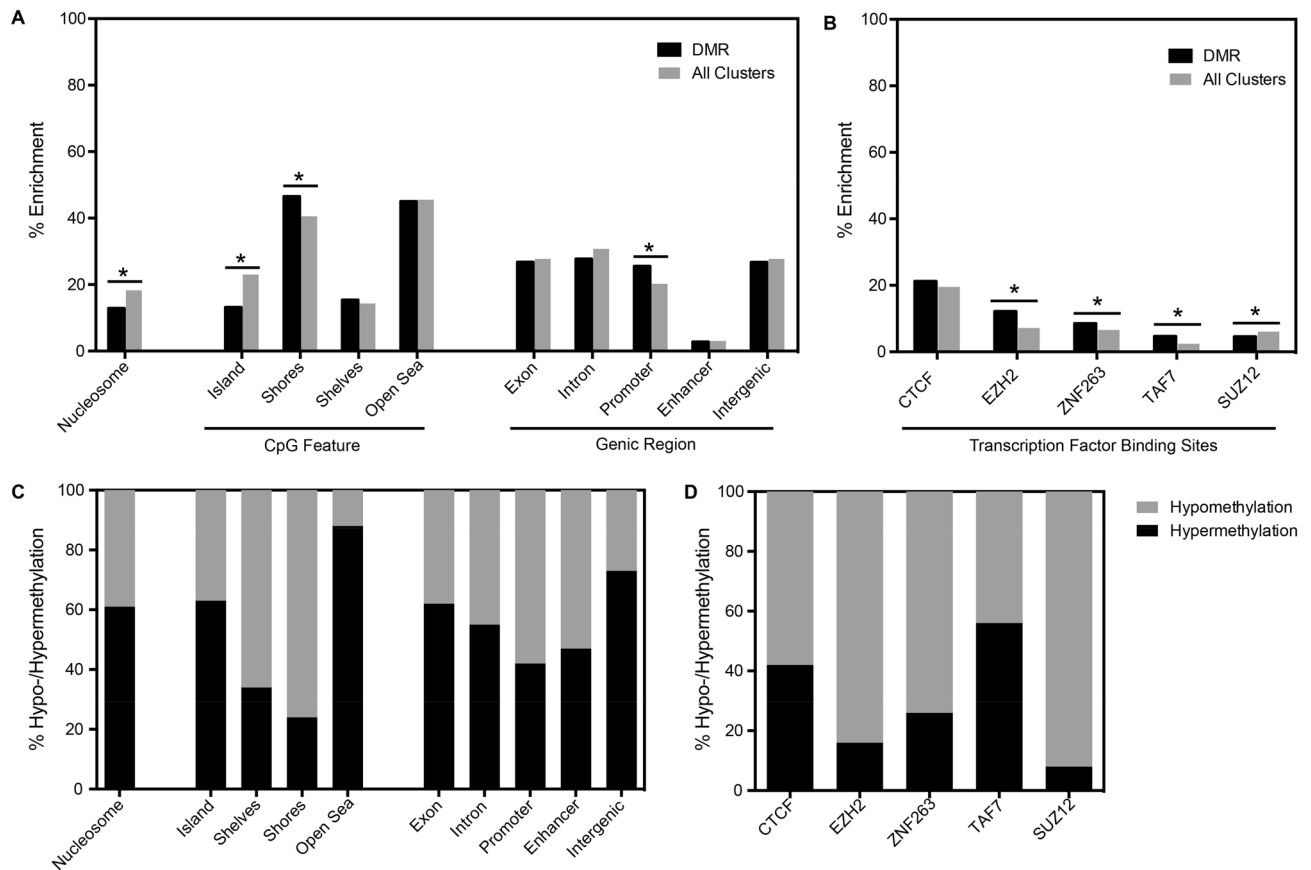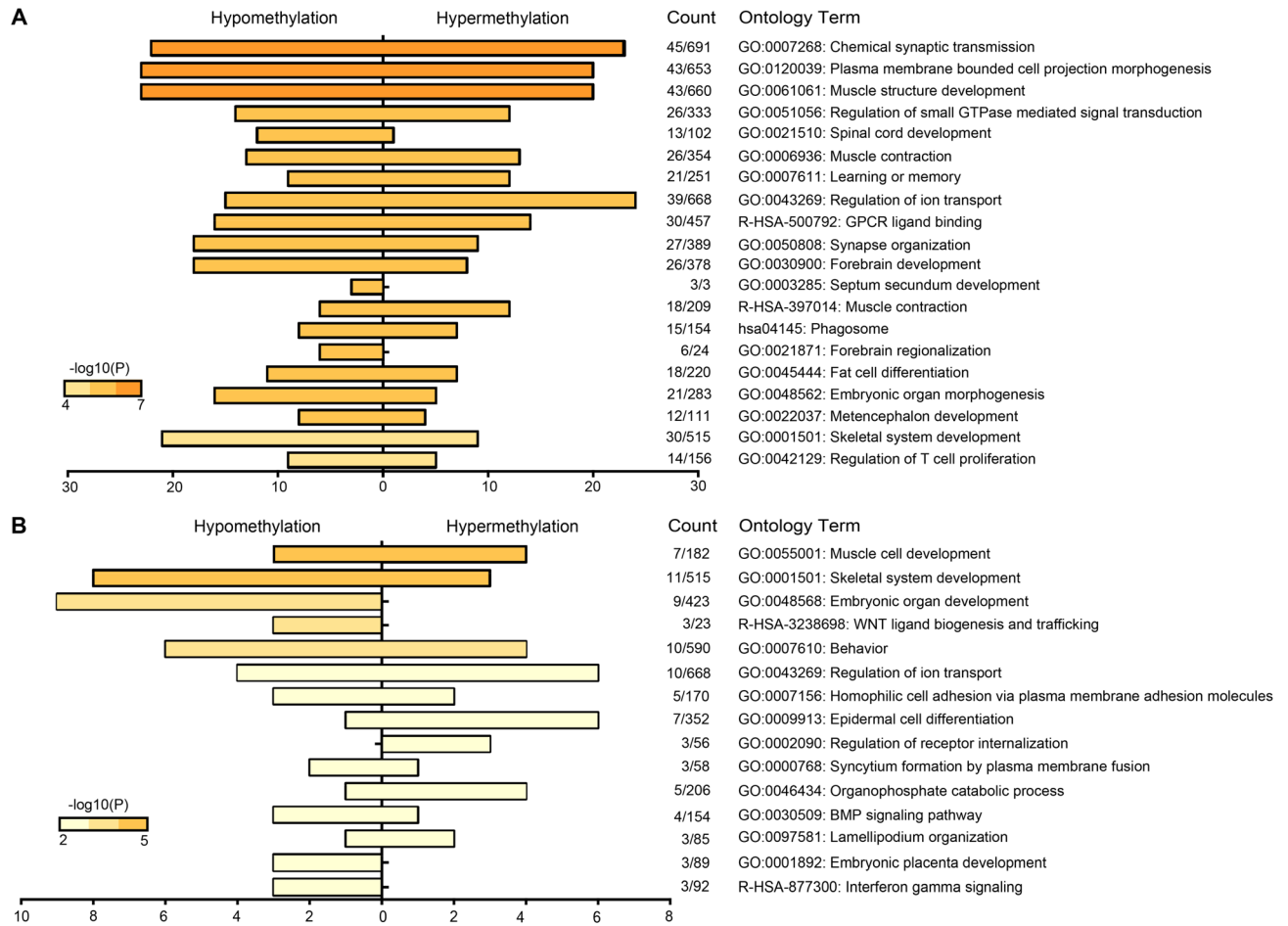
Figure 3. Localization of age-associated sperm DMRs. Enrichment analysis of all methylation clusters compared to DMRs at (**A**) nucleosome, CpG feature, and genic regions and (**B**) transcription factor binding sites and (**C,D**) their respective age-associated methylation patterns (**C,D**). Categories not exclusive, numbers may not add up to 100%.

statistically significant sperm DMRs at the Bonferroni correction level, behavior and developmental pathways were highly enriched including homophilic cell adhesion, epithelial cell differentiation, *WNT* ligand biogenesis and trafficking, and embryonic placenta development (Fig. 4B and Supplementary Table S6).

