



## ARTICLE

## Association of assisted reproductive technology, germline de novo mutations and congenital heart defects in a prospective birth cohort study

Cheng Wang<sup>1,2,3</sup>, Hong Lv<sup>1,2,4</sup>, Xiufeng Ling<sup>1,5</sup>, Hong Li<sup>1,6</sup>, Feiyang Diao<sup>1,7</sup>, Juncheng Dai<sup>1,2</sup>, Jiangbo Du<sup>1,2</sup>, Ting Chen<sup>8</sup>, Qi Xi<sup>9</sup>, Yang Zhao<sup>1,2</sup>, Kun Zhou<sup>1,2</sup>, Bo Xu<sup>1,2</sup>, Xiumei Han<sup>1,2</sup>, Xiaoyu Liu<sup>1,2</sup>, Meijuan Peng<sup>1,2</sup>, Congcong Chen<sup>1,2</sup>, Shiyao Tao<sup>1,2</sup>, Lei Huang<sup>1,2</sup>, Cong Liu<sup>1,2</sup>, Mingyang Wen<sup>1,2</sup>, Yangqian Jiang<sup>1,2</sup>, Tao Jiang<sup>1,2</sup>, Chuncheng Lu<sup>1,10</sup>, Wei Wu<sup>1,10</sup>, Di Wu<sup>1,10</sup>, Minjian Chen<sup>1,10</sup>, Yuan Lin<sup>1,4,11</sup>, Xuejiang Guo<sup>1</sup>, Ran Huo<sup>1</sup>, Jiayin Liu<sup>1,4,7</sup>, Hongxia Ma<sup>1,2</sup>, Guangfu Jin<sup>1,2</sup>, Yankai Xia<sup>1,10</sup>, Jiahao Sha<sup>1</sup>, Hongbing Shen<sup>1,2,4</sup> and Zhibin Hu<sup>1,2,4</sup>

Emerging evidence suggests that children conceived through assisted reproductive technology (ART) have a higher risk of congenital heart defects (CHDs) even when there is no family history. De novo mutation (DNM) is a well-known cause of sporadic congenital diseases; however, whether ART procedures increase the number of germline DNM (gDNM) has not yet been well studied. Here, we performed whole-genome sequencing of 1137 individuals from 160 families conceived through ART and 205 families conceived spontaneously. Children conceived via ART carried 4.59 more gDNMs than children conceived spontaneously, including 3.32 paternal and 1.26 maternal DNMs, after correcting for parental age at conception, cigarette smoking, alcohol drinking, and exercise behaviors. Paternal DNMs in offspring conceived via ART are characterized by C>T substitutions at CpG sites, which potentially affect protein-coding genes and are significantly associated with the increased risk of CHD. In addition, the accumulation of non-coding functional mutations was independently associated with CHD and 87.9% of the mutations were originated from the father. Among ART offspring, infertility of the father was associated with elevated paternal DNMs; usage of both recombinant and urinary follicle-stimulating hormone and high-dosage human chorionic gonadotropin trigger was associated with an increase of maternal DNMs. In sum, the increased gDNMs in offspring conceived by ART were primarily originated from fathers, indicating that ART itself may not be a major reason for the accumulation of gDNMs. Our findings emphasize the importance of evaluating the germline status of the fathers in families with the use of ART.

*Cell Research* (2021) 31:919–928; <https://doi.org/10.1038/s41422-021-00521-w>

## INTRODUCTION

The World Health Organization (WHO) predicts that infertility will be the third most serious disease worldwide in the 21st century.<sup>1</sup> In the past decade, assisted reproductive technology (ART) has gradually become a common practice in reproductive medicine clinics, and over 1.0% of children in China are conceived through ART.<sup>2</sup> Although ART has been widely used, emerging evidence has suggested that children conceived through ART have an elevated risk of adverse perinatal outcomes<sup>3–6</sup> and congenital diseases.<sup>3,7</sup> However, the underlying reasons remain largely unknown.

