scientific reports

Check for updates

OPEN

Short-term diet intervention comprising of olive oil, vitamin D, and omega-3 fatty acids alters the small non-coding RNA (sncRNA) landscape of human sperm

Candida Vaz[1,17]✉, Mark Burton[2,17], Alexandra J. Kermack[2,3,4,5,17], Pei Fang Tan[1,16], Jason Huan[1], Tessa P. X. Yoo[2], Kerry Donnelly[4], Susan J. Wellstead[4,5], Dennis Wang[1,16], Helena L. Fisk[2,3], Franchesca D. Houghton[2], Sheena Lewis[6,7], Yap Seng Chong[1,8], Peter D. Gluckman[1,9], Ying Cheong[2,3,4,5], Nicholas S. Macklon[2,3,10], Philip C. Calder[2,3], Anindya Dutta[11,12], Keith M. Godfrey[3,13], Pankaj Kumar[11], Karen A. Lillycrop[2,14] & Neerja Karnani[1,15,16]✉

Offspring health outcomes are often linked with epigenetic alterations triggered by maternal nutrition and intrauterine environment. Strong experimental data also link paternal preconception nutrition with pathophysiology in the offspring, but the mechanism(s) routing effects of paternal exposures remain elusive. Animal experimental models have highlighted small non-coding RNAs (sncRNAs) as potential regulators of paternal effects. Here, we characterised the baseline sncRNA landscape of human sperm and the effect of a 6-week dietary intervention on their expression profile. This study involves sncRNAseq profiling, that was performed on a subset (n = 17) of the participants enrolled in the PREPARE trial: 9 from the control group and 8 from the intervention group. 5'tRFs, miRNAs and piRNAs were the most abundant sncRNA subtypes identified; their expression was associated with age, BMI, and sperm quality. Nutritional intervention with olive oil, vitamin D and omega-3 fatty acids altered expression of 3 tRFs, 15 miRNAs and 112 piRNAs, targeting genes involved in fatty acid metabolism and transposable elements in the sperm genome.

PREPARE Trial registration number: ISRCTN50956936, Trial registration date: 10/02/2014.

Keywords Sperm, Diet intervention, Small non-coding RNA, miRNA, piRNA, tRF

[1]Institute for Human Development and Potential (IHDP), Agency for Science, Technology and Research (A*STAR), 30 Medical Drive, Singapore 117609, Republic of Singapore. [2]School of Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. [3]NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust and University of Southampton, Southampton, UK. [4]Complete Fertility, Princess Anne Hospital, Southampton, UK. [5]University Hospital Southampton NHS Foundation Trust, Southampton, UK. [6]Queen's University, Belfast, Northern Ireland, UK. [7]Examen Lab Ltd, Belfast, Northern Ireland, UK. [8]Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore. [9]Liggins Institute, University of Auckland, Auckland, New Zealand. [10]London Women's Clinic, London, UK. [11]Department of Biochemistry and Molecular Genetics, University of Virginia, School of Medicine, Charlottesville, VA, USA. [12]Department of Genetics, U. Alabama, Birmingham, AL 35294, USA. [13]MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK. [14]School of Biological Sciences, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK. [15]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore (NUS), Singapore, Republic of Singapore. [16]Agency for Science, Technology and Research (A*STAR), Bioinformatics Institute (BII), 30 Biopolis Street #07-01 Matrix, Singapore 138671, Republic of Singapore. [17]Candida Vaz, Mark Burton, Alexandra J. Kermack and Pei Fang Tan have contributed equally to this work. ✉email: candida_vaz@sics.a-star.edu.sg; neerja_karnani@sics.a-star.edu.sg

nature portfolio

1

Inheritance of genetic material has long been acknowledged as the primary route of transmitting health related information from parents to the offspring. However, emerging evidence from the DOHaD (Developmental Origins of Health and Disease) field[1,2] identifies epigenetic mechanisms as an additional route of transgenerational inheritance, especially in transferring the effects of parent's lifestyle and environmental exposures, with implications for the offspring's risk of non-communicable diseases (NCDs) in later life[3]. Epidemiological, clinical, and basic science research has characterised the period around conception as being critical in the processes linking parental environmental exposures with the health of the next generation[4]. Since these exposures and the related epigenetic perturbations constitute modifiable factors, they can serve as promising targets for developing effective interventions against NCDs. Current DOHaD research is primarily focused at understanding the role of maternal factors, such as nutrition[5], in shaping the development of the offspring. Less is known about the mechanisms involved in the transmission of paternal exposures through the male germline.

A variety of lifestyle and environmental factors ranging from nutrition[6–8], stress[9–11] and exposure to toxicants such as fungicides, pesticides[12–15] and air pollution[16] are known to alter sperm concentration and quality. It is also becoming apparent that these exposures may leave an imprint on the sperm epigenome, that may be transmitted to the new life at conception[17–19]. Studies from rodent models have provided some insights into the mechanisms involved in mediating these effects to the offspring[4,20,21]. These primarily involve alterations to DNA methylation and small non-coding RNA (sncRNA) expression[18]. Sperm sncRNA expression in rodents is known to be sensitive to paternal exposures such as diet, and capable of signaling its effects to the offspring[6,17,22].

sncRNAs include a variety of small RNA species, that are typically < 100 nucleotides long, and regulate diverse cellular processes[23]. Their size and structure can give sncRNAs greater stability, and mobility across tissues, making them ideal candidates for intercellular signaling and gene regulation[24]. These characteristics also make sncRNAs pivotal regulators in developmental and physiological processes. MicroRNAs (miRNAs) are an example of such small RNA species, that are typically 22–24 nucleotides long, and can post-transcriptionally regulate gene expression by base-pairing with complementary sequences on mRNA. They can modulate other epigenetic regulators, and conversely themselves be targets of epigenetic regulation[25,26]. Studies in mouse models have shown that stress[27] and diet[28] can alter miRNA expression in the male germline. Furthermore, injecting the environment responsive sperm miRNAs into the zygote/embryo can remodel offspring gene expression, and propagate paternal phenotypes to the subsequent generation[27,28].

tRNA-derived small RNAs (tsRNAs)[29], also known as tRNA-derived fragments (tRFs)[30–32], constitute another class of sncRNAs implicated in the trans-generational epigenetic inheritance in rodent model systems. One of the earliest discovered classes of tRFs were the tiRNAs (tiRs), which are typically 31–40 nucleotides long. The other tRF classes constitute the relatively shorter subtypes, that are 14–30 nucleotides long and map to the 5' and 3' ends of the mature or primary tRNA transcripts. tRF-5 and tRF-3 are generated from 5' and 3' ends of the mature tRNAs, while tRF-1 is generated from the 3' end of the primary tRNA transcripts[33]. Depending upon their length, tRF-5 and tRF-3 can be further divided into a, b or c sub-classes[33,34]. This heterogeneity in the tRNA derivatives provides a sense of complexity in gene regulation mediated by these small RNA molecules.

Chen et al.[35], using a rodent model system, reported that feeding a high fat diet (HFD; 60% fat) to male mice altered the expression of their sperm tsRNAs. Injection of these sperm-derived tsRNAs from the HFD-fed males into normal zygotes generated metabolic disorders in the F1 offspring. Similarly, Sharma et al.[36], also showed the effects of diet (low protein) on the levels of miRNAs and tsRNAs in mature mouse sperm. In a recent study by Natt et al.[37], the short-term effects of a 2-step diet intervention on human sperm was demonstrated. Their study reported that human sperm is highly sensitive to a sugar-rich diet, which affected sperm motility and biogenesis of tsRNAs, suggesting an interplay between nutrition and male reproductive health.

In addition to miRNAs and tRFs, piRNAs also constitute an abundant class of sncRNAs in the male germline. piRNAs range from 25–33 nucleotides in size, with 5' sequences enriched in uridine and 2'O-methyl modifications at the 3'end[38]. The primary function of piRNAs is to silence the transposable elements to protect genome integrity[39,40], but they have also been reported to play a role in DNA methylation and mRNA stability. piRNAs can also respond to various environmental stressors, raising the possibility of existing mechanisms that directly signal the effects of environmental changes to piRNA biogenesis machinery[41]. However, there are currently no studies reporting the effects of diet on piRNA expression in human sperm.

As is evident from animal model systems, sperm borne sncRNA may provide an epigenetic route for inter-generational signaling of paternal exposures. However, little is known about the existence of similar or related mechanisms in humans. Recently, research has shown that a "Mediterranean diet" rich in vegetables and legumes, olive oil, and fish (a source of omega-3 fatty acids) and low in carbohydrate rich snacks was related to a 40% increase in the probability of pregnancy and improved fertility outcomes[42,43]. In another study, significant correlations were observed between dietary intake of omega-3 fatty acids and embryo morphology after IVF[44]. However, there have been no large, randomised intervention studies that have investigated the impact of optimizing the periconceptional diet on fertility, perinatal and child outcomes. In addition, the mechanisms by which nutritional interventions impact fertility, embryo quality, implantation and subsequently periconceptional and perinatal development remain unknown and under explored. Therefore, to better understand the link between diet and human fecundity, the PREPARE trial was designed[45]. The PREPARE trial aims to address the effect of a diet rich in olive oil (as spread and cooking oil), marine omega-3 fatty acids EPA (0.8 g daily) DHA (1.2 g daily) and vitamin D (10 μg daily) on semen and sperm (in males), intrauterine environment (in females)[46] and on the developing embryo[47]. In this study, we aimed to characterise the baseline sncRNA landscape of human sperm, and study the effects of a paternal diet, enriched in olive oil, vitamin D, and omega-3 fatty acids on sperm sncRNA expression.

Results

Sperm sncRNA profiling was performed on a subset (n = 17) of participants enrolled in the PREconception dietary suPplements in Assisted REproduction (PREPARE) study in the UK (n = 111 were recruited and 102 completed the study; see Methods for details)[45]. Sperm samples were collected twice during the study, i.e. before and after 6 weeks of dietary intervention (**Supplementary File S1**). The 17 participants selected for sncRNA analysis were those with sufficient RNA both pre- and post-intervention for small RNAseq analysis; 9 were participants in the control arm, and 8 were participants in the intervention arm of the trial. The intervention group used an olive oil-based spread and olive oil for cooking plus a fruit-based drink enriched in vitamin D and omega-3 fatty acids (EPA and DHA as a component of fish oil) for 6 weeks, while the control group received a sunflower oil-based spread and sunflower oil for cooking and the fruit-based drink without added vitamin D and omega-3 fatty acids (the drink contained the same amount of sunflower oil as there was fish oil in the intervention drink) for the same duration of time (Fig. 1). Thus, participants and researchers were blinded to participant allocation to intervention or control group.

We checked whether the subgroup of participants being studied for sncRNA expression were representative of the physiological and anthropometric characteristics of the full study population by comparing the traits of the sncRNA subgroup (n = 17) with those of the full cohort minus the subgroup (n = 85). We found no significant differences in age and BMI between these two groups. However, some semen and sperm quality measures such as semen volume, sperm concentration, percentage of sperm motility and percentage of sperm progressive motility were significantly higher in the sncRNA subgroup, probably representing that the quality of sperm RNA that is needed for sequencing depends on the quality of the sperm itself: the lower the sperm quality, the lower the yield of RNA. The sncRNA subgroup had enough RNA pre- and post-intervention and hence this would mean having better sperm/semen quality at both stages. The comparison of the traits amongst the full and sub population are presented as **Supplementary Figure S1**.

Participant characteristics and demographics of the subpopulation (n = 17; control = 9, intervention = 8) are presented in Table 1. To get an enhanced understanding of the study subpopulations, comparison of the anthropometric and physiological traits was carried out between the control and intervention groups at Pre-intervention stage. There were no significant differences observed between the two groups, including the semen and sperm quality measures (Fig. 2, **Supplementary Figure S2**).