Scatter plots depict the associations between male age and sperm DNA methylation of a subset of genes important for early embryonic development and behavior (Fig. 5). Male age was associated with a loss of sperm methylation at loci of key development genes (Fig. 5A) such as of wingless/integrated 1 (*WNT1*; $rho = -0.56$, $p < 0.001$), GATA binding protein 4 (*GATA4*; $rho = -0.45$, $p = 0.007$), zinc finger of the cerebellum 1 (*ZIC1*; $rho = -0.49$, $p = 0.001$) and homeobox B6 (*HOXB6*; $rho = -0.44$, $p = 0.002$). For behavior-associated genes, male age was significantly associated with increases in sperm methylation in dopamine receptor D3 (*DRD3*; $rho = 0.51$, $p < 0.001$), and 5-hydroxytryptamine receptor 2A (*HTR2A*; $rho = 0.46$, $p = 0.001$). Altered methylation was also observed in glutamate ionotropic receptor NMDA type subunit 1 (*GRIN1*; $rho = -0.28$, $p = 0.059$) and synaptic Ras GTPase activating protein 1 (*SNYGAP1*; $rho = 0.26$, $p = 0.079$), although borderline significant with spearman correlations (Fig. 5B).

Age-associated sperm methylome profiles and ART outcomes. We next utilized unsupervised clustering of the 1698 significant individual sperm CpGs associated with male age to examine their relationships with the likelihood of fertilization and live birth (Fig. 6). Unsupervised clustering of age-associated methylome profiles resulted into four main groups with groups A and B having the youngest and oldest age (mean ± SD: 28.3 ± 4.6 years and 40.5 ± 3.2 years, respectively). The oldest group (group B) had distinct CpG hypermethylation profiles compared to the youngest age group while intermediate methylation was observed for groups C and D, suggesting a linear dose–response between male age and sperm methylation. When evaluating live birth in these groups defined by age-associated methylome profiles, 57% (n = 4 of 7) in the youngest group and 35% (n = 11 of 31) in the two intermediate age groups had successful live birth; however, no live births (n = 0 of 9) was observed in oldest group of men (group B). Interestingly, this group with no live births not only had the highest overall mean age, but also included younger men who had older sperm methylome profiles (Fig. 6).

Sperm DNA methylation as a mediator of male age on ART outcomes. We performed high dimensional mediation analyses to evaluate the role of sperm DNA methylation in male age-associated prob-
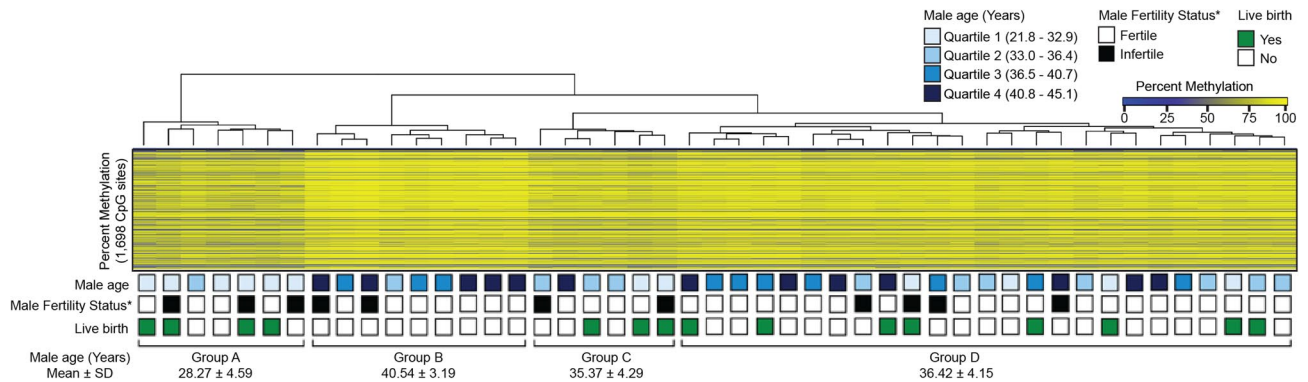
Figure 4. Pathway analysis of age-associated sperm DMRs. (**A**) FDR-adjusted DMRs (n = 1146) and (**B**) Bonferroni-adjusted DMRs (n = 156).
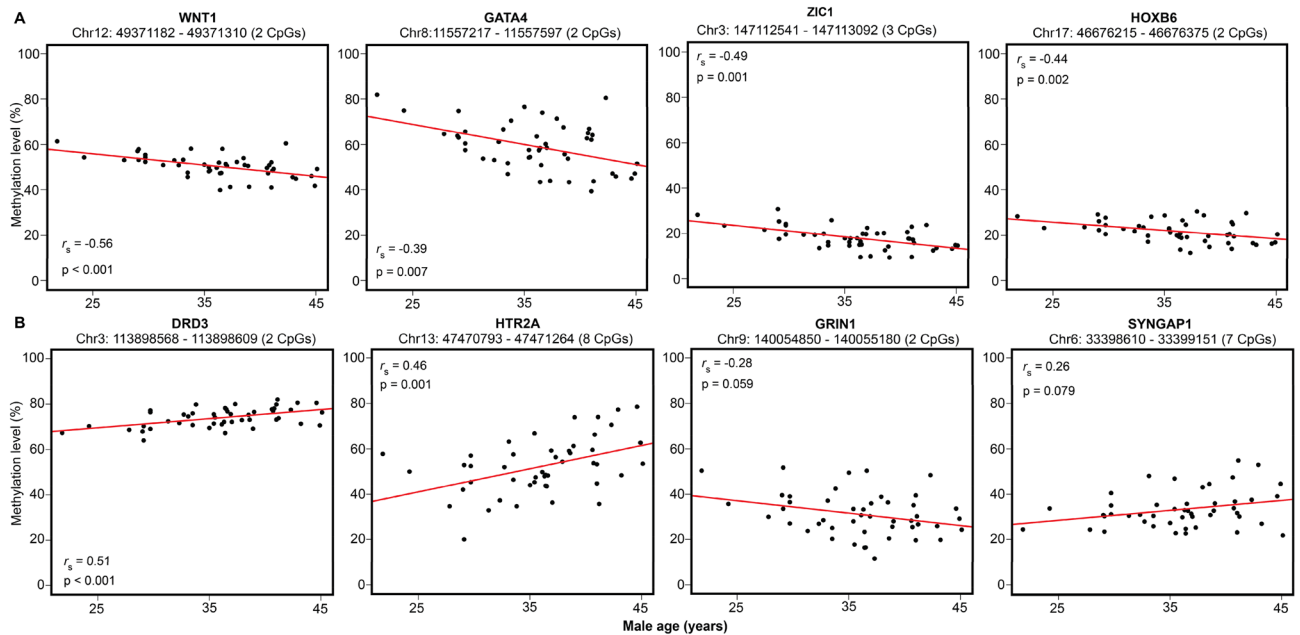

Figure 5. Scatterplots of age-related sperm DNA methylation. Selected genes associated with (**A**) early embryonic development (**B**) behavioral patterns. Genomic locations based on GRCh37/hg19 2009 human assembly.