Germline de novo mutations (gDNMs) are new mutations that have occurred within one generation, provide offspring with

genetic variants in addition to those inherited from their parents,<sup>8</sup> and account for nearly 90% of total DNMs.<sup>9</sup> Multiple studies have reported a considerable role of deleterious DNMs in unfavorable perinatal outcomes as well as in congenital and developmental diseases,<sup>10–14</sup> such as congenital heart defects (CHDs).<sup>15–18</sup> Recent evidence has shown that DNMs can accumulate with DNA replication during spermatogonium division or the double-strand break repair process of oocytes.<sup>8</sup> Thus, it is necessary to evaluate whether ART, which includes the artificial manipulation of both gametes and early embryos, contributes to the occurrence of DNMs and the risk of subsequent diseases. However, such evaluation is challenging because the risk factors of infertility may

<sup>1</sup>State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing, Jiangsu, China; <sup>2</sup>Department of Epidemiology and Biostatistics, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China; <sup>3</sup>Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, Jiangsu, China; <sup>4</sup>State Key Laboratory of Reproductive Medicine (Suzhou Centre), The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou, Jiangsu, China; <sup>5</sup>Department of Reproduction, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Hospital, Nanjing, Jiangsu, China; <sup>6</sup>Reproductive Genetic Center, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou, Jiangsu, China; <sup>7</sup>Clinical Center of Reproductive Medicine, The First Affiliated Hospital of Nanjing Medical University, Nanjing Medical University, Nanjing, Jiangsu, China; <sup>8</sup>Scientific Education Section, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Hospital, Nanjing, Jiangsu, China; <sup>9</sup>Department of Obstetrics, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou, Jiangsu, China; <sup>10</sup>Key Laboratory of Modern Toxicology of Ministry of Education, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China and <sup>11</sup>Department of Maternal, Child and Adolescent Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China

Correspondence: Hongbing Shen (hbshen@njmu.edu.cn) or Zhibin Hu (zhibin\_hu@njmu.edu.cn)

These authors contributed equally: Cheng Wang, Hong Lv, Xiufeng Ling, Hong Li, Feiyang Diao

Received: 14 January 2021 Accepted: 17 May 2021

Published online: 9 June 2021

also increase the occurrence of DNMs. For example, parental age has been linked to the risk of infertility<sup>19</sup> and the accumulation of DNMs.<sup>9,20–23</sup> Thus, the effect of parental factors on DNMs should be carefully evaluated and controlled in the analyses of association between ART and DNMs.

Here, we performed whole-genome sequencing (WGS) of 1137 individuals from 365 families (including 407 offspring) from the China National Birth Cohort (CNBC) to investigate the effect of ART-related elements on gDNMs and their associations with CHD.

## RESULTS

### General description of DNMs

Among 365 parent-offspring families, 202 offspring (including 86 twin offspring from 44 families; Supplementary information, Fig. S1) were born from ART pregnancies (ARTP) and 205 from spontaneous pregnancies (SP). The information of all offspring and their parents were listed in Supplementary information, Table S1. After stringent quality control (QC) of WGS data (see Materials and Methods), we identified a total of 15,862 high-confidence point autosomal gDNMs in all offspring for the subsequent analysis, with an average of 38.97 mutations per offspring. Then, we defined the parental origin of the DNMs by tracing the read pairs near the phased variants, resulting in a total of 4879 gDNMs (30.76%), including 3785 (23.86%) paternal DNMs and 1094 (6.90%) maternal DNMs in all offspring. The standardized DNM number of each parental origin was estimated by dividing the number of phased DNMs by the proportion phased in each trio and the callable base number, and then the number was standardized to the whole genome length ( $2.70 \times 10^9$ ) for the following association analyses.