Small RNA sequencing was performed on the total RNA extracted from the 17 sperm samples pre and post intervention. The number of sequenced reads per sample ranged from 10 to 21 M, and the sequence length ranged from 10 – 46 nucleotides. sncRNAseq reads were mapped to the human genome and annotated into different subtypes using the Genboree software and additional miRNA and tRF annotators detailed in the methods. The reads were mapped to the human genome using Genboree software (using the default > = 18nt cut-off). On an average 56–59% of the reads mapped to the genome, 18% mapped to rRNA and 22–25% were unmapped (**Supplementary Figure S3**).
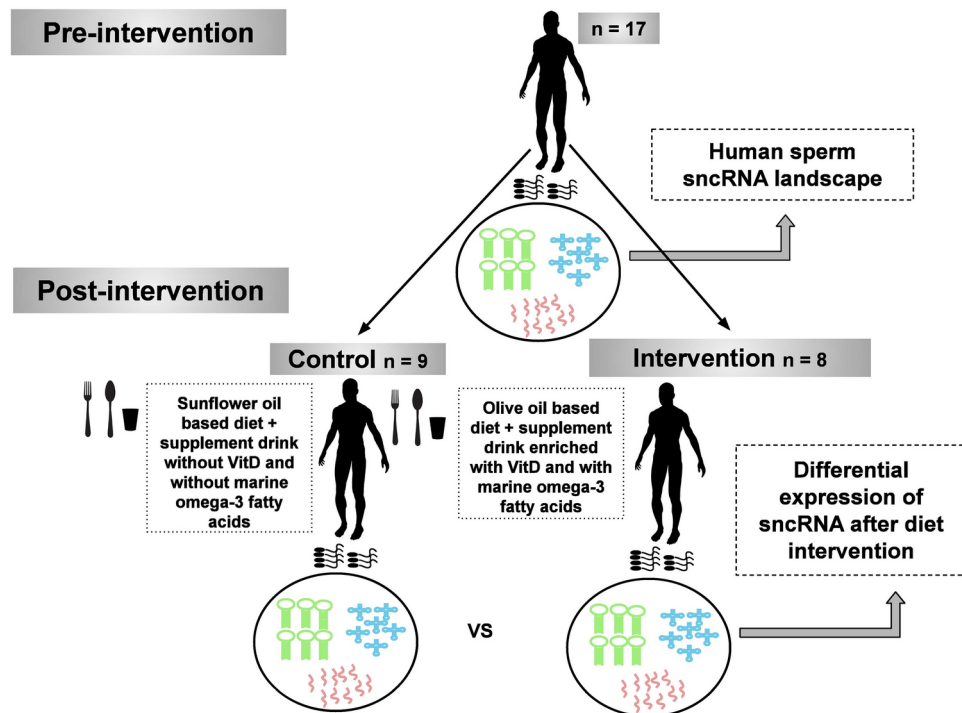
Fig. 1. Illustration of the Study design. Representation of the study design indicating the number of subjects and sncRNA analysis conducted in the pre- and post-intervention groups.

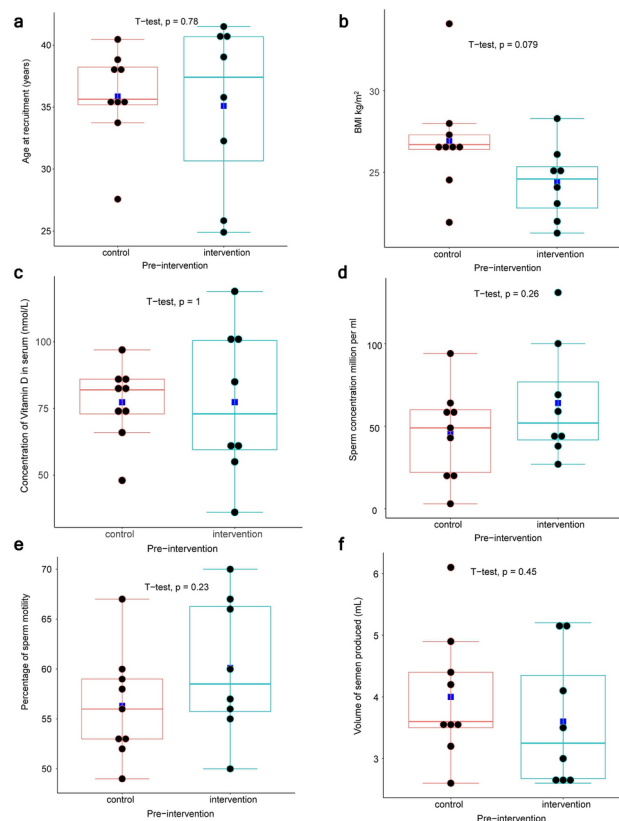|	Pre-intervention			Post-intervention		
	Control (n = 9) N/Average (SD)	Intervention (n = 8) N/Average (SD)	*P*-value	Control (n = 9) N/Average (SD)	Intervention (n = 8) N/Average (SD)	*P*-value
Characteristics						
Height (m)	1.81 (0.06)	1.81 (0.05)	9.57E-01	1.82 (0.07)	1.82 (0.05)	9.58E-01
Weight (kg)	88.71 (13.49)	81.51 (11.05)	2.45E-01	87.77 (14.41)	82.9 (11.76)	4.56E-01
BMI (kg/m^2)	26.93 (3.23)	24.39 (2.27)	7.93E-02	26.52 (3.94)	24.74 (2.60)	2.86E-01
Age (years)	35.87 (3.76)	35.09 (6.73)	7.79E-01	36.03 (3.74)	35.33 (6.70)	8.01E-01
Sperm:						
Sperm concentration (million/mL)	45.56 (27.77)	64.00 (35.22)	2.55E-01	52.33 (28.05)	65.88 (28.76)	3.43E-01
Sperm motility (%)	56.33 (5.39)	60.12 (6.92)	2.33E-01	51.78 (6.10)	59.88 (9.28)	5.80E-02
Sperm progressive motility (%)	45.11 (6.33)	50.5 (5.29)	7.54E-02	43.33 (6.89)	49.38 (11.31)	2.16E-01
Sperm with normal morphology (%)	11.25 (4.17)	8.62 (3.38)	1.89E-01	8.50 (4.93)	11.00 (5.88)	3.73E-01
Antisperm antibody IgG binding (%)	16.62 (29.83)	18.50 (21.75)	8.88E-01	19.17 (10.21)	15.00 (0.00)	3.63E-01
Average sperm damage (%)	30.91 (10.83)	29.21 (5.06)	7.58E-01	28.12 (8.02)	23.12 (11.64)	4.09E-01
EPA (%)	0.05 (0.02)	0.06 (0.03)	3.28E-01	0.06 (0.03)	0.08 (0.03)	1.18E-01
DHA (%)	23.68 (7.05)	24.23 (7.12)	8.75E-01	19.97 (8.96)	21.35 (7.16)	7.29E-01
Seminal fluid:						
EPA (%)	0.08 (0.04)	0.09 (0.07)	7.35E-01	0.14 (0.10)	0.08 (0.05)	1.68E-01
DHA (%)	10.12 (6.04)	10.24 (4.53)	9.64E-01	9.81 (6.09)	9.55 (5.71)	9.31E-01
Blood serum:						
Vitamin D2 (nmol/L)	0.00 (0.00)	1.50 (2.78)	1.71E-01	2.44 (2.92)	0.00 (0.00)	3.63E-02
Vitamin D3 (nmol/L)	77.33 (14.16)	75.88 (28.56)	8.99E-01	82.78 (21.57)	186.75 (36.17)	1.89E-05
Vitamin D total concentration (nmol/L)	77.33 (14.16)	77.38 (28.43)	9.97E-01	85.22 (20.12)	186.75 (36.17)	2.55E-05
RBC:						
EPA (%)	1.19 (0.74)	0.96 (0.30)	3.95E-01	1.06 (0.34)	2.52 (0.46)	5.74E-06
DHA (%)	5.07 (1.59)	4.38 (1.10)	3.09E-01	5.19 (1.41)	7.14 (1.03)	5.22E-03

Table 1. Participant characteristics and demographics, sperm analysis, semen analysis and blood measures for the sncRNA subset (n = 17) pre-intervention and post-intervention. BMI—Body mass index, EPA—Eicosapentaenoic acid (20:5n-3), DHA—Docosahexaenoic acid (22:6n-3). Simple t-test was performed for all continuous variables and fishers' test was performed for all categorical variables.

Pre-intervention sncRNA landscape of human sperm

sncRNA profiles of the 17 pre-intervention sperm samples were used to determine the baseline sncRNA landscape of the subjects in the study. tRFs were the most abundant sncRNA species (28%) detected in sperm, followed by miRNAs (5%) and piRNAs (1%). Small amounts of mitochondrial tRNAs and rRNAs were also detected (Fig. 3a, **Supplementary Figure S3**). High stringency analysis, based on the criterion of detecting sncRNA expression in all 17 subjects, identified a total of 143 tRFs, 401 miRNAs, and 2290 piRNAs (**Supplementary File S2**). Most of the sperm tRFs (63.6%, 91/143) originated from multiple genomic loci, while miRNAs (86.3%, 346/401) and piRNAs (79.7%, 1826/2290) were predominantly expressed from unique genomic locations (Fig. 3b-d). The sncRNA can arise from multiple genomic locations i.e. they have a tendency of multimapping on the genome. However, since we use a sequence centric approach, these multi-mapped sequences are counted only once per sequence.

Chromosome-based enrichment analysis revealed expression of a higher number of miRNAs from chromosomes 19 and X (Fig. 3b and **Supplementary Figure S4**). Likewise, a higher number of tRFs was observed from 4 chromosomes (5, 7, 15 and 16; Fig. 3c and **Supplementary Figure S5**), and piRNAs from 6 chromosomes (6, 9 15, 17, 19 and 22; Fig. 3d and **Supplementary Figure S6**). We also noted expression of miRNAs and piRNAs from the Y chromosome. Genomic annotation analysis revealed enrichment of the sperm sncRNA from the regulatory regions of the genome, and depletion from the coding exons. The majority of the miRNAs originated from the UTRs, while tRFs and piRNA were predominantly expressed from CpG islands of the sperm genome (Fig. 3e). There was also enrichment of sncRNA from the promoters (2 Kb upstream of the transcription start site) and 2 Kb downstream of the transcription termination sites. Introns showed a slight enrichment of miRNA expression but were generally depleted of tRFs and piRNAs.

Ninety-seven percent of the tRFs were derived from the 5′-end of tRNAs, including 5′-tiRs and 5′-tRFs (**Supplementary File S2**). In contrast, 3′-tiRs, 3′-tRFs, and 1'-tRFs were underrepresented in sperm. For each tRF id, the most abundant genomic sequence (having the highest read count) was extracted and compared across the pre-intervention samples, demonstrating sequence as well as length conservation across the samples. Around 80% of the most abundant tRF sequences were almost the same length (+ -2 nucleotides) across the samples. The most abundant sequence was taken as the representative sequence for a tRF (**Supplementary File S3**).

Fig. 2. Characterisation of the subpopulation (n = 17); comparison of traits amongst the intervention and control groups at pre-intervention. Box plots depicting the comparison of the anthropological and physiological traits amongst the intervention and control groups, based on a simple t-test analysis, for (**a**) Age (**b**) BMI (**c**) Concentration of Vitamin D in serum (**d**) Sperm concentration (**e**) Percentage of sperm motility and (**f**) Volume of semen produced.