Figure 6. Unsupervised clustering of age-related sperm methylation at individual CpG sites (n = 1698). Boxes below the heatmap are color-coded representing individual-level data (n = 47) including quartiles of male age, male fertility status (fertile vs. infertile), and successful live birth (yes or no).
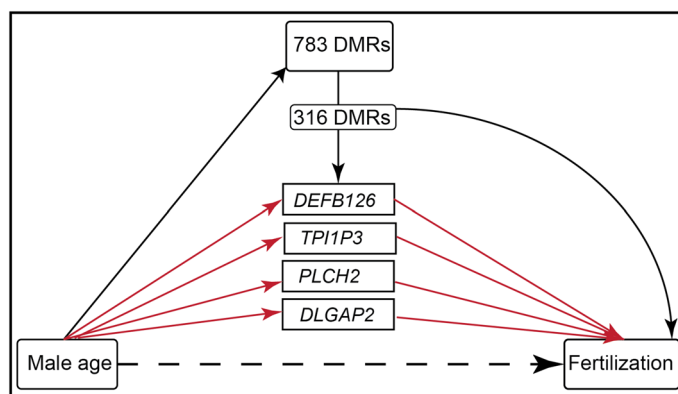
Figure 7. Schematic representation of mediation analyses for fertilization. The estimated association is provided on each pathway: the association between male age and each mediator; the association between each mediator and the outcome, and finally the remaining association of male age on the outcome, which is not mediated the DMRs (i.e. the direct effect). Total effects (TE) of age on outcomes is considered to be the sum of natural direct effect (dashed arrow) and natural indirect effect (red arrows).

ability of fertilization and live birth, and to identify candidate DMRs as mediators. Among the 783 age-associated sperm DMRs within 1500 bp of the transcriptional start site of genes, 316 sperm DMRs were significantly associated (q < 0.05) with fertilization. Further screening by sure independence screening (SIS) identified four potential sperm DMR mediators: *DEFB126, TPI1P3, PLCH2, and DLGAP2* (Fig. 7 and Supplementary Table S7). The overall natural indirect effect (NIE) for the four DMR mediators is associated with lower odds of fertilization (NIE: 0.96, p < 0.05), suggesting that each 1 year higher male age was associated with approximately 4% lower odds of fertilization acting through sperm DNA methylation at these four DMRs. Additionally, the NIE of age on fertilization acting through these four sperm DMRs accounted for 63% (95% CI 41%, 86%; p < 0.05) of the total effect of male age on fertilization (OR = 0.92; 95% CI 0.89, 0.96).

A similar process was used to evaluate mediators of age effect on live birth as described previously for fertilization. High dimensional screening of all 783 age-associated sperm DMRs suggested three potential mediators consisting of four genes, *CHRNB1, FGF11, AC005017.2*, and *MYOF* (Supplementary Figure S1 and Table S7). The mediation analysis suggested that each 1 year higher male age was indirectly associated with lower odds of live birth, acting through methylation of these genes (NIE: 0.95, p < 0.05). The proportion of the OR representing the total effect of male age on live birth mediated by these DMRs was 81% (95% CI − 36.4%, 198%).

Discussion

With secular trends toward delayed parenthood in developed countries, gamete quality is of clinical significance among couples or partners with advanced reproductive age. The utility of epigenetic aging based on measurements in somatic tissue has been highlighted in numerous studies[33,34]. In contrast, epigenetic aging in male germs cells on reproductive outcomes has been largely unaddressed to date. In this study, we show that male age is associated with extensive changes in sperm DNA methylation profiles, both at individual CpG sites and regional levels. Age-associated methylation was enriched at genes known for their role in early-life development and neurobehavior among others. To our knowledge, our results are the first to identify sperm DNA methylation as a likely pathway by which male age is associated with diminished reproduction outcomes in humans.

Spermatogonia are highly proliferative and self-renewing. Spermatogenesis commences at puberty and continues throughout adulthood, so that stochastic or induced changes in DNA methylation likely accumulate with age during the reproductive lifespan of the adult male. In regional analyses, sperm DMRs were associated with 784 genes that were significantly enriched in pathways important for biological functions such as embryonic, structural and central nervous system development, and cellular signaling. Of particular relevance are the findings that similar gene ontology (GO) terms were found among DMRs in CD4+ T cells comparing newborns and centenarians[39] as well as in CD4+ cells comparing middle aged (40–55 years) and long-lived (> 90 years) individuals[40]. Within these broad GO categories, age-associated sperm DMRs were identified in several key genes that encode for transcription factors essential embryogenesis and neurodevelopment such as *WNT1, GATA4, ZIC1*, and *HOXB6*. Interestingly, age-associated methylation changes in mouse sperm were also enriched in genes important for Wnt/β-catenin signaling pathway[13]. Furthermore, while *ZIC1* encodes a transcription factor necessary for normal cerebellar development in human[41], its mutation has been linked to cranial malformation and learning impairment in children[41]. Lastly, previous research has indicated that nucleosome-associated regions of sperm DNA, as compared to protamine-associated regions, may be important for embryonic development[42,43]. However, we found that age-associated sperm DMRs, contrary to our a priori hypothesis, were depleted, not enriched, in nucleosome regions, suggesting that sperm DNA methylation may mediate the influence of male age on reproductive outcomes independent of retained nucleosomes.

Advanced male age at conception has also been linked to an increase in risk in neurobehavioral issues in offspring. Interestingly, accumulating epidemiological evidence shows methylation signatures of neuro-developmental genes in sperm are particularly sensitive to advanced male age[23,44]. Similarly, in mice, age-associated sperm DNA methylation were transmitted to offspring with effect on both behavior and gene expression at regions known to be linked with schizophrenia and bipolar disorder[45]. In our analyses, we also identified several genes that were differentially methylated by male age including *DRD3, SYNGAP1, HTR2A*, and *GRIN1*. Hypermethylation of *DRD3*[46] and *HTR2A*[47] has been reported to be significantly associated with the risk of schizophrenia, while hypomethylation of *GRIN1* has been linked to depression in children[48]. Genetic alterations of *SYNGAP1* has also been linked to social dysfunctions and intellectual disability[49,50]. Together, these findings suggest that DNA methylation profiles of sperm from older fathers may contribute to the increase risk of abnormal embryo development and may play an important role in the neurodevelopment of offspring.

Numerous studies have attributed the association between higher male age and poorer reproductive outcomes to be driven by age-associated increases in sperm DNA damage[51], chromosomal abnormalities[52], as well as poor semen parameters[53,54]. Although some studies have reported no associations between sperm methylation and ART outcomes, they are limited by the use of global measures of methylation[55] and/or limited to targeted approaches at imprinted and Y chromosome loci[56]. In contrast, genome-wide analyses showed that distinct sperm methylation profiles associated with embryo quality status among participants seeking infertility treatment[57], suggesting that sperm methylation profiles could provide predictive value on subseqent embryo development. In our mediation analyses, we found that sperm DMRs in four genes, *DEFB126, PLCH2, TPI1P3* and *DLGAP2*, significantly mediated the relationship between male age and diminished fertilization. Of particular significance are the findings for *DEFB126*, which is a member of the beta-defensin protein family and is involved in several aspects of sperm function. DEFB126 is a main component of glycocalyx, which covers the entire surface of sperm and confers protection against the female immune system, enhances motility and facilitates binding to the zona pellucida[58–60]. Not surprisingly, loss of function of DEFB126 in males is associated with lower reproductive success. For example, men who are homozygous for the common *DEFB126* 2-nt frameshift deletion (*del/del*), have longer time-to-pregnancy among couples from the general population[61] and lower clinical pregnancy rates among couples undergoing intrauterine insemination[62]. Additionally, sperm from *del/del* males exhibited an 84% reduction in penetration assays, a surrogate for cervical mucus, compared to other genotypes[61]. Similarly, a 4-nt frameshift deletion in *DEFB126* has also been associated with male infertility[63]. Most recently, co-culture of immotile testicular sperm with full-length DEFB126 protein compared to mutated isoforms resulted in increased sperm motility[64]. Our results suggest that the age-dependent increase in *DEFB126* methylation might also contribute to declines in reproductive success.