Offspring conceived through ART showed significantly more gDNMs than those conceived spontaneously

We found that the ARTP offspring carried significantly more gDNMs than the SP offspring (ARTP vs SP, total gDNM: Beta = 7.29,  $P = 2.23 \times 10^{-11}$ ; paternal DNM: Beta = 5.43,  $P = 1.17 \times 10^{-7}$ ; maternal DNM: Beta = 1.90,  $P = 1.48 \times 10^{-3}$ ; Fig. 1). As parental age at conception was the well-known risk factor of gDNMs<sup>9,20–23</sup> as well as infertility,<sup>24</sup> we investigated the association between parental age and DNMs and re-evaluated the effect of ART after controlling for parental age. In line with previous studies, the number of paternal DNMs specifically increased with paternal age (Beta = 0.91 per year,  $P = 2.96 \times 10^{-18}$ ; Supplementary information, Fig. S2). The effect of maternal age on maternal DNMs was significant but smaller (Beta = 0.39 per year,  $P = 1.03 \times 10^{-6}$ ; Supplementary information, Fig. S2). After adjusting for the

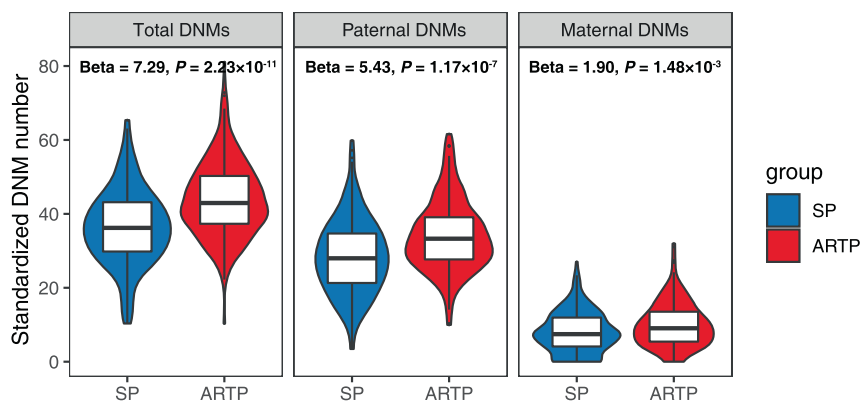
parental age at conception, the ARTP offspring still carried significantly more DNMs than the SP offspring (ARTP vs SP, total gDNM: Beta = 4.67,  $P = 1.52 \times 10^{-6}$ ; paternal DNM: Beta = 3.37,  $P = 4.99 \times 10^{-4}$ ; maternal DNM: Beta = 1.29,  $P = 0.032$ ).

Parental lifestyle behaviors were associated with gDNMs of their offspring

Parental lifestyle data obtained from the baseline survey were presented in Supplementary information, Table S1. After adjustment for parental age and conception type, paternal drinking behavior was significantly associated with an increased number of paternal gDNMs, while paternal exercise was associated with decreased paternal gDNMs (Supplementary information, Table S2).

In the multivariable model, fathers' drinking behavior was associated with increased paternal DNMs in the offspring (Beta = 1.88,  $P = 0.047$ ), while fathers' exercise habit was associated with decreased gDNMs (Beta = -2.20,  $P = 0.048$ , Table 1). We observed that there is no heterogeneity of the effects in the ARTP and SP groups (Cochran's Q-tests for heterogeneity,  $P > 0.05$ ; Supplementary information, Fig. S3). Paternal DNMs were specifically associated with paternal alcohol consumption and exercise, suggesting that these environmental factors might influence generation of the gDNMs during the spermatogenesis process. In our cohort, we observed a low proportion of maternal smoking (1.37%) and drinking (11.78%). Thus, the effect of maternal smoking and drinking behaviors could not be accurately evaluated in the current study.

We noticed that the effect of ART remained significant in the model (ARTP vs SP, total gDNM: Beta = 4.59,  $P = 2.38 \times 10^{-6}$ ; paternal DNM: Beta = 3.32,  $P = 6.52 \times 10^{-4}$ ; maternal DNM: Beta = 1.26,  $P = 0.039$ ; Table 1). Similar results were seen in the model with the correction for parental age as categorical variables (Supplementary information, Table S3). The results were consistent if we adjusted parental singleton mutation loads, which represent mutation accumulation in recent generations and parental germline mutation rates,<sup>25</sup> as covariates in the multivariable model (Supplementary information, Table S4). Because the age of ARTP parents was much higher than that of parents in the SP group (Supplementary information, Table S1), we conducted a matching on the parental age, which resulted in 192 ARTP offspring and 152 SP offspring. In the matched datasets, there was no significant difference in parental age in ARTP and SP groups (Student's *t*-test,  $P_{\text{paternal age}} = 0.14$ ,  $P_{\text{maternal age}} = 0.30$ ; Supplementary information, Fig. S4). The effect of ART in the matched data was similar to that in Table 1 (Supplementary information, Table S5).