422 out of 2290 (18.4%) baseline piRNAs were located within the transposable elements (LINE:106, SINE:66, LTR:195, Satellites:3, Others:53), and 17 out of 2290 (0.7%) originated from tRNAs (**Supplementary File S4**). Transposable elements (TEs) associated with piRNAs comprised of LINE elements such as L1, L2 and L3, SINE elements such as the Alu family, LTRs such as ERV1, ERVL, ERVL-MaLR, and Gypsy, DNA transposons such as Tigger3b|DNA|TcMar-Tigger and MER|DNA|hAT-Charlie, and Satellite elements (**Supplementary File S5**). 1041 of the remaining baseline piRNAs mapped to the genes, long intergenic non-coding RNAs (lincRNAs) and pseudogenes (**Supplementary File S5**). We found 234 genes related to the 1041 baseline piRNAs to be enriched in pathways such as Golgi and Endomembrane system organization (**Supplementary Table S1**).

Factors contributing to inter-individual variation in the sncRNA profiles

To investigate the factors contributing to the inter-individual variation in sperm sncRNA profiles of the subjects in the pre-intervention group, we studied the effects of age, BMI, and sperm fertility measures such as sperm concentration and sperm motility (Table 1). Sperm miRNA expression was significantly correlated with age, BMI, sperm concentration and sperm motility (p-value < 0.05). Expression levels of 36 miRNAs (13 positive and 23 negative) were correlated with age, 4 miRNAs (3 positive and 1 negative) with BMI, 171 miRNAs (89 positive and 82 negative) with sperm concentration and 10 miRNAs (7 positive and 3 negative) with sperm motility (**Supplementary File S6**). The top two most significantly correlated miRNAs associated with the four factors are shown in Fig. 4a: age (i) & (ii), BMI (iii) & (iv), sperm concentration (v) & (vi) and sperm motility (vii) & (viii). To ascertain the genes targeted by miRNAs correlated with age, experimentally validated and/or high scoring predicted targets of these miRNAs were identified. Pathway enrichment analysis of the target genes by IPA core analysis (details under Methods) identified the age dependent miRNA targets to be enriched in the pathways known to be linked with ageing, such as TGF-β signaling, mitochondrial dysfunction and oxidative phosphorylation (**Supplementary Figure S7**).

Sperm tRF expression was significantly correlated with age, BMI, sperm concentration and sperm motility (p-value < 0.05). Expression levels of 23 tRFs (10 positive and 13 negative) were correlated with age, 3 tRFs (2 positive and 1 negative) with BMI, 31 tRFs (10 positive and 21 negative) with sperm concentration and 3 tRFs (all 3 negative) with sperm motility (**Supplementary File S6**). The top two most significantly correlated tRFs associated with the four factors are shown in Fig. 4b: age (i) & (ii), BMI (iii) & (iv), sperm concentration (v) & (vi) and sperm motility (vii) & (viii).

a

b

c

d

e

Similarly, piRNA expression was significantly correlated with age, BMI, sperm concentration and sperm motility (p-value < 0.05). Expression levels of 103 piRNAs (48 positive and 55 negative) were correlated with age, 71 piRNAs (43 positive and 28 negative) with BMI, 233 piRNAs (108 positive and 125 negative) with sperm concentration, and 59 piRNAs (26 positive and 33 negative) with sperm motility. The top two most significantly correlated piRNAs associated with the four factors are shown in Fig. 4c: age (i) & (ii), BMI (iii) & (iv), sperm concentration (v) & (vi) and sperm motility (vii) & (viii).

In summary, sperm concentration had a more profound impact on the different sncRNA subtypes followed by age and BMI. This pre-intervention assessment of inter-individual variation in sncRNA profiles also helped identify the factors that could confound the post-intervention analysis. Sperm motility was not found to be that significant in our study as compared to sperm concentration, age, and BMI, but since it has been previously reported to affect sncRNA expression[48], we included it as an additional covariate in the subsequent analysis.

Effects of dietary intervention on vitamin D and omega-3 fatty acid levels

To study the effects of dietary intervention, we measured the levels of vitamin D and omega-3 fatty acids in blood samples from all male participants (n = 102) collected before and after 6 weeks of intervention. To also assess the sampling bias, these blood measures for the 17 subjects included in sncRNA analysis were compared with the full dataset. There was a significant increase in the serum vitamin D concentrations of the participants in the intervention group as compared to the control group for both the full set and sncRNA subset (p-value 5.71e-14 (full set) and p-value 2.55e-05 (sncRNA subset); Fig. 5a). Likewise, the dietary intervention also caused

◄**Fig. 3**. Baseline characterization of sncRNA landscape of human sperm. (**a**) Relative abundance of sncRNA subtypes identified during the baseline characterization of the sncRNA landscape of sperm samples collected at the pre-intervention visit. tRFs were the most abundant sncRNAs (mentioned as tRNA by Genboree tool), followed by miRNAs and piRNAs. Small amounts of mitochondrial tRNAs and rRNAs were also present. The misc_RNA included additional small RNA subtypes and fragmented mRNA. 'Other' includes small RNAs identified in this study, but not yet annotated in available sncRNA databases. (**b**) Phenogram showing chromosome-based enrichment analysis of 401 baseline miRNAs. Lighter shade of green represents the miRNA expressed from unique genomic locations, and the darker shade represents the miRNAs that mapped to multiple sites in the genome. *: chromosomes with over-represented miRNA expression (Z-score ≥ 2). #: chromosomes with under-represented miRNA expression (Z-score ≤ -2). (**c**) Phenogram showing chromosome-based enrichment analysis of 143 baseline tRFs. Lighter shade of blue represents the tRF expressed from unique genomic locations, and the darker shade represents the tRF that mapped to multiple sites in the genome. *: chromosomes with over-represented tRF expression (Z-score ≥ 2). #: chromosomes with under-represented tRF expression (Z-score ≤ -2). (**d**) Phenogram showing chromosome-based enrichment analysis of 2290 baseline piRNAs. Lighter shade of pink represents the piRNA expressed from unique genomic locations, and the darker shade represents the piRNA that mapped to multiple sites in the genome. *: chromosomes with over-represented piRNA expression (Z-score ≥ 2). #: chromosomes with under-represented piRNA expression (Z-score ≤ -2) (**e**) Genomic annotation enrichment of the sperm baseline sncRNA. The genomic regions analysed comprised of genes, introns exons, CpG islands, 5'UTRs, 3'UTRs, 2 Kb downstream of transcription termination site (Down) and 2 Kb upstream of transcription start site (Up). To statistically evaluate the associations between the baseline sncRNA and the genomic region sets, permutation tests were used. The z-scores generated indicate the enrichment/depletion of the baseline sncRNA in a particular genomic region and the * indicates p-value of < 0.01.

a significant increase in the percentage of EPA in red blood cells (p-value 6.01e-26 (full set) and p-value 5.74e-06 (sncRNA subset), Fig. 5b) and in the percentage of DHA in red blood cells (p-value 8.45e-19 (full set) and p-value 5.22e-03 (sncRNA subset); Fig. 5c) for both the full set and sncRNA subset.

These increases in the omega-3 fatty acids and vitamin D in the circulation provides evidence for the effectiveness of the 6-week dietary intervention. The lack of change in omega-3 fatty acids and vitamin D in the control group supports that the combination of sunflower oil and fruit-based drink does not impact the nutrients under study. We also tested if the intervention altered the EPA and DHA content of seminal fluid and sperm. There was no significant increase in the percentages of EPA and DHA in seminal fluid for both the full set and subset (Fig. 5d & 5e), The percentage of EPA measured in the sperm was significantly higher in the intervention group as compared to the control group for the full set (p-value 1.01e-02 (full set)), however owing to the small sample size increase in EPA was not significant for the subset, but the trend persisted (Fig. 5f). There was no significant increase in the percentage of DHA in the sperm for both full set and subset (Fig. 5g).

Hence, there are differences in EPA, DHA, and vitamin D levels in the circulation and in EPA levels in sperm that reflect changes in paternal diet; our data suggest that altered paternal nutrition might be captured and signaled to the male germline via an epigenetic route.

Effect of dietary intervention on sperm sncRNA expression

We next studied the effects of the dietary intervention on the sncRNA profile. Similar to the pre-intervention sncRNA composition of sperm, the post-intervention sncRNA profiles also showed a high abundance of tRFs (28%), followed by miRNAs (5%) and piRNAs (1%) (Fig. 5h, **Supplementary Figure S3**), indicating no global changes to the relative abundance of small RNA subtypes. However, we did identify differentially expressed sncRNAs between the control and intervention groups. **Supplementary File S7** provides the list of these differentially expressed sncRNAs detected after adjusting for the effects of age, BMI, sperm concentration and motility.