In addition, *DLGAP2* functions in synapse organization and signaling in neuronal cells and is implicated as a candidate gene for autism and schizophrenia[65]. Aberrant methylation of *DLGAP2* in sperm has been linked to male infertility[66], as well as cannabis use in humans and THC exposure in rats[67]. Interestingly, rat offspring whose fathers were exposed to THC displayed differential methylation in the nucleus accumbens, suggesting an intergenerational inheritance of sperm methylation[67]. Despite the cell signaling function of *PLCH2*[68], as well as the overexpression of the pseudogene, *TPI1P3*, in the testis, their direct involvement in mediating the effects of male age on fertilization is unknown.

Our high dimensional mediation analyses for male age on live birth also identified three sperm DMRs that were associated with four genes, *CHRNB1, FGF11, MYOF*, and *AC005017.2*. Of potential relevance, *FGF11* is a member of the fibroblast growth factor gene family, which is involved in a numerous biological functions such as embryonic and nervous system development[69,70]. *MYOF*, a member of the ferlin gene family, displays calcium mediated membrane functions for muscle functions in humans[71]; however, ferlin genes in other species such as *C. elegans* and Drosophila melanogaster are essential for reproductive success as mutations result in male sterility due to defects in fertilization[72,73]. *CHRNB1*, which shares a CpG island with *FGF11*, encodes the beta subunit of the muscle acetylcholine receptor; however, it is unknown how age-associated changes in sperm methylation of *CHRNB1* or the pseudogene, *AC005017.2*, influences the probability of live birth.

Age-associated alterations in DNA methylation are profound, providing the basis for use of statistical models to predict chronological age from methylation profiles in most somatic tissue[30,31]. The 1698 CpGs we identified with age-associated methylation in sperm are distinct from those used for chronological age prediction in somatic tissue, such that only 3 CpGs overlapped with Levine et al.'s 513 phenoage CpGs[33] (Supplementary Table S8)

while zero CpGs overlapped with Hannum's 89 CpGs[31] and Horvath's 353 CpGs[30]. To date, one study by Jenkins et al. has examined the predictive ability of sperm methylation profiles on chronological age. Among 329 samples from infertile patients, sperm donors, and individuals from the general population, sperm methylation of 51 genomic regions via Illumina's 450K that reproducibly predicted individual's chronological age regardless of fertility status[35]. However, only 30% of these unique genomic regions used for these prediction models overlapped with our age-associated sperm DMRs (Supplementary Table S9). Possible explanations for the minimal overlap between age-associated sperm DMRs include differences in age distribution, such that our study only included men seeking IVF treatment and resulted in a narrower but clinically relevant age range of 22–45 years compared to Jenkins et al. with a male age range of 22–70 years and included men from the general population and those seeking fertility treatment. Additionally, technical variation in data processing, normalization, and statistical analysis approach (i.e., modeling m-values vs. beta-values for DNA methylation) could also account for differences in sperm DMRs. For example, 24 CpGs used in the Jenkins' age prediction calculator were removed prior to our analyses in our multi-step data preprocessing pipeline due to potential cross-hybridization and localization near known SNPs. Not surprisingly, Jenkins' age prediction calculator yielded a slightly attenuated correlation in our samples (actual vs predicted age $R^2 = 0.59$; $p < 0.001$) compared with their 10 individual independent cohort validation ($R^2 = 0.77$; $p = 0.0016$)[35]. Interestingly, when we stratified our results based on male infertility status or successful live birth, similar associations between actual and predicted age were found compared to Jenkins' age prediction calculator (Supplementary Figure S2: infertile group; $R^2 = 0.75$; $p < 0.001$ or no successful live birth; $R^2 = 0.77$; $p < 0.001$). Lastly, it must be noted that the age-associated sperm DMRs in our study were slightly biased towards hypermethylation (57.0%), which is consistent with a recent study by Denomme et al.[44] who also reported a slight bias toward hypermethylation (56.7%) of sperm DMRs. However, these findings contradict those of Jenkins et al. where only 5.5% of age-related sperm DMRs were hypermethylated[23]. The discrepancy of the direction of methylation change with age between these studies warrants further investigation in additional cohort studies.

Our study has several strengths. First, one unique feature of this study was the utilization of gradient centrifugation to isolate only motile fraction of the sperm, which represent the fraction with highest fertilization capacity. In a non-clinical natural pregnancy setting, sperm with higher motility have greater likelihood of fertilizing oocytes, which may help make these findings applicable to the general population. Additionally, our methylation analyses utilized the aliquots of the same semen samples that were used for ART procedures thus eliminating potential biases associated with obtaining independent semen samples. Moreover, use of gradient centrifugation represents standard clinical practice, so that these results have clinical relevance. Nevertheless, we recognize that there are some limitations to our study. Our novel findings are from a modest sample size of 47 couples. Additionally, participants include a mix of couples with male and/or female factor infertility seeking infertility treatment, and infertility status may play a complex role in the relationship between epigenetics and reproductive outcomes. To address this issue, we adjusted for infertility status and utilized stratified models in sensitivity analyses. Stratification by either male or female fertility status did not appreciably change the overall interpretation of our IVF outcome findings (Supplementary Table S1). We also recognize that assessment of DNA damage, such as DNA strand breaks, were not measured in our study and could possibly influence the effect of age on reproductive outcomes. Taken together, larger multi-center studies are warranted to validate and further refine our results with the inclusion of additional clinical information, such as hormone profiles and DNA damage, and to ensure generalizability across different infertility treatments.

Conclusion

Through the use of genome-wide methylation analyses, our study is the first, to our knowledge, to identify methylation profiles that may serve as a potential mechanism by which male age negatively affects reproductive success (e.g., fertilization and live birth rates) in an ART setting. Such information may inform clinical care by identifying couples who have poor prognosis for fertilization success from infertility treatments such as timed intercourse, intrauterine insemination and traditional IVF compared to ICSI.

Materials and methods

Study population. The Sperm Environmental Epigenetics and Developments Study (SEEDS) is a prospective observational cohort study aimed at investigating the associations of male preconception endocrine disrupting chemical exposure with sperm epigenetics and subsequent early-life development. Participants were recruited from couples undergoing fertility treatment at Baystate Medical Center located in Springfield, Massachusetts. The inclusion criteria were male partners between 18 and 55 years old without vasectomy and fresh ejaculate sperm used for IVF treatment. Relevant demographics (race, age, height, and weight), lifestyle factors (current and past alcohol and cigarette use), and medical history (diagnoses of infertility) data were collected by clinic personnel during the IVF cycle for both partners. Male fertility diagnoses were based on semen quality parameters according to 2010 World Health Organization (WHO) reference values[74] while female fertility diagnoses were based on CDC criteria such as polycystic ovarian syndrome, anovulation, tubal factors, diminished ovarian reserve, endometriosis and other factors[75]. Females received ovarian stimulation protocols based on infertility diagnoses and prior stimulation history, and egg retrieval was carried out about 36 h after the leading follicle reached 18–20 mm diameter as previously described[76]. The analytic sperm sample considered for the analyses described here comprised the first 48 men recruited for SEEDS with available data for sperm 450K methylation. One sample was removed from analyses based on assessment of outliers for DNA methylation. Written informed consent was obtained from eligible males who showed voluntary interest in participating in the study by the attending physicians. The study is approved by the institutional review boards at Baystate Medical Center and at the University of Massachusetts Amherst (IRB number: BH-12-190).