**Fig. 1** Offspring conceived through ART showed significantly more total (left), paternal (middle), and maternal (right) DNMs than those conceived spontaneously. ART assisted reproductive technology, ARTP assisted reproductive technology pregnancies, SP spontaneous pregnancies, DNMs, de novo mutations.

**Table 1.** Multivariable regression of parental age at conception, lifestyles, and conception type in all offspring.

Classification	Factors	Levels	Total DNMs ( $n = 407$ ) <sup>a</sup>		Paternal DNMs ( $n = 407$ ) <sup>b</sup>		Maternal DNMs ( $n = 407$ ) <sup>c</sup>	
			Beta (95%CI)	$P$	Beta (95%CI)	$P$	Beta (95%CI)	$P$
Parental age at conception	Paternal age at conception	–	0.87 (0.58–1.17)	<b><math>1.54 \times 10^{-8}</math></b>	0.82 (0.52–1.12)	<b><math>1.36 \times 10^{-7}</math></b>	0.05 (–0.13–0.24)	0.570
	Maternal age at conception	–	0.30 (–0.08–0.67)	0.124	–0.03(–0.41–0.35)	0.871	0.33 (0.09–0.57)	<b><math>7.66 \times 10^{-3}</math></b>
Paternal lifestyles	Smoking status	No	ref	–	ref	–	ref	–
		Yes	1.49 (–0.35–3.34)	0.114	0.75 (–1.11–2.62)	0.429	0.74 (–0.43–1.91)	0.217
	Drinking behavior (> 1/month)	No	ref	–	ref	–	ref	–
		Yes	0.84 (–0.99–2.68)	0.369	1.88 (0.03–3.73)	<b>0.047</b>	–1.02 (–2.19–0.14)	0.086
Exercise habit (> 3/week)	No	ref	–	ref	–	ref	–	
	Yes	–3.13 (–5.28––0.98)	<b><math>4.53 \times 10^{-3}</math></b>	–2.20 (–4.37––0.03)	<b>0.048</b>	–0.93 (–2.30–0.43)	0.181	
Maternal lifestyles	Exercise habit (> 3/week)	No	ref	–	ref	–	ref	–
		Yes	–1.29 (–3.11–0.53)	0.165	–1.72 (–3.56–0.11)	0.067	0.42 (–0.73–1.58)	0.476
Conception type	Conception type	SP	ref	–	ref	–	ref	–
		ARTP	4.59 (2.72–6.47)	<b><math>2.38 \times 10^{-6}</math></b>	3.32 (1.43–5.22)	<b><math>6.52 \times 10^{-4}</math></b>	1.26 (0.07–2.46)	<b>0.039</b>

Bold indicates statistically significant ( $P < 0.05$ ).

DNMs de novo mutations, *ref* reference group, *ARTP* ART pregnancy, *SP* spontaneous pregnancy.

<sup>a</sup>Total DNM number ~ paternal age + maternal age + paternal smoking status + paternal drinking behavior + paternal exercise habit + maternal exercise habit + conception type.

<sup>b</sup>Paternal DNM number ~ paternal age + maternal age + paternal smoking status + paternal drinking behavior + paternal exercise habit + maternal exercise habit + conception type.

<sup>c</sup>Maternal DNM number ~ paternal age + maternal age + paternal smoking status + paternal drinking behavior + paternal exercise habit + maternal exercise habit + conception type.