Differentially expressed miRNAs (DE-miRNAs)
mirDeep2 was used to characterise sperm miRNA expression profiles of the participants in the control and intervention groups and identify the DE-miRNAs. To study the specific effects of diet, we adjusted the analysis for the confounding effects of age, BMI, sperm motility and sperm concentration. Fifteen miRNAs were identified to be differentially expressed between the control and intervention groups, with log2(1.5) fold change and –log10 nominal p-value ≥ 2 (Table 2, Fig. 6a, and **Supplementary File S7**). DIANA miRPath identified 7 of these 15 miRNAs to target genes in fatty acid biosynthesis and metabolism pathways (Fig. 6b). For example, hsa-miR-506-3p had predicted target sites in 8 genes namely: *ACAA2, ACSL1, CPT1A, ELOVL5, HADH, OXSM, PECR* and *SCD*, followed by hsa-miR-513a-3p and hsa-miR-513c-3p targeting 5 genes: *ACADSB, ACOX1, ELOVL5, FASN, PTPLB* (Fig. 6b). Additional IPA analysis of the genes targeted by 7 out of 15 DE-miRNAs reconfirmed the enrichment of specific pathways related to fatty acid metabolism, such as fatty acid β-oxidation, mitochondrial – L carnitine shuttle pathway, stearate biosynthesis, fatty acid activation, γ-linolenate biosynthesis, palmitate biosynthesis, fatty acid biosynthesis initiation and oleate biosynthesis (Fig. 6c). Besides fatty acid metabolism pathways, the LXR/RXR activation, TGF-beta signaling, Hippo signaling and Wnt signaling pathways were also found to be significantly enriched (**Supplementary Table S2**).
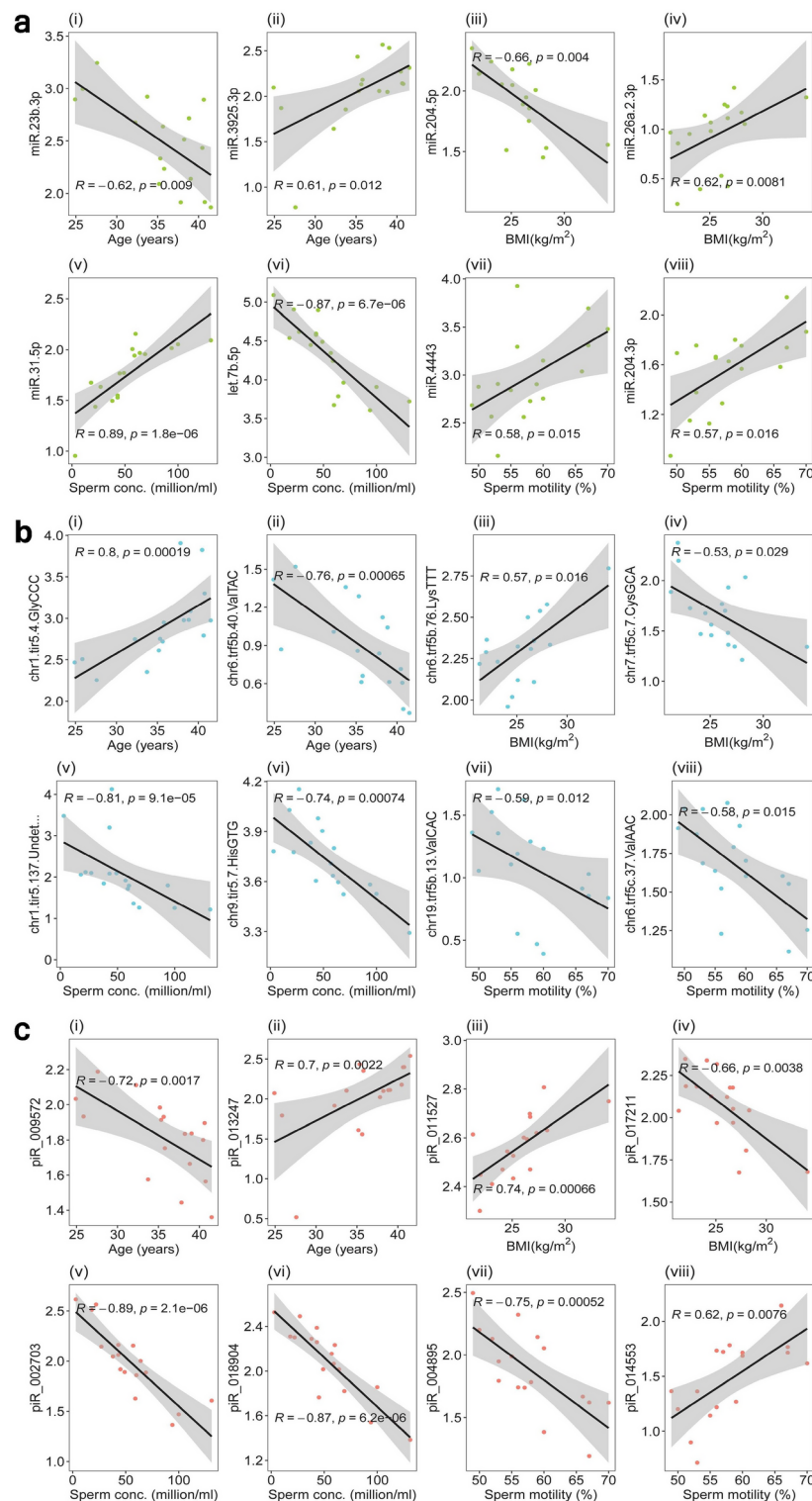
Fig. 4. sncRNA correlations with age, BMI, sperm concentration and sperm motility. (**a**) Correlation plots showing the top two most significantly correlated miRNAs associated with (i & ii) age, (iii & iv) BMI, (v & vi) sperm concentration and (vii & viii) sperm motility. (**b**) Correlation plots showing the top two most significantly correlated tRFs associated with (i & ii) age, (iii & iv) BMI, (v & vi) sperm concentration and (vii & viii) sperm motility. (**c**) Correlation plots showing the top two most significantly correlated piRNAs associated with (i & ii) age, (iii & iv) BMI, (v & vi) sperm concentration and (vii & viii) sperm motility. Y-axis represents the specific sncRNA expression in log CPM, while X-axis represents the factor.
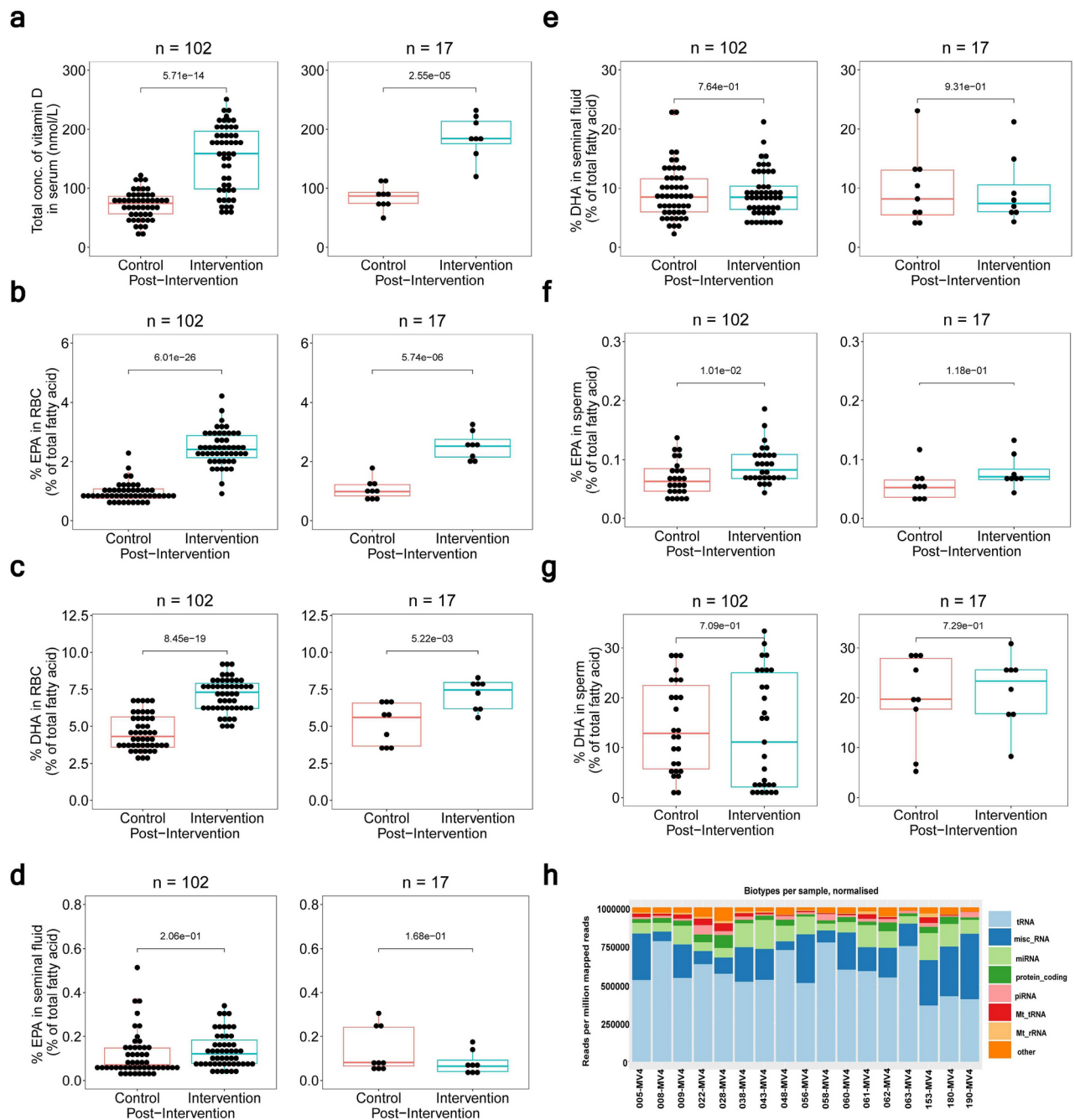
Fig. 5. Graphical depiction of the effect of diet intervention on vitamin D and omega-3 fatty acid levels. Boxplots depicting the comparison among the participants of the control and intervention groups post-intervention, based on a simple t-test analysis. (**a**) Total concentration of vitamin D in the serum (nmol/l) (**b**) Percentage EPA in RBCs (**c**) Percentage DHA in RBCs (**d**) Percentage EPA in seminal fluid (**e**) Percentage DHA in seminal fluid (**f**) Percentage EPA in sperm (**g**) Percentage DHA in sperm. The plots in the left panel represent the full set (n = 102) and the plots in the right panel represent the sncRNA subset (n = 17). (**h**) Relative abundance of sncRNA subtypes post-intervention.

Differentially expressed tRFs (DE-tRFs)

Expression profiling of tRNA derived fragments (tRFs) was done using the analysis pipeline developed by Kumar et al.[34]. Quantitative expression of 8 different categories of tRNA derived tRFs (5'tir, tRF-5 (a/b/c), tRF-3(a/b/c) and tRNA trailers (tRF-1)) was performed. After adjusting for age, BMI, sperm motility and sperm concentration, 3 tRFs were found to be differentially expressed between the intervention and control groups (log2(1.5) fold change and –log10 nominal p-value ≥ 2; Table 3, Fig. 7a, **Supplementary File S7**). These three differentially expressed tRFs included derived trf5b-2-TyrGTA, tir5-29-CysGCA and trf5b-24-AlaAGC.

miRNAs	logFC	Nominal P value	FDR-corrected P value
miR-513c-3p	2.54	4.25E-04	0.50
miR-4774-3p	-6.39	1.25E-03	0.51
miR-509–3-5p	1.65	1.37E-03	0.51
miR-1229-3p	-4.57	2.51E-03	0.51
miR-656-5p	-3.74	2.88E-03	0.51
miR-4760-5p	-1.91	2.97E-03	0.51
miR-506-3p	2.40	3.02E-03	0.51
miR-503-5p	2.34	3.74E-03	0.55
miR-6511b-5p	-2.11	4.76E-03	0.60
miR-513a-3p	2.08	5.08E-03	0.60
miR-136-3p	2.05	5.99E-03	0.64
miR-508-3p	1.63	7.01E-03	0.65
miR-6503-3p	-4.03	7.18E-03	0.65
miR-4507	-1.88	8.79E-03	0.72
miR-3147	-3.03	9.88E-03	0.72

Table 2. Differentially expressed miRNAs following the dietary intervention. These DE miRNAs were identified after adjusting for age, BMI, sperm concentration and sperm motility using log2(1.5) fold change cut-off and a -log10 nominal p-value ≥ 2.

Differentially expressed piRNAs (DE-piRNAs)
Differential expression analysis of piRNAs adjusted for age, BMI, sperm motility and sperm concentration identified 112 piRNAs (log2(1.5) fold change and –log10 nominal p-value ≥ 2) to differ in expression between the intervention and control groups (Table 4, Fig. 7b and **Supplementary File S7**). 10.7% (12/112) of these DE-piRNAs were located within the transposable elements (LINE:4, SINE:2, LTR:4, Others:2), and another 48.2% (54/112) overlapped with genomic regions encoding genes or LINC RNA (**Supplementary Files S8 and S9**).

Discussion

Small non-coding RNA of sperm contribute to diverse functions such as: development of germ cells[49], regulation of fertility and transmission of epigenetic information. Both piRNAs and miRNA play an important role during spermatogenesis, and hence any disruptions in their biogenesis pathway can lead to male infertility, or influence embryo development[50]. The epididymosomes secreted by the epididymal epithelial cells contribute to sperm maturation as well as transferring environmental effects to the sperm through the ncRNA components. As the sperm transits from testis to epididymis tsRNAs (tRFs) become the dominant sncRNA and ultimately sperm acquires a sncRNA profile that serves as "sperm RNA code"[51]. This sperm RNA code is highly sensitive to environmental and lifestyle factors and has the potential to transmit the acquired effects to the next generation. A study on men living with obesity revealed increased expression of miR-155 and miR-122, that were previously associated with inflammation and iron metabolism, respectively, at both the systemic and sperm level[52]. Besides transmitting adverse phenotypes, the sperm borne sncRNA can also transmit paternally acquired cognitive benefit to the offspring[53]. Endurance training was found to acutely affect the expression of eight piRNAs, which regressed after the cessation of training[53].

Our study provides a high-resolution mapping of the three most abundant sncRNA species (miRNAs, tRFs and piRNAs) in human sperm, and identifies the effects of 6 weeks of healthy dietary intervention on their expression. In the intervention arm of the study, the participants were given a diet enriched in olive oil, vitamin D and omega-3 fatty acids, known for their anti-inflammatory and beneficial roles in metabolism[54].

Spermatogenesis takes an average of 74 days; however, this cycle is continuous, with millions of sperm cells produced daily. Given our 6-week intervention period (42 days), we can expect a mixture of sperm cells with varying degrees of dietary exposure. Even extending the intervention to 74 days or longer would not ensure a homogeneous population of sperm with consistent exposure, as mature sperm cells are either ejaculated or reabsorbed by the body, leading to the production of new sperm cells. Additionally, factors such as ejaculation frequency and abstinence duration would affect the composition of the sperm population.