Embryo quality assessment. Embryos were evaluated on a five-point scale (1 being the best and 5 being the worst quality) on days 3 and 5 post-insemination, during the cleavage stage and blastocyst stage, respectively, as previously described[77]. A separate category was reserved for all arrested embryos at both time points. At the cleavage stage, embryos were evaluated morphologically for the presence or absence of blastomere multinucleation, symmetry, cell number, and amount of fragmentation. At the blastocyst stage, embryos were evaluated for the developmental stage including the expansion of blastocoel and quality of trophectoderm and inner cell mass. At both time points, grades 1–2 were classified as high quality embryos, similar to that in our previously published study[77], while at day 5, grades 1–4 were additionally classified as transferrable quality.

Sperm collection and DNA isolation. Fresh semen samples were collected in a sterile plastic specimen cup after a recommended abstinence period of at least 48 h according to standard IVF protocol. To separate motile sperm from somatic cells as well as abnormal and non-motile sperm cells, semen samples were processed using a two-step gradient fractionation (40% and 80%). Sperm DNA from motile-enriched fraction were isolated according to our previously published protocol[78] and used to generate sperm methylation data. As part of the routine protocol, trained embryologists microscopically examined all samples post-gradient fractionation, which were observed to be free of somatic cell contamination.

To further confirm findings were not influenced by somatic contamination, we analyzed methylation at imprinted genes, which have been published in a previous publication[28]. We showed that across identified 203 imprinted loci of 18 imprinted genes[79], methylation levels of all maternal and paternal imprinted loci were < 10% or > 85% respectively, indicating negligible somatic cell contamination in the sperm samples. Here, we present methylation of maternal imprinted gene, *DLK1*, known to be differentially methylated between sperm and somatic tissues[35] (Supplementary Figure S3). Average methylation for all participants was below 5% and male participants who were classified as infertile (shown in red) according to 2010 WHO cutoffs[74] displayed no difference in average methylation at *DLK1* locus compared to fertile men (mean (SD): 2.87% (1.22) and 2.85% (1.06), respectively; p = 0.90). This data provides strong evidence that the reliability and validity of our results are not compromised by somatic contamination and participants' fertility status.

450K analyses. Genomic sperm DNA (400 ng) was analyzed using Illumina 450K Infinium Methylation Beadchip, which allows a genome-wide coverage of over 450,000 methylation sites. Batch effects were minimized by randomizing samples within and across beadchip. The beadchip was run at the Applied Genomics Technology Center at Wayne State University. We performed within-array normalization and dye-bias equalization of Type I and Type II probes[80] using normal-exponential convolution method (Noob)[81] via minfi package in R[82]. Batch effects were corrected with the ComBat function in the sva package[83] while cross-hybridizing probes and sex chromosome probes were removed using the DMRcate package[84]. Of the total 485,512 individual CpG sites interrogated for each participant, 75,164 sites were excluded after preprocessing, leaving a total of 410,348 CpG sites available for downstream analyses. The raw data can be found on GEO Accession Viewer (Accession Number GSE102970).

Sequenom validation. Validation of our 450K data was performed using Sequenom mass array for both our significant and non-significant CpGs and the results are reported in our previous publication[28]. All show excellent agreement between the two platforms (e.g., cg14156792, Spearman r = 0.939, p < 2.2E−16; cg24997886, r = 0.911, p < 2.2E−16; cg19101893, r = 0.785, p = 6.25E−11). The two CpGs associated with male age (cg07462448 and cg04473763) displayed strong correlations (Spearman r = 0.89, p < 2.2E−16 and r = 0.77, p = 2.43E−10, respectively).

All methods were performed in accordance with the relevant guidelines and regulations.

Statistical analyses. We first sought to determine the association between male age and ART outcomes: fertilization status, high quality embryos, and transferable quality status. Since we have repeated measurements of ART outcomes (multiple oocytes per couple), we modeled these outcomes as a proportion. The fertilization outcome is defined as the proportion of successful fertilization attempts for each couple; and similarly, for high quality embryo and transferable quality outcomes. This modeling approach makes it explicit that our unit of analysis is a couple and not an oocyte. We also included the number of oocytes per couple as observation weights in all logistic regression analyses performed involving fertilization, high quality embryos and transferable quality outcomes. The weighted logistic regression model accounts for the variability in the number of repeated measurements for each couple. All estimates represent the increase in the odds of the outcome associated with a 1-year higher male age, after adjusting for male body mass index (BMI), male infertility status based on 2010 WHO reference values[74], male smoking and female age. These covariates were selected based on biological plausibility and after univariate screening with the outcome (p < 0.1). Statistical significance was defined as p-values < 0.05.

We then sought to determine the association between male age and sperm DNA methylation at individual CpG sites and regionally. To balance both validity and interpretability of our analyses, we conducted statistical analyses using both β-values and M-values, defined as the logit transformation of the β-values log (β/(1 − β)). To identify differences in percent methylation associated with male age, we used M-values due to its better adherence to homoscedasticity in linear models[85]. However, to facilitate the biological interpretation of our results, we reported effect estimates from models using beta-values, which provide CpG methylation values between 0 and 100%.

Associations between male age and sperm methylation were first analyzed at individual CpG sites using CpGassoc[86]. Next, regional analyses used A-clustering algorithm[87] to identify co-regulated CpGs (≥ 2 correlated CpG sites) within 1000 base pairs. A total of 7771 co-regulated regions were identified and formed the unit of our

regional methylation analyses. General estimated equations (GEE) were used to identify differential methylated regions (DMRs) by using a linear link and an exchangeable correlation to account for the correlated structure of the CpGs sites within a region[88]. All methylation analyses were adjusted for male covariates based on biological plausibility and included BMI, smoking status, and male infertility status. All p-values were adjusted for multiple testing via Benjamini-Hochberg (q < 0.05) and Bonferroni correction.

Bioinformatics analyses. Prior to ontology analyses, each methylation region was assigned the closest gene within 1500 bp of TSS using GRCh37 assembly data from ENSEMBL via the annotatePeakInBatch function from the ChIPpeakAnno R package (version 3.6.5). We used metascape (http://metascape.org)[89] to determine the enrichment of age-associated sperm DMRs with known biological pathways and regulation. We determined genic (enhancers, promoters, exons, introns, intergenic regions) and CpG features (island, shelves, shores, open sea) of methylated regions via annotations from Ensembl. For nucleosome retention vs. protamine designation of methylation regions, the publicly available dataset from Donkin et al. was used[38]. Binding site locations for transcription factors such as *EZH2* and *CTCF* were retrieved from ENCODE. Fisher's exact test was used to test for significant enrichment or depletion of each feature in the statistically significant DMRs (GEE models; q < 0.05) versus the entire set of methylated regions.