#### Characterized mutation spectrum of gDNMs in the offspring

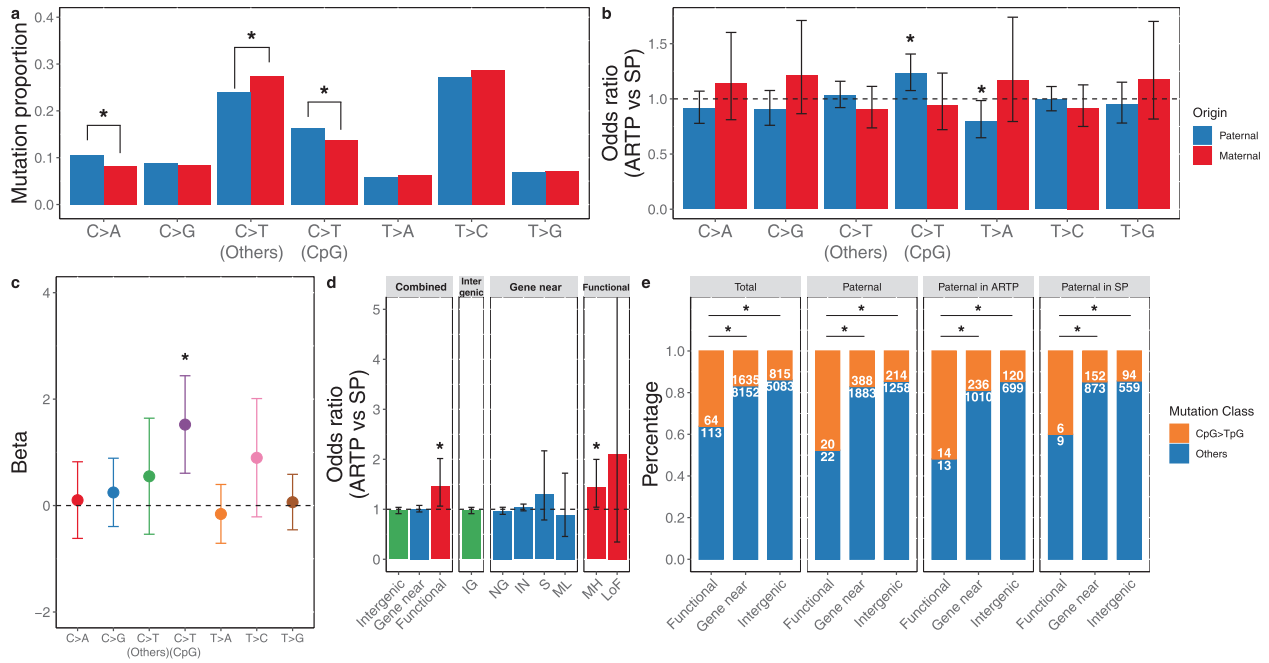
In all offspring, the classification of gDNMs by mutation type revealed that gDNMs predominantly comprised C>T and T>C transitions and that the relative frequencies of C>T transitions at the non-CpG sites were greater in maternal than paternal transmissions (Fisher's exact test,  $P_{C>T \text{ at non-CpG}} = 2.30 \times 10^{-5}$ ), while C>T transitions at the CpG sites and C>A transversions were rarer (Fisher's exact test,  $P_{C>A} = 3.56 \times 10^{-5}$ ,  $P_{C>T \text{ at CpG}} = 8.15 \times 10^{-5}$ ; Fig. 2a). Moreover, we found that paternal C>T transitions at the CpG sites were significantly elevated in the ARTP group after Bonferroni correction ( $P_{adj \text{ C>T at CpG}} = 2.00 \times 10^{-4}$ ; Fig. 2b). Interestingly, the ART effect on paternal gDNMs was also significant for C>T transitions at the CpG sites (Beta = 1.52,  $P = 0.001$ ; Fig. 2c). These results suggested that the increase in paternal DNMs in the ARTP offspring was primarily characterized by C>T transitions at the CpG sites, which were mutation hotspots induced by cytosine deamination damage. For maternal DNMs, we did not observe a significant difference in the proportion of mutation classes between ARTP and SP groups (Fig. 2b).