A longer duration of exposure could potentially have led to stronger signals, however the PREPARE trial was conducted on couples undergoing IVF, hoping to achieve conception at the earliest opportunity. The intervention duration was designed to be long enough to influence gamete maturation but not so long as to be burdensome or delay treatment, which might deter participation and compliance to the interventions. This duration is supported by evidence indicating that certain lipid pools could reach maximal or near-maximal changes in EPA and DHA content within 6 weeks[55]. Additionally, rodent studies show that dietary manipulations during a very short window around implantation can significantly impact early development[56]. Moreover, previous studies with shorter dietary intervention durations than ours have shown to have a significant impact on sperm. For example, Natt et al.[37] demonstrated the acute effects of a 2-step diet intervention on human sperm involving 1 week of a healthy diet to establish a baseline, followed by 1 week of additional sugar intake. They observed a dramatic increase in sperm motility and a concurrent rise in tsRNA levels in all individuals. The study concluded
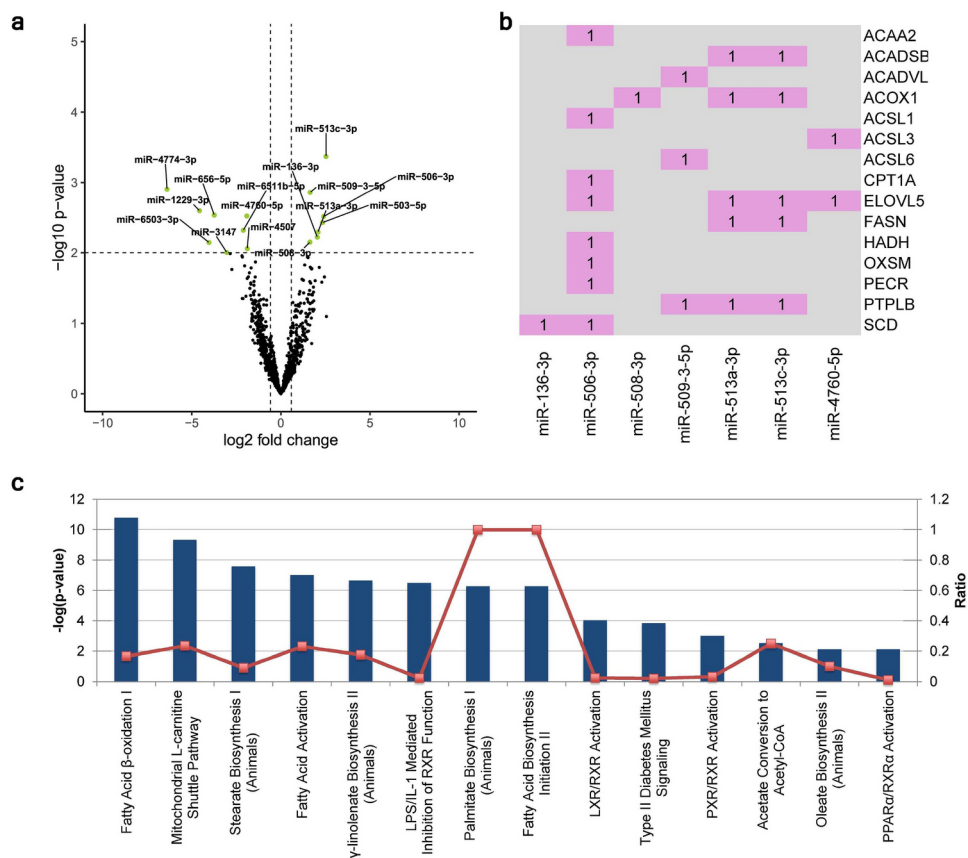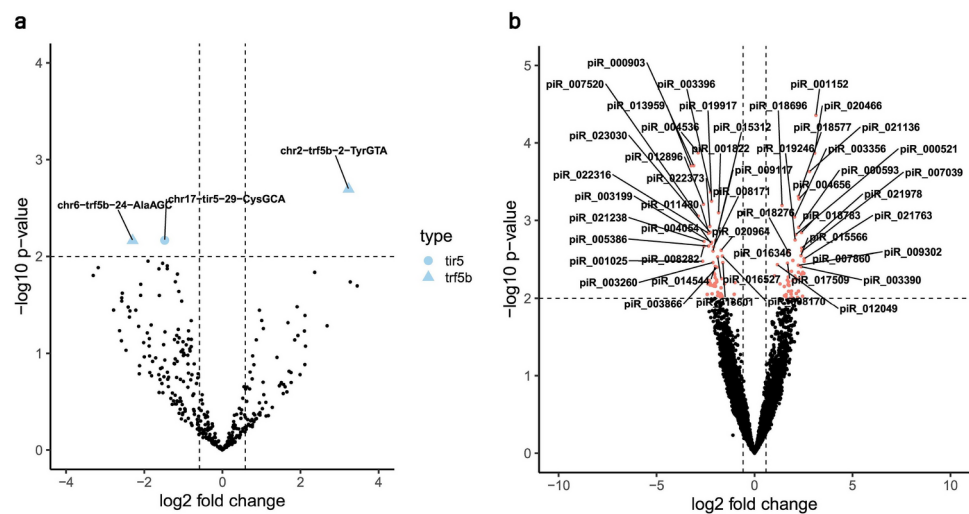
Fig. 6. Effect of the dietary intervention on miRNA expression. (**a**) Volcano plot showing differentially expressed miRNAs (intervention vs control) identified after adjusting for age, BMI, sperm concentration and motility. Horizontal dashed line represents the -log10 nominal p-value = 2 significance thresholds of association, and the vertical dashed lines represent the log2(1.5) fold change (**b**) Differentially expressed miRNAs and their corresponding target genes predicted by DIANA miRPath. Purple color (1) indicates the presence of the miRNA-target interaction, while grey color indicates the absence of a predicted interaction. (**c**) Ingenuity Pathway Analysis (IPA) of genes targeted by differentially expressed miRNA. Pathways with the most overrepresented genes and an enrichment -log10 FDR-corrected p-value > 2 are plotted by the decreasing order of significance. The dotted red line represents the ratio of the miRNA target genes to the total number of genes known to be present in that pathway.

tRFs	General tRF name	logFC	Nominal P value	FDR-corrected P value
chr2-trf5b-2-TyrGTA	trf5b-TyrGTA	3.23	2.02E-03	0.44
chr17-tir5-29-CysGCA	tir5-CysGCA	-1.48	6.84E-03	0.44
chr6-trf5b-24-AlaAGC	trf5b-AlaAGC	-2.30	6.87E-03	0.44

Table 3. Differentially expressed tRFs following the dietary intervention. These DE tRFs were identified after adjusting for age, BMI, sperm concentration and sperm motility using log2(1.5) fold change cut-off and a -log10 nominal p-value ≥ 2.

that sperm motility and the sncRNA population in human sperm are highly sensitive to nutritional flux, exhibiting a rapid and specific response to dietary modifications. Despite the relatively short 6-week duration of the intervention, there was a significant increase in the levels of vitamin D, EPA and DHA in the circulation, and a change in the sperm sncRNA expression profile. This study also generated a comprehensive baseline sncRNA landscape of human sperm and identified the small RNAs associated with age, BMI, and sperm concentration.

Baseline characterisation of the sncRNA landscape in human sperm revealed multiple subtypes of small RNAs, with varying abundance. These sncRNA were predominantly expressed from the regulatory regions of the genome, such as the UTRs, CpG islands and promoters suggesting their potential role in chromatin remodeling[57] and transcriptional regulation in sperm. The expression levels of the tRFs were relatively high as compared to miRNAs and piRNAs, making them the most abundant sncRNA expressed in male gametes.

Fig. 7. Effect of dietary intervention on tRF and piRNA expression (**a**) Volcano plot showing differentially expressed tRFs (intervention vs. control) identified after adjusting for age, BMI, sperm concentration and motility. (**b**) Volcano plot showing the differentially expressed piRNAs (intervention vs. control). Horizontal dashed line represents the -log10 nominal p-value = 2 significance thresholds of association, and the vertical dashed lines represent the log2(1.5) fold change.

piRNAs	logFC	Nominal *P* value	FDR-corrected *P* value
piR_001152	3.13	4.40E-05	0.33
piR_004536	-2.88	1.34E-04	0.33
piR_020466	3.07	1.36E-04	0.33
piR_000903	-3.11	1.96E-04	0.33
piR_013959	-3.24	1.96E-04	0.33
piR_021136	2.81	2.33E-04	0.33
piR_003396	-2.25	4.29E-04	0.45
piR_018577	2.25	4.87E-04	0.45
piR_004656	2.27	5.27E-04	0.45
piR_019917	-2.19	5.62E-04	0.45
piR_012896	-2.61	6.18E-04	0.45
piR_018696	1.40	6.39E-04	0.45
piR_001822	-1.83	7.94E-04	0.51
piR_007520	-2.88	8.79E-04	0.51
piR_003356	2.04	9.04E-04	0.51

Table 4. Differentially expressed piRNAs following dietary intervention. These DE piRNAs were identified after adjusting for age, BMI, sperm concentration and sperm motility using log2(1.5) fold change cut-off and a -log10 nominal p-value ≥ 2. Only the top 15 most significant DE piRNA have been listed out in the table.

tRF5b and tRF5c were the dominant tRNA derivatives, that have been previously reported to be involved in epigenetic inheritance[33]. We also detected tRF3a, b and c forms that not only have lengths comparable to mature miRNAs but are also known to function like miRNAs and follow similar target binding rules. The most highly expressed tRFs included Gly-GCC, Glu-CTC, Glu-TTC, Val-CAC and His-GTG, which were also reported to be the most abundant tRFs in mouse model systems[35,36]. tRF-Gly-GCC has been shown to repress expression of genes regulated by endogenous retroelement MERVL in early embryo development[36]. Among the differentially expressed tRFs, Cys-GCA was previously reported to be downregulated in a high-fat diet mouse model[35].

The effectiveness of the 6 weeks dietary intervention was evident by a significant increase in the vitamin D levels of serum, and the EPA and DHA percentage in RBCs of the participants in the intervention group. 7 out of 15 differentially expressed miRNAs were found to target genes enriched in fatty acid biosynthesis and metabolism, demonstrating the potential of the dietary intervention to generate a physiological change sufficient to be signaled to the male germline and be epigenetically captured by alterations to sncRNA.

During spermatogenesis in mammalian testis, piRNA is the most dominant species among the sncRNA population and on the basis of the stages in which it appears it can be classified into three categories: the first set of piRNAs (26-28nt) arising in the fetal testis brings about transposon silencing via de novo DNA methylation[58];

the second set of piRNAs (26-27nt) are called "pre-pachytene piRNAs" that originate from the 3'UTR of mRNA and have unknown functions[59]; and the third set of piRNAs (26-30nt) are called "pachytene" piRNAs since they arise in the pachytene phase of meosis[60]. Pachytene piRNAs have diverse roles: (a) Regulation of meotic progression (b) Elimination of bulk mRNA in spermatids (c) Activation of the translation of spermiogenic mRNAs. Pachytene piRNAs play a significant role in regulation of spermatogenesis.

In our study we found around 18.4% (422/2290) of baseline piRNAs and 10.7% (12/112) of diet induced piRNAs are expressed from transposable elements (TE) such as LINES, SINES and LTRs. Since one of the functions of piRNAs is to protect genome integrity during spermatogenesis by silencing transposable elements[61], some of them were found to originate directly from TE transcripts in the sense orientation.

piRNAs are also known to have non-TE related functions, such as regulation of mRNA and lincRNA expression through post-transcriptional gene silencing[62]. We found genes related to baseline piRNAs to be enriched in pathways such as Golgi and Endomembrane system organization. Strikingly, we also found 17 baseline piRNAs to originate from tRNAs, suggesting the existence of a regulatory crosstalk between the two.