Mediation analysis. Analysis of mediation was conducted to evaluate sperm DNA methylation as a pathway through which male age might be related to ART outcomes. Analysis of the overall association of male age with these outcomes was described previously. Having identified the male age-associated DMRs, further screening was conducted to focus on those 783 DMRs that fall within 1500 bp of the TSS of the nearest gene, making them more likely to play a role in gene regulation. We used a two-stage screening procedure to reduce the number of DMRs to be evaluated in mediation analysis. In the first stage, we ran multiple logistic regressions of the outcome (as described previously) on each DMR adjusting for female age, male BMI, male smoking and male infertility status, and computed BH-adjusted p-values (q < 0.05) for the significance of each DMR. In the second stage, we used the R package Sure Independence Screening (SIS) followed by Minimum Concave Penalty (MCP) to further reduce the set of DMRs found significant in the first stage to be smaller than the sample size[90]. We performed a 'leave-one-out' approach where we iteratively removed one observation at a time from the data set and performed DMR screening using SIS; then we selected the DMRs that were identified by SIS in at least 20% of the iterations to be the final set of DMRs used in the mediation analyses. We then performed a regression-based multiple mediation analysis[91] to assess mediation of the association between male age and ART outcomes (fertilization proportion and live birth status) by sperm methylation using this final set of DMRs. Analyses controlled for female age, male BMI, male smoking, and male infertility status. As described previously, analyses for the fertilization outcome included weights to account for the total number of oocytes per couple; however, weights were not used for models of live birth since only a single embryo was considered. For each outcome, we reported estimates and confidence intervals for the *natural indirect effect (NIE)*—defined as the effect of a 1 year increase in male age on odds of the outcome transmitted through the sperm DMR, and the *mediation proportion*—the proportion of the total effect of male age on odds of the outcome represented by the NIE indicating the association of male age with outcomes via sperm DNA methylation.

We note that our mediation analysis utilized male age and DMR values that were on continuous scales while fertilization and live birth outcomes were modeled on logistic scales. We followed the adjustment procedure proposed[92]. This approach entails computing a rescaling factor to correct for this incompatibility for the sake of validity of statistical inference and interpretation of the natural indirect effect and proportion mediated.

References

1. Khandwala, Y. S., Zhang, C. A., Lu, Y. & Eisenberg, M. L. The age of fathers in the USA is rising: An analysis of 168 867 480 births from 1972 to 2015. *Hum. Reprod.* **32**, 2110–2116. https://doi.org/10.1093/humrep/dex267 (2017).
2. Waldenström, U. Postponing parenthood to advanced age. *Upsala J. Med. Sci.* **121**, 1–9. https://doi.org/10.1080/03009734.2016.1201553 (2016).
3. Bray, I., Gunnell, D. & Smith, G. D. Advanced paternal age: How old is too old?. *J. Epidemiol. Community Health* **60**, 851–853. https://doi.org/10.1136/jech.2005.045179 (2006).
4. Mills, M. *et al.* Why do people postpone parenthood? Reasons and social policy incentives. *Hum. Reprod. Update* **17**, 848–860. https://doi.org/10.1093/humupd/dmr026 (2011).
5. Hemminki, K., & Kyyrönen, P. Parental age and risk of sporadic and familial cancer in offspring: Implications for germ cell mutagenesis. *Epidemiology* **10**(6), 747–751 (1999).
6. Jacobsson, B., Ladfors, L. & Milsom, I. Advanced maternal age and adverse perinatal outcome. *Obstet. Gynecol.* **104**, 727–733. https://doi.org/10.1097/01.AOG.0000140682.63746.be (2004).
7. Kemkes-Grottenthaler, A. Parental effects on offspring longevity—Evidence from 17th to 19th century reproductive histories. *Ann. Hum. Biol.* **31**, 139–158. https://doi.org/10.1080/03014460410001663407 (2004).
8. Hassan, M. A. M. & Killick, S. R. Effect of male age on fertility: Evidence for the decline in male fertility with increasing age. *Fertil. Steril.* **79**, 1520–1527. https://doi.org/10.1016/S0015-0282(03)00366-2 (2003).
9. Horta, F. *et al.* Male ageing is negatively associated with the chance of live birth in IVF/ICSI cycles for idiopathic infertility. *Hum. Reprod.* **34**, 2523–2532. https://doi.org/10.1093/humrep/dez223 (2019).
10. Montgomery, S. M., Olsson, T. & Ekbom, A. Parental age, family size, and risk of multiple sclerosis. *Epidemiology* **15**, 717–723. https://doi.org/10.1097/01.ede.0000142138.46167.69 (2004).
11. Puleo, C. M. *et al.* Advancing paternal age and autism. *Autism* **16**, 367–380. https://doi.org/10.1177/1362361311427154 (2012).

12. Saha, S. *et al.* Advanced paternal age is associated with impaired neurocognitive outcomes during infancy and childhood. *PLoS Med.* **6**, e40–e40. https://doi.org/10.1371/journal.pmed.1000040 (2009).

13. Xie, K. *et al.* Epigenetic alterations in longevity regulators, reduced life span, and exacerbated aging-related pathology in old father offspring mice. *Proc. Natl. Acad. Sci. USA.* **115**, E2348–E2357. https://doi.org/10.1073/pnas.1707337115 (2018).

14. Contreras, Z. A. *et al.* Parental age and childhood cancer risk: A Danish population-based registry study. *Cancer Epidemiol.* **49**, 202–215. https://doi.org/10.1016/j.canep.2017.06.010 (2017).

15. Sartorelli, E. M. P., Mazzucatto, L. F. & de Pina-Neto, J. O. M. Effect of paternal age on human sperm chromosomes. *Fertil. Steril.* **76**, 1119–1123 (2001).

16. Gratten, J. *et al.* Risk of psychiatric illness from advanced paternal age is not predominantly from de novo mutations. *Nat. Genet.* **48**, 718–724. https://doi.org/10.1038/ng.3577 (2016).

17. Reichenberg, A. *et al.* Advancing paternal age and autism. *Arch. Gen. Psychiatry* **63**, 1026–1032 (2006).

18. Snajderova, M. *et al.* The importance of advanced parental age in the origin of neurofibromatosis type 1. *Am. J. Med. Genet. Part A* **158A**, 519–523. https://doi.org/10.1002/ajmg.a.34413 (2012).

19. Johnson, S. L., Dunleavy, J., Gemmell, N. J. & Nakagawa, S. Consistent age-dependent declines in human semen quality: A systematic review and meta-analysis. *Ageing Res. Rev.* **19**, 22–33. https://doi.org/10.1016/j.arr.2014.10.007 (2015).

20. Rosiak-Gill, A. *et al.* Age-related changes in human sperm DNA integrity. *Aging* https://doi.org/10.18632/aging.102120 (2019).

21. Kong, A. *et al.* Rate of de novo mutations, father's age, and disease risk. *Nature* **488**, 471–475. https://doi.org/10.1038/nature11396.Rate (2013).

22. Miller, B. *et al.* Advanced paternal age and parental history of schizophrenia. *Schizophr. Res.* **133**, 125–132. https://doi.org/10.1016/j.schres.2011.08.013 (2011).

23. Jenkins, T. G., Aston, K. I., Pflueger, C., Cairns, B. R. & Carrell, D. T. Age-associated sperm DNA methylation alterations: Possible implications in offspring disease susceptibility. *PLoS Genet.* https://doi.org/10.1371/journal.pgen.1004458 (2014).

24. Watkins, A. J. *et al.* Paternal diet programs offspring health through sperm- and seminal plasma-specific pathways in mice. *Proc. Natl. Acad. Sci.* **115**, 10064–10069. https://doi.org/10.1073/pnas.1806333115 (2018).

25. Murphy, S. K. *et al.* Cannabinoid exposure and altered DNA methylation in rat and human sperm. *Epigenetics* **13**, 1–14. https://doi.org/10.1080/15592294.2018.1554521 (2018).

26. Eva Benito, A. *et al.* RNA-dependent intergenerational inheritance of enhanced synaptic plasticity after environmental enrichment data and software availability GSE110183. *Cell Rep.* **23**, 546–554. https://doi.org/10.1016/j.celrep.2018.03.059 (2018).