The functional enrichment of paternal C>T mutations at CpG sites We classified all mutations into three major classes: functional coding gDNMs (Functional), other gDNMs near genes (Gene near), and intergenic mutations (Intergenic). The majority of the gDNMs were annotated as Gene near (9787/15,862 = 61.7%) and Intergenic (5898/15,862 = 37.2%), while only 177 mutations (1.1%) were classified as Functional. We observed that the mutations in the Functional group showed 1.46-fold enrichment in the ARTP group (proportion of functional DNMs in ARTP group: 112/8606 = 1.30%; in SP group: 65/7256 = 0.90%; Fig. 2d). Importantly, we found that the loss-of-function mutations (LoFs) had the highest odds ratio (OR = 2.11), and the ratio increased with the functional CADD score of missense mutations (Fig. 2d). Interestingly, C>T mutations at the CpG sites were significantly enriched in the Functional class (Functional vs Intergenic enrichment, OR = 3.53,  $P = 2.52 \times 10^{-13}$ ; Fig. 2e). The enrichment ratio increased in the analyses of paternal DNMs (Functional vs Intergenic enrichment, OR = 5.34,  $P = 6.02 \times 10^{-7}$ ; Fig. 2e). We observed a similar enrichment trend for paternal DNMs in the ARTP and SP groups,

and the enrichment ratio was higher in the ARTP (Functional vs Intergenic enrichment, OR = 6.25,  $P = 1.05 \times 10^{-5}$ ) group than in the SP group (Functional vs Intergenic enrichment, OR = 3.95,  $P = 0.016$ ; Fig. 2e).

#### The association between gDNMs and CHDs

We further investigated the associations between gDNMs and offspring diseases based on the cohort follow-up at age 1 year and information from the disease registry. CHD is the most common type of birth defect, with seven offspring from ARTP families (3.5%) and two from SP families (0.98%). We did not find any large copy number alterations ( $\geq 10$  Mb) or pathogenic variants inherited from parents in the reported genes for CHDs. Interestingly, we found that paternal C>T gDNMs at CpG sites (Supplementary information, Table S6), which were enriched in the Functional mutation class, significantly increased the risk of CHD (OR = 1.20,  $P = 0.002$ ). Functional mutations in known CHD-related genes contributed to a higher risk of CHD (OR = 22.95,  $P = 0.012$ ; Table 2). Given that 98.2% gDNMs were located in the non-coding regions of the genome, we further determined a total of 128 potential pathogenic non-coding DNMs in CHD based on a deep learning method<sup>26</sup> (Supplementary information, Table S6). The number of these DNMs was significantly associated with CHD (OR = 2.96,  $P = 0.005$ ; Table 2). Among these mutations, 33 had defined parental origin and the majority of these mutations (29/33, 87.9%) were originated from the father. Additional structural equation modeling suggested significantly indirect effects of paternal C>T gDNMs at CpG sites and non-coding functional gDNMs as mediators between ART and the risk of CHD (paternal C>T gDNMs: OR = 1.17,  $P = 0.008$ ; non-coding functional gDNMs, OR = 2.76,  $P = 0.016$ ), and the indirect effects were independent (Fig. 3). When we included parental age in the model, we observed a similar indirect effect of paternal C>T DNMs at CpG sites in the association between ART and CHD (Supplementary information, Fig. S5a). Thus, we concluded that offspring conceived by ART carried more paternal C>T DNMs at CpG sites, which greatly increased the risk of CHD. The effects of ART and parental age on non-coding functional DNMs were not significant (Supplementary information, Fig. S5b), suggesting that we cannot



**Fig. 2 The influence of ART on the mutational spectrum.** **a** Relative frequency of mutational classes by parental origins. **b** The odds ratio of proportion comparison for specific mutational classes between the ARTP and SP groups. The *P* value and odds ratios were adjusted by Bonferroni methods. **c** The association (mutation rate ratio) between specific mutation class and ARTP/SP in the multivariable models. **d** The odds ratio of proportion comparison for specific functional groups between the ARTP and SP groups (IG, intergenic; NG, nearby genes; IN, intron; S, synonymous; ML, missense mutations with CADD ≤ 15; MH, missense mutations with CADD scores > 15). **e** The proportion of C>T mutations at the CpG sites in each functional group. \**P* < 0.05.