There have been several reports characterizing the expression of tRNA derived ncRNAs, which suggests the concept that tRNAs are utilized to produce functional ncRNAs[63,64]. There is still a dearth of information regarding the need to utilize tRNAs for ncRNA production. The functional convergence on sncRNA expressed during spermatogenesis, in gametes and zygotes was demonstrated by Garcia-Lopez et al.[65]. They explored a set of piRNAs that mapped to rRNAs and tRNAs suggesting that these piRNAs could arise as secondary products of tRNA and rRNA processing machineries implying a cross-talk between translational and interfering pathways[66].

This study has some limitations. This trial was conducted on a specific population of individuals undergoing IVF, to enable a pre-fertilization intervention to study the effect on embryo quality[47]. Therefore, its generalizability to men outside of the IVF settings is uncertain. Nevertheless, the insights obtained from the study are valuable in progressing our knowledge regarding the impact of dietary intervention in the periconception phase and may be applicable to the general population as well. Statistical analysis involving comparison of traits between control and intervention groups for both pre- and post-intervention was performed using both Wilcoxon and Student's t-test. We found no significant difference in the outcomes; this could be attributed to the small sample size. To avoid small sample size errors and detection of false positives, we restricted the baseline analysis to the high confidence sncRNA detected in 17 participants from the original 102 in order that high quality RNA could be studied. This high stringency analysis may have limited the sensitivity of detection and missed some sncRNA. Also, due to the limited functional characterization of the downstream targets of human tRFs, we were unable to reflect on the regulatory outcomes of their altered expression following dietary intervention. However, the high-resolution tRF and piRNA mapping attempted in this study, revealed the not yet known complexity in the expression of their subtypes in human sperm. Expression of a subset of piRNAs from the tRNA locus also identified the potential crosstalk between different sncRNA, that might be essential to effectively capture the effects of external exposures in the male germline.

In conclusion, the short-term nutritional intervention with olive oil, vitamin D and omega-3 fatty acids altered expression of 3 tRFs, 15 miRNAs and 112 piRNAs, targeting genes involved in fatty acid metabolism and transposable elements in the sperm genome. As per the WHO 2021 parameters only 2 out of 17 men were sub-fertile; however, there was variation among the sperm concentration and sperm motility measures for all the 17 men. This variation could lead to detection of sperm fertility parameters related sncRNA instead of diet. Therefore, along with age and BMI, these were also adjusted for in the differential expression model to ensure that the differentially expressed sncRNA were not related to sperm concentration and motility.

The epigenetic programing of male gametes is sensitive to alterations in lifestyle, such as diet[6-8]. Effects of short-term dietary exposures can be captured as "environmental messages" in the form of altered sncRNA expression profiles in the male germline and be likely transmitted to the next generation if these changes occur around conception. Our data mirror experimental studies in rodents showing paternal nutritional exposures can be captured and signaled to the male germline via an epigenetic route, with implications for the health of the next generation[22]. A systematic review and meta-analysis were conducted to determine the effects of healthy dietary patterns on semen quality. These included studies examining the influence of the Mediterranean, Dietary Approaches to Stop Hypertension (DASH), and prudent diet patterns on semen quality parameters. By comparing high versus low consumption of healthy diets, the analysis revealed that men with higher consumption had significantly better sperm concentration, progressive sperm motility, and total sperm count compared to those with lower consumption[67].

In our study, we explored the effect of diet beyond fertility parameters and discovered several diet-altered sncRNAs. The altered miRNAs were found to target genes enriched in fatty acid biosynthesis and metabolism pathways. Fatty acids play multiple biological roles in sperm, from being a cell membrane component to supplying energy and functioning as signaling molecules. An optimal fatty acid profile, which can be enhanced by dietary fatty acids, is essential for normal spermatogenesis[68]. Thus, the intervention drink, enriched in omega fatty acids, led to changes not only in the %EPA and %DHA in RBCs but also in the %EPA in sperm, further influencing sperm sncRNAs.

Methods
Sample information and study design
The PREPARE trial is a single center, randomized controlled trial, in which 111 men and their partners about to undergo a cycle of IVF treatment provided written informed consent to be recruited to the study[45]. Ethical approval for this study was granted by an NHS Research Ethics Committee (13/SC/0544). All methods were performed in accordance with relevant guidelines and regulations. Exclusion criteria included eating oily fish (as defined by the UK Food Standards Agency) more than once a week, previously diagnosed diabetes, and any medical contraindication to the dietary components. Demographic data from the men and information on their

diet and physical activity were collected on a standard study proforma at the time of recruitment and their BMI was determined. The mean age of the participants was 36 ($\pm$ 5.5) years and mean BMI was 27.0 ($\pm$ 4.0) kg/m^2 (**Supplementary File S1**).

Study participants were randomized into the control and intervention diet groups. The intervention group were given a daily fruit-based dietary supplement drink containing 2 g DHA and EPA (1.2 g DHA and 0.8 g EPA) and 10 µg vitamin D (Nutrifriend 2000, Smartfish, Oslo, Norway) and olive oil and olive oil spread to use as part of their diet, while the control group received a placebo drink (the same fruit-based drink but without omega-3 fatty acids or vitamin D (Smartfish, Oslo, Norway)) and sunflower oil and sunflower oil based spread to use as part of their diet. The oils and spreads were to be used in place of the oils and spreads used habitually. The nutritional drink is a commercially available product called REMUNE from SMARTFISH company (OSLO, NORWAY). It contains (per 200 ml): 10 g whey protein, 11 g fat including omega-3 fatty acids ($\sim$ 1,200 mg DHA + $\sim$ 800 mg EPA) from fish oil, 20 g carbohydrate and 10 µg 25-hydroxy-vitamin D3. It also contains vitamins, minerals, and polyphenols from the fruit. The control drink was produced for the study by Smartfish and was identical except that it contained sunflower oil instead of fish oil and no vitamin D. So, intake of everything from the drinks was identical in both the groups apart from omega-3 FAs and vitamin D.

The Olive oil used was a supermarket-bought extra virgin olive oil that contained 65% of fatty acids as oleic acid and had an undetermined content of polyphenols and other naturally occurring ingredients.The olive oil, sunflower oil and both spreads were purchased from a local supermarket and were repackaged for use in the study. The drinks, oils and spreads were all provided in unmarked containers and the trial was double blind. Following inclusion into the study but prior to randomisation, consenting participants were invited to complete a preconception questionnaire and the short Southampton Food Frequency Questionnaire (FFQ) to characterise their lifestyle and diet prior to entry into the study. None of the participants were taking any prescribed medications, with the exception of PRN salbutamol inhalers for asthma. They were also excluded from participation if they were taking herbal remedies. It is therefore presumed that all male participants were generally fit and well.

At study entry, across the entire cohort of participants at study entry (n = 111 men), 7% of men had a total serum 25-hydroxy-vitamin D concentration < 25 nmol/L, indicative of deficiency. A further 20% of men had borderline deficiency (25–50 nmol/L). As expected, the season of recruitment affected the vitamin D concentration measured; men who were recruited in the summer (77.75 + 25.01 nmol/L) had higher mean total concentrations than those recruited in the winter (49.07 + 22.51 nmol/L; P < 0.001). There is no definition of omega-3 deficiency or sufficiency, but mean omega-3 index (EPA + DHA in red blood cells) at baseline in men was approx. 5.5 which is considered "intermediate" with 8 or more being desirable. None of the men had a desirable omega-3 index at study entry. The duration of the intervention was 6 weeks. 102 male participants completed the trial and had semen samples collected at both pre- and post-intervention visits. sncRNAseq profiling was performed on a subset of these participants (n = 17; 9 from the control group and 8 from the intervention group) (Fig. 1).

Semen analysis

The male participants were asked to abstain from sex or masturbation for three to five days prior to giving the semen sample. Semen analysis was performed by a blinded andrologist on all samples in accordance with the WHO guidelines[69]. Sample liquefaction and viscosity; the volume of semen produced; sperm concentration; percentage of motile and progressively motile sperm; and percentage of sperm with normal morphology were recorded. The sample was considered normal if at least 1.5 ml of semen was produced; it liquefied within one hour and was not viscous; and it had a least 15 million spermatozoa per ml of ejaculate with a progressive motility of at least 32% and a normal morphology of at least 4%.

Semen (2 ml or the maximum volume available if less) was placed on a density gradient (created using 1 ml of Nidacon PureSperm®100 90% solution overlaid with a 1 ml of Nicadon PureSperm®100 45% solution) and centrifuged at 400 g for 20 min to separate the seminal plasma from the sperm. The supernatant (seminal plasma) was removed and stored in 1 ml aliquots at -80 °C. The sperm pellet was washed with phosphate-buffered saline (PBS) and centrifuged at 400 g for a further 5 min, this sperm pellet was collected and stored at -80 °C prior to analysis. For DNA fragmentation analysis, semen was left to liquefy for at least 30 min following ejaculation. Between 100 µl and 200 µl was then placed into a tube and plunged into liquid nitrogen to snap freeze and then stored at -80 °C until analysis. Sperm DNA fragmentation was assessed using the SpermComet assay, a second-generation sperm DNA test[70]. This assay required approximately 5000 sperm and allowed the percentage of DNA damage in each sperm to be ascertained. It is a single-cell gel electrophoretic assay and detects global sperm DNA fragmentation i.e.both single and double stranded DNA breaks.

Blood sampling

Blood was collected into lithium heparin and serum separating tubes at the time of recruitment and on the day of partner's oocyte retrieval. Blood was collected in a non-fasting state. Blood collected into lithium heparin was centrifuged at 1000 g for 15 min at 20 °C. The plasma was removed, and the red cell pellet was washed in PBS and centrifuged at 400 g for 10 min at room temperature. The wash step was repeated twice. The red cell pellet was then stored as a 0.5 ml aliquot at -80 °C prior to fatty acid analysis. Blood collected into SST was mixed and centrifuged at 1500 g for 10 min. The supernatant (serum) was removed and stored in aliquots at -80 °C prior to vitamin D analysis.

Fatty acid analysis

Fatty acid analysis was performed on red blood cells, seminal plasma and sperm. In all cases, total lipid was extracted into chloroform: methanol (2:1, vol/vol). The lipid extract was heated to 50 °C for 2 h with 2%

methanol in sulphuric acid to produce fatty acid methyl esters (FAMEs). FAMEs were separated and identified by gas chromatography performed according to conditions described elsewhere[71]. FAMEs were identified by comparison of run times with those of authentic standards. Fatty acid concentrations are expressed as % of the total fatty acids present. Analysis on the sperm was only performed if the sperm concentration was greater than 15 million per ml and there was enough semen remaining following clinical use on the day of oocyte retrieval.

Vitamin D analysis
Serum vitamin D concentrations were determined by Liquid Chromatography/Tandem Mass Spectrometry (Waters, Milford, MA, USA) in the Chemical Pathology Laboratory, University Hospital Southampton NHS Foundation Trust, a member of the Vitamin D EQA scheme (DEQAS).