27. Wu, L. *et al.* Paternal psychological stress reprograms hepatic gluconeogenesis in offspring. *Cell Metab.* **23**, 735–743. https://doi.org/10.1016/J.CMET.2016.01.014 (2016).

28. Wu, H. *et al.* Preconception urinary phthalate concentrations and sperm DNA methylation profiles among men undergoing IVF treatment: A cross-sectional study. *Hum. Reprod.* **32**, 2159–2169. https://doi.org/10.1093/humrep/dex293 (2017).

29. Zhang, Y., Hapala, J., Brenner, H. & Wagner, W. Individual CpG sites that are associated with age and life expectancy become hypomethylated upon aging. *Clin. Epigenet.* **9**, 1–6. https://doi.org/10.1186/s13148-017-0315-9 (2017).

30. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115. https://doi.org/10.1186/gb-2013-14-10-r115 (2013).

31. Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367. https://doi.org/10.1016/j.molcel.2012.10.016 (2013).

32. Zhang, Y. *et al.* DNA methylation signatures in peripheral blood strongly predict all-cause mortality. *Nat. Commun.* **8**, 14617. https://doi.org/10.1038/ncomms14617 (2017).

33. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591. https://doi.org/10.18632/aging.101414 (2018).

34. Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327. https://doi.org/10.18632/aging.101684 (2019).

35. Jenkins, T. G., Aston, K. I., Cairns, B., Smith, A. & Carrell, D. T. Paternal germ line aging: DNA methylation age prediction from human sperm. *BMC Genomics* **19**, 763. https://doi.org/10.1186/s12864-018-5153-4 (2018).

36. Jenkins, T. G., Aston, K. I., Pflueger, C., Cairns, B. R. & Carrell, D. T. Age-associated sperm DNA methylation alterations: Possible implications in offspring disease susceptibility. *PLoS Genet.* **10**, e1004458. https://doi.org/10.1371/journal.pgen.1004458 (2014).

37. Jenkins, T. G. *et al.* Age-associated sperm DNA methylation patterns do not directly persist trans-generationally. *Epigenet. Chromatin* **12**, 74. https://doi.org/10.1186/s13072-019-0323-4 (2019).

38. Donkin, I. *et al.* Obesity and bariatric surgery drive epigenetic variation of spermatozoa in humans. *Cell Metab.* **23**, 369–378. https://doi.org/10.1016/j.cmet.2015.11.004 (2016).

39. Heyn, H. *et al.* Distinct DNA methylomes of newborns and centenarians. *Proc. Natl. Acad. Sci.* **109**, 10522–10527. https://doi.org/10.1073/pnas.1120658109 (2012).

40. Zhao, M. *et al.* Distinct epigenomes in CD4(+) T cells of newborns, middle-ages and centenarians. *Sci. Rep.* **6**, 38411. https://doi.org/10.1038/srep38411 (2016).

41. Twigg, S. R. F. *et al.* Gain-of-function mutations in ZIC1 are associated with coronal craniosynostosis and learning disability. *Am. J. Hum. Genet.* **97**, 378–388. https://doi.org/10.1016/J.AJHG.2015.07.007 (2015).

42. Arpanahi, A. *et al.* Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. *Genome Res.* **19**, 1338–1349. https://doi.org/10.1101/gr.094953.109 (2009).

43. Hammoud, S. S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**, 473–478. https://doi.org/10.1038/nature08162 (2009).

44. Denomme, M. M., Haywood, M. E., Parks, J. C., Schoolcraft, W. B. & Katz-Jaffe, M. G. The inherited methylome landscape is directly altered with paternal aging and associated with offspring neurodevelopmental disorders. *Aging Cell* **19**. https://doi.org/10.1111/acel.13178 (2020).

45. Milekic, M. H. *et al.* Age-related sperm DNA methylation changes are transmitted to offspring and associated with abnormal behavior and dysregulated gene expression. *Mol. Psychiatry* **20**, 995–1001. https://doi.org/10.1038/mp.2014.84 (2015).

46. Dai, D. *et al.* Significant association between DRD3 gene body methylation and schizophrenia. *Psychiatry Res.* **220**, 772–777. https://doi.org/10.1016/J.PSYCHRES.2014.08.032 (2014).

47. Cheah, S.-Y., Lawford, B. R., Young, R. M., Morris, C. P. & Voisey, J. mRNA expression and DNA methylation analysis of serotonin receptor 2A (HTR2A) in the human schizophrenic brain. *Genes* **8**, 14. https://doi.org/10.3390/genes8010014 (2017).

48. Weder, N. *et al.* Child abuse, depression, and methylation in genes involved with stress, neural plasticity, and brain circuitry. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 417-424.e415. https://doi.org/10.1016/j.jaac.2013.12.025 (2014).

49. Hamdan, F. F. *et al.* De novo SYNGAP1 mutations in nonsyndromic intellectual disability and autism. *Biol. Psychiatry.* **69**, 898–901. https://doi.org/10.1016/J.BIOPSYCH.2010.11.015 (2011).

50. Ozkan, E. D. *et al.* Reduced cognition in Syngap1 mutants is caused by isolated damage within developing forebrain excitatory neurons. *Neuron* **82**, 1317–1333. https://doi.org/10.1016/j.neuron.2014.05.015 (2014).

51. Moskovtsev, S. I., Willis, J. & Mullen, J. B. M. Age-related decline in sperm deoxyribonucleic acid integrity in patients evaluated for male infertility. *Fertil. Steril.* **85**, 496–499. https://doi.org/10.1016/j.fertnstert.2005.05.075 (2006).

52. Sloter, E., Nath, J., Eskenazi, B. & Wyrobek, A. J. Effects of male age on the frequencies of germinal and heritable chromosomal abnormalities in humans and rodents. *Fertil. Steril.* **81**, 925–943. https://doi.org/10.1016/j.fertnstert.2003.07.043 (2004).

12

53. Hossain, M. M., Fatima, P., Rahman, D. & Hossain, H. B. Semen parameters at different age groups of male partners of infertile couples. *Mymensingh Med. J.* **21**, 306–315 (2012).

54. Mukhopadhyay, D. *et al.* Semen quality and age-specific changes: A study between two decades on 3,729 male partners of couples with normal sperm count and attending an andrology laboratory for infertility-related problems in an Indian city. *Fertil. Steril.* **93**, 2247–2254. https://doi.org/10.1016/j.fertnstert.2009.01.135 (2010).

55. Benchaib, M. *et al.* Influence of global sperm DNA methylation on IVF results. *Hum. Reprod.* **20**, 768–773. https://doi.org/10.1093/humrep/deh684 (2005).

56. Montjean, D. *et al.* Methylation changes in mature sperm deoxyribonucleic acid from oligozoospermic men: Assessment of genetic variants and assisted reproductive technology outcome. *Fertil. Steril.* **100**, 1241–1247.e1242. https://doi.org/10.1016/j.fertnstert.2013.06.047 (2013).

57. Aston, K. I. *et al.* Aberrant sperm DNA methylation predicts male fertility status and embryo quality. *Fertil. Steril.* **104**, 1388-1397.e1385. https://doi.org/10.1016/j.fertnstert.2015.08.019 (2015).

58. Tollner, T. L., Bevins, C. L. & Cherr, G. N. Multifunctional glycoprotein DEFB126—A curious story of defensin-clad spermatozoa. *Nat. Rev. Urol.* **9**, 365–375. https://doi.org/10.1038/nrurol.2012.109 (2012).

59. Yudin, A. I. *et al.* Beta-defensin 126 on the cell surface protects sperm from immunorecognition and binding of anti-sperm antibodies. *Biol. Reprod.* **73**, 1243–1252. https://doi.org/10.1095/biolreprod.105.042432 (2005).