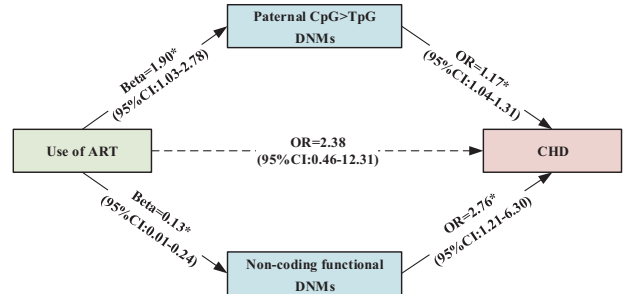
**Table 2.** The association between CHD and the number of DNMs.

DNM types	Mean standardized DNM number		OR	L95	U95	<i>P</i>
	Case (n = 9)	Control (n = 398)				
Total <sup>a</sup>	41.04	40.18	1.01	0.95	1.07	0.814
Paternal <sup>b</sup>	33.41	31.16	1.02	0.96	1.09	0.510
Maternal <sup>c</sup>	7.63	9.02	0.96	0.85	1.08	0.483
Paternal CpG>TpG <sup>d</sup>	10.31	5.10	1.20	1.07	1.34	<b>0.002</b>
Functional mutations in CHD-related genes <sup>e</sup>	0.11	0.01	22.95	2.00	263.46	<b>0.012</b>
Predicted non-coding pathogenic mutations <sup>f</sup>	0.92	0.31	2.96	1.39	6.34	<b>0.005</b>

Bold indicates statistically significant (*P* < 0.05).  
 L95 95% lower confidence limit, U95 95% upper confidence limit, CHD congenital heart defect.  
<sup>a</sup>Logistic regression: the event of CHD ~ total DNMs.  
<sup>b</sup>Logistic regression: the event of CHD ~ paternal DNMs.  
<sup>c</sup>Logistic regression: the event of CHD ~ maternal DNMs.  
<sup>d</sup>Logistic regression: the event of CHD ~ paternal CpG>TpG DNMs.  
<sup>e</sup>Logistic regression: the event of CHD ~ functional mutations in CHD-related genes.  
<sup>f</sup>Logistic regression: the event of CHD ~ predicted non-coding pathogenic mutations.

determine the causal factors of non-coding functional DNMs with the current sample size.

The relationship between ART-related elements and gDNMs  
 We further investigated whether ART-related elements, including twin pregnancy, ART procedures, and parental infertility reasons



**Fig. 3 Structural equation models showed that paternal CpG>TpG DNMs and non-coding functional DNMs acted as independent mediators between ART and risk of CHD.** The solid lines denoted significant (*P* < 0.05) relationships and dashed lines non-significant. \**P* < 0.05. CHD congenital heart defect, OR odds ratio, CI confidence interval.

were associated with gDNMs. With correction for parental age and lifestyle factors, we found that paternal infertility reason and usage of follicle-stimulating hormone (FSH) were specifically associated with gDNMs (Supplementary information, Table S7) and thus were entered into the multivariate analyses. We also included controlled ovarian stimulation protocols, embryo transfer strategies, fertilization methods, and trigger strategies in the following models because they were important procedures of ART. We observed that usage of both recombinant and urinary FSH (Beta = 2.69, *P* = 0.023) and usage of high-dosage human chorionic gonadotropin (hCG) trigger (Beta = 2.76, *P* = 0.032) were independently associated with an increased number of maternal gDNMs (Table 3). Similar results were seen in the model with the correction for parental age as categorical variables (Supplementary information, Table S8). In addition, we found that offspring with a father diagnosed with oligoasthenoteratozoospermia carried significantly more paternal gDNMs (Beta = 2.91, *P* =

**Table 3.** Multivariable regression including parental infertility-related diseases and detailed information on ART procedures.