Semen sample processing and RNA extraction
Total RNA including sncRNA was extracted from frozen sperm pellets using the RNeasy Micro Kit (Qiagen, UK) with a modified protocol adapted from Goodrich, et al.[72], which was optimised for recovery of sperm RNA. Briefly, sperm pellets were thawed, mixed with Qiazol (Qiagen, UK) and manually disrupted. Chloroform was added, samples incubated (room temperature for 5 min) and centrifuged ($12,000 \times g$) for 20 min. The upper aqueous layer was removed and mixed with ethanol. Samples were added to micro RNeasy spin columns (Qiagen, UK) and centrifuged for 15 s ($8000 \times g$). Spin columns were washed using supplied buffers and ethanol. Samples were transferred to RNAse free collection tubes and total RNA eluted. RNA concentrations were quantified using Qubit total and microRNA assays (Thermo Fisher, UK), RNA was snap frozen on dry ice and stored at -80 degrees.

sncRNA sequencing
Sperm RNA samples were selected from both control and dietary intervention groups based on the quantity of RNA post extraction (n = 17). 10 µl of RNA was used to perform small RNA-seq (~ 15 million reads) using the smallRNA workflow (Oxford Genomics, OGC). NEBNext Small RNA Library Prep Set for Illumina (E7330) was used as per manufacturers' instructions. Briefly, 3' SR adapter ligation was carried out followed by RT primer hybridisation to the mRNA. Denatured 5' SR adapter was then ligated, and first strand cDNA reverse transcription was performed. After second strand synthesis with indexed (TruSeq-based single indexed i7) and P5 primers, dsDNA was then cleaned-up using Ampure XP beads, and the libraries were size selected by Pippin (125-160 bp). The concentration of the small RNA libraries was determined by Picogreen before multiplexing for sequencing. The final size distribution of the pooled library was determined using Agilent Tapestation, and quantified by Qubit (Thermofisher, UK), before sequencing on the HiSeq2500 (Illumina, UK) (2X runs) as 50 bp single reads. Data were aligned to the reference genome and FASTQ files generated.

Processing of sncRNA sequencing data
FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to perform quality check (QC) on the sequencing reads. All samples passed the mean per base sequence quality scores ≥ 30. Average length of sequenced reads was 51 nt, inclusive of the adaptors. Trimmomatic 0.38[73] was used to remove the adapter sequences, and sequence reads < 10 nt were excluded from the subsequent analysis. exceRpt small RNA-seq analysis pipeline available under the Genboree workbench (https://www.genboree.org/site/) was used to profile and assess the relative abundance of sncRNA subtypes. Since the current version of Genboree does not include the latest release for miRNA and tRF databases, we additionally used specific annotator tools to report the final results. Mapper.pl and quantifier.pl modules of the miRDeep2 tool[74] and the latest miRBase (http://mirbase.org/) release 22 were used to annotate and quantify the known human miRNA.

Following parameters were used for the mapper and quantifier module taking sample 005mv1-249 as an example:

mapper.pl SRR12858025-trim-orig.fastq -e -g S25 -h -i -j -l 15 -m -s SRR12858025-trim-orig-col.fa -v

SRR12858025-trim-orig.fastq is the input fastq file and SRR12858025-trim-orig-col.fa is the output file with the collapsed reads

quantifier.pl -p hsa-release-22.1-prec-miRNA.fa -m hsa-release-22.1-mat-miRNA.fa -r SRR12858025-trim-orig-col.fa -t hsa

Input file for mapping: SRR12858025-trim-orig-col.fa (collapsed file output of mapper module)

For tRF analysis, human tRNA sequences were downloaded from the tRNA database (http://gtrnadb.ucsc.edu/). For each tRNA gene, the DNA sequences extracted from the genome included additional 100 bases upstream and 200 bases downstream of the mature tRNA. Since 'CCA' is added to the 3' end of the tRNA, by tRNA nucleotidyltransferase during the tRNA maturation process, this triplet nucleotide sequence was also included at the 3' end of the extracted DNA sequences to avoid sequence mismatch during the mapping process. The sperm sncRNA sequencing reads were then mapped to the extracted tRNA sequence by using the BLASTn tool. Sequence alignments with ≤ 1 nt mismatch were retained for subsequent analysis. The blast output file was parsed to get information on the mapped position of small RNA on the tRNA genes and segregate the sequences into specific subtypes: 5' tiRs, 5' (a,b,c), 3' (a,b,c) and 1-tRFS.

For piRNA expression analysis, Genboree platform was used to quantify and profile the human piRNAs from the piRNABank.

Statistical analysis involving comparison of traits
All statistical analyses were conducted using R version 4.0.3. Comparisons of traits between control and intervention groups for both the pre- (Fig. 2) and post-intervention (Fig. 5), were performed using Student's t-tests. This includes analysis of participant characteristics, demographics, semen analysis parameters, fatty

acid levels, and vitamin D measurements. A p-value of < 0.05 was considered statistically significant for all comparisons. We also carried out the statistical analysis using Wilcoxon test and found no significant difference in the outcomes, this could be attributed to the small sample size.

Over and under—representation analysis

Figures 3b-d have some chromosomes marked with * (over-represented), # (under-represented), this indicates the chromosomal enrichment of the particular sncRNA (miRNA/tRF/piRNA). The enrichment score was calculated using z-score approach. The z-score for each chromosome was calculated using randomisation (bed tools shuffle) to obtain the expected mean and standard deviation, while the actual sncRNA coordinates served as observed data. A z-score $> = 2$ represents over-representation and z-score $< = -2$ represents under-representation.

sncRNA association with paternal factors

For baseline characterization of sncRNA, raw QCed reads were processed to generate expression values in counts per million (cpm). sncRNA with cpm of ≥ 1 in all 17 subjects in the pre-intervention group were retained and normalized using Trimmed Mean of M (TMM) method. Using the normalized expression data for the miRNAs, piRNAs and tRFs, correlation with factors such as age, BMI, sperm concentration and sperm motility was carried out using spearman correlation test in R.

Differential expression of sncRNA between the control and intervention groups

Differentially expressed sncRNA between the control and intervention groups were analyzed using the Limma Bioconductor package[75]. sncRNA that had a cpm of ≥ 1 in at least 8 samples within each group were retained and normalized using the Trimmed Mean of M (TMM) method. The normalized data were subjected to double voom with sample specific quality weights transformation to down weigh the outlier samples. Since the samples were collected at two time points from the same individual, a paired analysis was performed using Limma's duplicate correlation method and blocking the ID. This analysis was adjusted for the confounding effects of age, BMI, sperm concentration and motility. The moderated t-statistics of differential expression was computed by empirical Bayes (eBayes), using a linear model fit (Lmfit) and comparing the **overall effect of intervention over control using the following equation: (Intervention.T2—Intervention.T1)—(Control.T2—Control. T1),** where T2 is the time point after the 6 weeks of intervention (MV4) and T1 is the time point of the first visit (MV1). Since none of the sncRNA passed the FDR corrected p-value of ≤ 0.05, we reported the sncRNA that passed the nominal p-value cut-off of ≤ 0.01 and log2(1.5) fold change.

Besides carrying out the differential expression analysis using the overall method (intervention_over_control) we also carried out the analysis for the control and intervention group separately as follows:

interventionEffect = (intervention.MV4—intervention.MV1)

controlEffect = (control.MV4—control.MV1)

Where interventionEffect and controlEffect are analysis for the control and intervention groups separately, comparing the expression profiles at the two-time points: pre intervention (MV1) and 6-week (MV4) post-intervention. We found that for miRNAs, the comparisons carried out for the intervention group separately resulted in a list of differentially expressed miRNAs (**Supplementary File S10, Sheet 1**) that were found to target genes enriched in pathways as shown in **Supplementary File S10, Sheet 2**, amongst these pathways, both Fatty Acid Biosynthesis and Metabolism are found to be enriched, the top most pathway being "Fatty Acid Biosynthesis" which was also detected in our original analysis (intervention over control) that considers the control and intervention group together. However, the comparisons carried out for the control group separately resulted in a very different list of differentially expressed miRNAs (**Supplementary File S10, Sheet 3**) that were found to target genes enriched in pathways as shown in **Supplementary File S10, Sheet 4**. The two pathways: Fatty Acid Biosynthesis and Metabolism do not show up in this list. This is indicative of the fact that our original model is very close to the intervention only model.

Determining the biological significance of sncRNA

miRNA target information and pathway enrichment using DIANA miRPath and IPA

DIANA-mirPath v3.0 (http://snf-515788.vm.okeanos.grnet.gr/) is an online software suite that makes use of the predicted miRNA targets provided by the DIANA-microT-CDS algorithm or experimentally validated miRNA-target interactions obtained from DIANA-TarBase. These predicted and/or experimentally validated targets are subsequently used for KEGG and Gene Ontology analysis. This software suite was used for the evaluation of miRNA target genes and the related functional pathways.

In addition to DIANA-mirPath we also used IPA (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) to assess the functional roles of miRNAs. This tool derives miRNA target information from databases such as TarBase, miRecords and TargetScan. Experimentally known as well as high confidence (high scoring) targets were filtered and used for the core analysis in IPA that reports the significantly enriched pathways. A threshold –log10 p-value of 2 was used to report the pathways in which the target genes were significantly enriched.

Determining the targets of the baseline and differentially expressed piRNAs

Two approaches were used to determine the functional annotation of piRNAs identified in this study. The first approach involved annotation of the genomic coordinates of the DE-piRNAs and baseline piRNAs using the annotatePeaks.pl program in the Homer motif analysis software package (http://homer.ucsd.edu/homer/). The second approach involved comparison of the piRNAs (piRNABank)[76] with the piRBase database[77] that comprises of information on repeat related and genes related piRNAs.

Data availability

All raw and processed sncRNA sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE159752. Due to ethical concerns, supporting clinical data cannot be made openly available. However, the PREPARE study team can provide the data upon request subject to appropriate approvals after a formal application to the PRE-PARE Study Oversight Group through the corresponding author. In-depth details of the analysis pipeline in the README file are provided at the Github page (https://github.com/Candida05/Diet_Intervention_Sperm).