60. Tollner, T. L. *et al.* Beta-defensin 126 on the surface of macaque sperm mediates attachment of sperm to oviductal epithelia. *Biol. Reprod.* **78**, 400–412. https://doi.org/10.1095/biolreprod.107.064071 (2008).

61. Tollner, T. L. *et al.* A common mutation in the defensin DEFB126 causes impaired sperm function and subfertility. *Sci. Transl. Med.* **3**, 92ra65. https://doi.org/10.1126/scitranslmed.3002289 (2011).

62. Boroujeni, P. B. *et al.* The role of DEFB126 variation in male infertility and medically assisted reproduction technique outcome. *Reprod. Biomed. Online* https://doi.org/10.1016/j.rbmo.2019.05.012 (2019).

63. Duan, S. *et al.* Another functional frame-shift polymorphism of DEFB126 (rs11467497) associated with male infertility. *J. Cell Mol. Med.* **19**, 1077–1084. https://doi.org/10.1111/jcmm.12502 (2015).

64. Aram, R., Chan, P. T. K. & Cyr, D. G. Beta-defensin126 is correlated with sperm motility in fertile and infertile men. *Biol. Reprod.* https://doi.org/10.1093/biolre/ioz171 (2019).

65. Rasmussen, A. H., Rasmussen, H. B. & Silahtaroglu, A. The DLGAP family: Neuronal expression, function and role in brain disorders. *Mol. Brain* **10**, 1–13. https://doi.org/10.1186/s13041-017-0324-9 (2017).

66. Sujit, K. M. *et al.* Genome-wide differential methylation analyses identifies methylation signatures of male infertility. *Hum. Reprod.* **33**, 2256–2267. https://doi.org/10.1093/humrep/dey319 (2018).

67. Schrott, R. *et al.* Cannabis use is associated with potentially heritable widespread changes in autism candidate gene DLGAP2 DNA methylation in sperm. *Epigenetics* https://doi.org/10.1080/15592294.2019.1656158 (2019).

68. Zhou, Y., Wing, M. R., Sondek, J. & Harden, T. K. Molecular cloning and characterization of PLC-eta2. *Biochem. J.* **391**, 667–676. https://doi.org/10.1042/BJ20050839 (2005).

69. Smallwood, P. M. *et al.* Fibroblast growth factor (FGF) homologous factors: New members of the FGF family implicated in nervous system development. *Proc. Natl. Acad. Sci. USA.* **93**, 9850–9857. https://doi.org/10.1073/pnas.93.18.9850 (1996).

70. Zhong, W. *et al.* FGF ligand family mRNA expression profile for mouse preimplantation embryos, early gestation human placenta, and mouse trophoblast stem cells. *Mol. Reprod. Dev.* **73**, 540–550. https://doi.org/10.1002/mrd.20417 (2006).

71. Davis, D. B., Doherty, K. R., Delmonte, A. J. & McNally, E. M. Calcium-sensitive phospholipid binding properties of normal and mutant ferlin C2 domains. *J. Biol. Chem.* **277**, 22883–22888. https://doi.org/10.1074/jbc.M201858200 (2002).

72. Washington, N. L. & Ward, S. FER-1 regulates Ca^{2+}-mediated membrane fusion during *C. elegans* spermatogenesis. *J. Cell Sci.* **119**, 2552–2562. https://doi.org/10.1242/jcs.02980 (2006).

73. Smith, M. K. & Wakimoto, B. T. Complex regulation and multiple developmental functions of misfire, the *Drosophila melanogaster* ferlin gene. *BMC Dev. Biol.* **7**, 21. https://doi.org/10.1186/1471-213x-7-21 (2007).

74. Cooper, T. G. *et al.* World Health Organization reference values for human semen characteristics. *Hum. Reprod. Update* **16**, 231–245. https://doi.org/10.1093/humupd/dmp048 (2010).

75. Wright, V. C., Schieve, L. A., Reynolds, M. A. & Jeng, G. Assisted reproductive technology surveillance–United States, 2000. *MMWR Surveill. Summ.* **52**, 1–16 (2003).

76. Barsky, M., St Marie, P., Rahil, T., Markenson, G. R. & Sites, C. K. Are perinatal outcomes affected by blastocyst vitrification and warming?. *Am. J. Obstet. Gynecol.* **215**, 603.e601–603.e605. https://doi.org/10.1016/j.ajog.2016.06.002 (2016).

77. Wu, H. *et al.* Parental contributions to early embryo development: Influences of urinary phthalate and phthalate alternatives among couples undergoing IVF treatment. *Hum. Reprod.* **32**, 65–75. https://doi.org/10.1093/humrep/dew301 (2017).

78. Wu, H., de Gannes, M. K., Luchetti, G. & Pilsner, J. R. Rapid method for the isolation of mammalian sperm DNA. *Biotechniques* **58**, 293–300 (2015).

79. Nazor, K. L. *et al.* Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell* **10**, 620–634. https://doi.org/10.1016/j.stem.2012.02.013 (2012).

80. Niu, L., Xu, Z. & Taylor, J. A. RCP: A novel probe design bias correction method for Illumina Methylation BeadChip. *Bioinformatics* **32**, 2659–2663. https://doi.org/10.1093/bioinformatics/btw285 (2016).

81. Triche, T. J., Weisenberger, D. J., Van Den Berg, D., Laird, P. W. & Siegmund, K. D. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res.* **41**, 1–11. https://doi.org/10.1093/nar/gkt090 (2013).

82. Aryee, M. J. *et al.* Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369. https://doi.org/10.1093/bioinformatics/btu049 (2014).

83. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883. https://doi.org/10.1093/bioinformatics/bts034 (2012).

84. Peters, T. J. *et al.* De novo identification of differentially methylated regions in the human genome. *Epigenet. Chromatin* **8**, 6–6. https://doi.org/10.1186/1756-8935-8-6 (2015).

85. Du, P. *et al.* Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.* **11**, 587. https://doi.org/10.1186/1471-2105-11-587 (2010).

86. Barfield, R. T., Kiларu, V., Smith, A. K. & Conneely, K. N. CpGassoc: An R function for analysis of DNA methylation microarray data. *Bioinformatics* **28**, 1280–1281. https://doi.org/10.1093/bioinformatics/bts124 (2012).

87. Sofer, T., Schifano, E. D., Hoppin, J. A., Hou, L. & Baccarelli, A. A. Gene expression A-clustering: A novel method for the detection of co-regulated methylation regions, and regions associated with exposure. *Bioinformatics* **29**, 2884–2891. https://doi.org/10.1093/bioinformatics/btt498 (2013).

88. Liang, K.-Y. & Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22. https://doi.org/10.1093/biomet/73.1.13 (1986).

89. Tripathi, S. *et al.* Meta-and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. *Cell Host Microbe.* **18**(6), 723–735. https://doi.org/10.1016/j.chom.2015.11.002 (2015).

90. Saldana, D. F. & Feng, Y. SIS: An R package for sure independence screening in ultrahigh-dimensional statistical models. *J. Stat. Softw.* **83**, 1–25. https://doi.org/10.18637/jss.v083.i02 (2018).

91. VanderWeele, T. J. Mediation analysis: A practitioner's guide. *Annu. Rev. Public Health* **37**, 17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402 (2016).

92. MacKinnon, D. P. *Introduction to Statistical Mediation Analysis* 306 (Taylor & Francis, London, 2008).

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/s41598-020-80857-2.

Correspondence and requests for materials should be addressed to J.R.P.

Reprints and permissions information is available at www.nature.com/reprints.

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