References

1. McMillen, I. C. et al. Developmental origins of adult health and disease: the role of periconceptional and foetal nutrition. *Basic Clin. Pharmacol. Toxicol.* **102**, 82–89 (2008).
2. Velazquez, M. A., Fleming, T. P. & Watkins, A. J. Periconceptional environment and the developmental origins of disease. *J. Endocrinol.* **242**, T33–T49 (2019).
3. Fleming, T. P. et al. Environmental exposures around conception: developmental pathways leading to lifetime disease risk. *Int. J. Environ. Res. Public Health* https://doi.org/10.3390/ijerph18179380 (2021).
4. Fleming, T. P. et al. Origins of lifetime health around the time of conception: causes and consequences. *Lancet* **391**, 1842–1852 (2018).
5. Fleming, T. P., Eckert, J. J. & Denisenko, O. The role of maternal nutrition during the periconceptional period and its effect on offspring phenotype. *Adv. Exp. Med. Biol.* **1014**, 87–105 (2017).
6. Carone, B. R. et al. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* **143**, 1084–1096 (2010).
7. Ng, S. F. et al. Chronic high-fat diet in fathers programs beta-cell dysfunction in female rat offspring. *Nature* **467**, 963–966 (2010).
8. Ng, S. F. et al. Paternal high-fat diet consumption induces common changes in the transcriptomes of retroperitoneal adipose and pancreatic islet tissues in female rat offspring. *FASEB J* **28**, 1830–1841 (2014).
9. Crews, D. et al. Epigenetic transgenerational inheritance of altered stress responses. *Proc. Natl. Acad. Sci. U S A* **109**, 9143–9148 (2012).
10. Skinner, M. K. Environmental stress and epigenetic transgenerational inheritance. *BMC Med.* **12**, 153 (2014).
11. McCreary, J. K. et al. Altered brain morphology and functional connectivity reflect a vulnerable affective state after cumulative multigenerational stress in rats. *Neuroscience* **330**, 79–89 (2016).
12. Manikkam, M., Tracey, R., Guerrero-Bosagna, C. & Skinner, M. K. Pesticide and insect repellent mixture (permethrin and DEET) induces epigenetic transgenerational inheritance of disease and sperm epimutations. *Reprod. Toxicol.* **34**, 708–719 (2012).
13. Manikkam, M., Tracey, R., Guerrero-Bosagna, C. & Skinner, M. K. Plastics derived endocrine disruptors (BPA, DEHP and DBP) induce epigenetic transgenerational inheritance of obesity, reproductive disease and sperm epimutations. *PLoS One* **8**, e55387 (2013).
14. Skinner, M. K. et al. Ancestral dichlorodiphenyltrichloroethane (DDT) exposure promotes epigenetic transgenerational inheritance of obesity. *BMC Med.* **11**, 228 (2013).
15. Skinner, M. K. et al. Alterations in sperm DNA methylation, non-coding RNA and histone retention associate with DDT-induced epigenetic transgenerational inheritance of disease. *Epigenetics Chromatin* **11**, 8 (2018).
16. Deng, Z. et al. Association between air pollution and sperm quality: A systematic review and meta-analysis. *Environ. Pollut.* **208**, 663–669 (2016).
17. Donkin, I. & Barres, R. Sperm epigenetics and influence of environmental factors. *Mol. Metab.* **14**, 1–11 (2018).
18. Siddeek, B., Mauduit, C., Simeoni, U. & Benahmed, M. Sperm epigenome as a marker of environmental exposure and lifestyle, at the origin of diseases inheritance. *Mutat. Res.* **778**, 38–44 (2018).
19. Bedi, Y., Chang, R. C., Gibbs, R., Clement, T. M. & Golding, M. C. Alterations in sperm-inherited noncoding RNAs associate with late-term fetal growth restriction induced by preconception paternal alcohol use. *Reprod. Toxicol.* **87**, 11–20 (2019).
20. Fullston, T. et al. Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content. *FASEB J.* **27**, 4226–4243 (2013).
21. Gapp, K. et al. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* **17**, 667–669 (2014).
22. de Castro Barbosa, T. et al. High-fat diet reprograms the epigenome of rat spermatozoa and transgenerationally affects metabolism of the offspring. *Mol. Metab.* **5**, 184–197 (2016).
23. Farazi, T. A., Juranek, S. A. & Tuschl, T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **135**, 1201–1214 (2008).
24. Patil, V. S., Zhou, R. & Rana, T. M. Gene regulation by non-coding RNAs. *Crit. Rev. Biochem. Mol. Biol.* **49**, 16–32 (2014).
25. Shukla, G. C., Singh, J. & Barik, S. MicroRNAs: Processing, maturation, target recognition and regulatory functions. *Mol. Cell. Pharmacol.* **3**, 83–92 (2011).
26. Gruber, A. J. & Zavolan, M. Modulation of epigenetic regulators and cell fate decisions by miRNAs. *Epigenomics* **5**, 671–683 (2013).
27. Rodgers, A. B., Morgan, C. P., Leu, N. A. & Bale, T. L. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *Proc. Natl. Acad. Sci. U S A* **112**, 13699–13704 (2015).
28. Grandjean, V. et al. RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders. *Sci. Rep.* **5**, 18193 (2015).
29. Peng, H. et al. A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Res.* **22**, 1609–1612 (2012).
30. Cole, C. et al. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* **15**, 2147–2160 (2009).
31. Lee, Y. S., Shibata, Y., Malhotra, A. & Dutta, A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes. Dev.* **23**, 2639–2649 (2009).
32. Haussecker, D. et al. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* **16**, 673–695 (2010).
33. Kumar, P., Kuscu, C. & Dutta, A. Biogenesis and Function of Transfer RNA-Related Fragments (tRFs). *Trends Biochem. Sci.* **41**, 679–689 (2016).
34. Kumar, P., Mudunuri, S. B., Anaya, J. & Dutta, A. tRFdb: a database for transfer RNA fragments. *Nucleic Acids Res.* **43**, D141-145 (2015).
35. Chen, Q. et al. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* **351**, 397–400 (2016).
36. Sharma, U. et al. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* **351**, 391–396 (2016).
37. Natt, D. et al. Human sperm displays rapid responses to diet. *PLoS Biol.* **17**, e3000559 (2019).

38. Weick, E. M. & Miska, E. A. piRNAs: from biogenesis to function. *Development* **141**, 3458–3471 (2014).
39. Czech, B. & Hannon, G. J. One loop to rule them all: The ping-pong cycle and piRNA-guided silencing. *Trends Biochem. Sci.* **41**, 324–337 (2016).
40. Ernst, C., Odom, D. T. & Kutter, C. The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity. *Nat. Commun.* **8**, 1411 (2017).
41. Belicard, T., Jareosettasin, P. & Sarkies, P. The piRNA pathway responds to environmental signals to establish intergenerational adaptation to stress. *BMC Biol.* **16**, 103 (2018).
42. Vujkovic, M. et al. The preconception Mediterranean dietary pattern in couples undergoing in vitro fertilization/intracytoplasmic sperm injection treatment increases the chance of pregnancy. *Fertil. Steril.* **94**, 2096–2101 (2010).
43. Toledo, E. et al. Dietary patterns and difficulty conceiving: a nested case-control study. *Fertil. Steril.* **96**, 1149–1153 (2011).
44. Hammiche, F. et al. Increased preconception omega-3 polyunsaturated fatty acid intake improves embryo morphology. *Fertil. Steril.* **95**, 1820–1823 (2011).
45. Kermack, A. J., Calder, P. C., Houghton, F. D., Godfrey, K. M. & Macklon, N. S. A randomised controlled trial of a preconceptional dietary intervention in women undergoing IVF treatment (PREPARE trial). *BMC Womens Health* **14**, 130 (2014).
46. Kermack, A. J. et al. The fatty acid composition of human follicular fluid is altered by a 6-week dietary intervention that includes marine omega-3 fatty acids. *Lipids* **56**, 201–209 (2021).
47. Kermack, A. J. et al. Effect of a 6-week "Mediterranean" dietary intervention on in vitro human embryo development: the Preconception Dietary Supplements in Assisted Reproduction double-blinded randomized controlled trial. *Fertil. Steril.* **113**, 260–269 (2020).
48. Capra, E. et al. Small RNA sequencing of cryopreserved semen from single bull revealed altered miRNAs and piRNAs expression between High- and Low-motile sperm populations. *BMC Genomics* **18**, 14 (2017).
49. Zhu, Q., Kirby, J. A., Chu, C. & Gou, L. T. Small Noncoding RNAs in Reproduction and Infertility. *Biomedicines* https://doi.org/10.3390/biomedicines9121884 (2021).
50. Deng, W. & Lin, H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2**, 819–830 (2002).
51. Zhang, Y., Shi, J., Rassoulzadegan, M., Tuorto, F. & Chen, Q. Sperm RNA code programmes the metabolic health of offspring. *Nat. Rev. Endocrinol.* **15**, 489–498 (2019).
52. Lopez, P. et al. miR-155 and miR-122 Expression of Spermatozoa in Obese Subjects. *Front. Genet.* **9**, 175 (2018).
53. Ingerslev, L. R. et al. Endurance training remodels sperm-borne small RNA expression and methylation at neurological gene hotspots. *Clin. Epigenetics* **10**, 12 (2018).
54. Calder, P. C. Very long-chain n-3 fatty acids and human health: fact, fiction and the future. *Proc. Nutr. Soc.* **77**, 52–72 (2018).
55. Browning, L. M. et al. Incorporation of eicosapentaenoic and docosahexaenoic acids into lipid pools when given as supplements providing doses equivalent to typical intakes of oily fish. *Am. J. Clin. Nutr.* **96**, 748–758 (2012).
56. Fleming, T. P., Lucas, E. S., Watkins, A. J. & Eckert, J. J. Adaptive responses of the embryo to maternal diet and consequences for post-implantation development. *Reprod. Fertil. Dev.* **24**, 35–44 (2011).
57. Francia, S. Non-Coding RNA: Sequence-Specific Guide for Chromatin Modification and DNA Damage Signaling. *Front. Genet.* **6**, 320 (2015).
58. De Fazio, S. et al. The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* **480**, 259–263 (2011).
59. Robine, N. et al. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19**, 2066–2076 (2009).
60. Yu, T. et al. Long first exons and epigenetic marks distinguish conserved pachytene piRNA clusters from other mammalian genes. *Nat. Commun.* **12**, 73 (2021).
61. Toth, K. F., Pezic, D., Stuwe, E. & Webster, A. The piRNA pathway guards the germline genome against transposable elements. *Adv. Exp. Med. Biol.* **886**, 51–77 (2016).
62. Larriba, E. & Del Mazo, J. An integrative piRNA analysis of mouse gametes and zygotes reveals new potential origins and gene regulatory roles. *Sci. Rep.* **8**, 12832 (2018).
63. Li, S., Xu, Z. & Sheng, J. tRNA-derived small RNA: A novel regulatory small non-coding RNA. *Genes (Basel)* https://doi.org/10.3390/genes9050246 (2018).
64. Schimmel, P. The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat. Rev. Mol. Cell Biol.* **19**, 45–58 (2018).
65. Garcia-Lopez, J. et al. Diversity and functional convergence of small noncoding RNAs in male germ cell differentiation and fertilization. *RNA* **21**, 946–962 (2015).
66. Sobala, A. & Hutvagner, G. Transfer RNA-derived fragments: origins, processing, and functions. *Wiley Interdiscip Rev. RNA* **2**, 853–862 (2011).
67. Cao, L. L. et al. The effect of healthy dietary patterns on male semen quality: a systematic review and meta-analysis. *Asian J. Androl.* **24**, 549–557 (2022).
68. Collodel, G., Castellini, C., Lee, J. C. & Signorini, C. Relevance of fatty acids to sperm maturation and quality. *Oxid Med. Cell Longev.* **2020**, 7038124 (2020).
69. World Health Organization. (2010). WHO laboratory manual for the examination and processing of human semen, 5th ed. World Health Organization. https://iris.who.int/handle/10665/44261.
70. Lewis, S. E. et al. The impact of sperm DNA damage in assisted conception and beyond: recent advances in diagnosis and treatment. *Reprod. Biomed. Online* **27**, 325–337 (2013).
71. Fisk, H. L., West, A. L., Childs, C. E., Burdge, G. C. & Calder, P. C. The use of gas chromatography to analyze compositional changes of fatty acids in rat liver tissue during pregnancy. *J. Vis. Exp.* https://doi.org/10.3791/51445 (2014).
72. Goodrich, R. J., Anton, E. & Krawetz, S. A. Isolating mRNA and small noncoding RNAs from human sperm. *Methods Mol. Biol.* **927**, 385–396 (2013).
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).
75. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
76. Sai Lakshmi, S. & Agrawal, S. piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res.* **36**, D173-177 (2008).
77. Wang, J. et al. piRBase: a comprehensive database of piRNA sequences. *Nucleic Acids Res.* **47**, D175–D180 (2019).

Acknowledgements

Author contributions

Funding

Declarations

Competing interests

N.K. (Agency for Science, Technology and Research, Singapore), K.M.G. (University of Southampton), K.A.L. (University of Southampton), and Y.S.C. (National University of Singapore) are part of an academic consortium that has received research funding from Abbott Nutrition, Nestec, BenevolentAI Bio Ltd. and Danone. P.C.C. serves on the Scientific Advisory Board of Smartfish, the producer of the drinks used in the study. P.C.C. (University of Southampton,) has received advisory and/or speaking honoraria from Fresenius-Kabi, B. Braun, Baxter Healthcare, Abbott Nutrition, and Danone/Nutricia, sellers of parenteral and enteral feeds, and from DSM, Pronova BioPharma/BASF AS and Smartfish, sellers of products containing omega-3 fatty acids. K.M.G. (University of Southampton,) has received reimbursement for speaking at conferences sponsored by companies selling nutritional products. The remaining authors declare no competing interests. The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Ethics approval

Ethical approval for the PREPARE trial was granted by an NHS Research Ethics Committee (13/SC/0544). All participants provided written informed consent.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-83653-4.

Correspondence and requests for materials should be addressed to C.V. or N.K.

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.