Class	Factors	Levels	Total DNMs (n = 202) <sup>a</sup>		Paternal DNMs (n = 202) <sup>b</sup>		Maternal DNMs (n = 202) <sup>c</sup>	
			Beta (95%CI)	P	Beta (95%CI)	P	Beta (95%CI)	P
ART procedure details	COS protocols	Antagonist	ref	-	ref	-	ref	-
		Short agonist	1.14 (-3.27-5.54)	0.614	1.19 (-3.17-5.55)	0.595	-0.06 (-3.00-2.88)	0.967
	ET strategies	Long agonist	1.90 (-3.29-7.09)	0.474	2.67 (-2.47-7.81)	0.311	-0.75 (-4.22-2.72)	0.672
		Microstimulation	-2.26 (-9.23-4.71)	0.526	-5.59 (-12.50-1.32)	0.115	3.32 (-1.33-7.98)	0.164
	Fertilization methods	Frozen	ref	-	ref	-	ref	-
		Fresh	0.51 (-2.82-3.84)	0.764	-1.57 (-4.87-1.72)	0.351	2.08 (-0.14-4.30)	0.068
	Subtype of FSH	ICSI	ref	-	ref	-	ref	-
		IVF	0.05 (-3.44-3.53)	0.979	-0.65 (-4.11-2.81)	0.713	0.70 (-1.63-3.03)	0.556
	Trigger strategies	rFSH	ref	-	ref	-	ref	-
		uFSH	0.41 (-3.32-4.15)	0.828	2.11 (-1.59-5.81)	0.265	-1.70 (-4.19-0.80)	0.185
Both		0.64 (-2.79-4.07)	0.715	-2.05 (-5.45-1.35)	0.239	2.69 (0.40-4.98)	<b>0.023</b>	
Single hCG (Low)		ref	-	ref	-	ref	-	
Single hCG (High)		-0.63 (-4.36-3.10)	0.741	-3.38 (-7.08-0.32)	0.075	2.76 (0.27-5.25)	<b>0.032</b>	
Single GnRH <sub>a</sub>		-1.31 (-6.39-3.77)	0.614	-0.86 (-5.89-4.18)	0.739	-0.46 (-3.85-2.93)	0.791	
Parental diseases	Double	0.46 (-4.72-5.63)	0.863	-0.61 (-5.73-4.52)	0.817	1.07 (-2.38-4.53)	0.544	
	OAT	ref	-	ref	-	ref	-	
Ovarian factors	No	2.86 (-0.01-5.74)	0.053	2.91 (0.06-5.76)	<b>0.047</b>	-0.06 (-1.98-1.86)	0.954	
	Yes	ref	-	ref	-	ref	-	
Tubal factors	Others	0.66 (-3.34-4.66)	0.747	0.20 (-3.76-4.16)	0.922	0.46 (-2.21-3.13)	0.737	
	PCOS/DOR/POI	2.56 (-1.00-6.12)	0.161	1.44 (-2.09-4.96)	0.426	1.12 (-1.26-3.49)	0.357	
Endometriosis	No	ref	-	ref	-	ref	-	
	Yes	-0.01 (-3.54-3.52)	0.996	-1.88 (-5.38-1.61)	0.293	1.87 (-0.48-4.23)	0.122	
	No	ref	-	ref	-	ref	-	
	Yes	1.50 (-2.95-5.96)	0.509	1.16 (-3.27-5.58)	0.609	0.34 (-2.64-3.33)	0.821	

Bold indicates statistically significant ( $P < 0.05$ ).  
 ref reference group, COS controlled ovarian stimulation, ET embryo transfer, ICSI intracytoplasmic sperm injection, IVF in vitro fertilization, FSH follicle-stimulating hormone, rFSH recombinant FSH, uFSH urinary FSH, LH luteinizing hormone, hCG human chorionic gonadotropin, GnRH<sub>a</sub> gonadotropin-releasing hormone agonist, OAT oligoasthenoteratozoospermia, PCOS polycystic ovarian syndrome, DOR diminished ovarian reserve, POI premature ovarian insufficiency.  
<sup>a</sup>Total DNM number + ET strategies + fertilization methods + subtype of FSH + trigger strategies + OAT + ovarian factors + tubal factors + endometriosis + paternal age + maternal age + paternal smoking status + paternal drinking behavior + paternal exercise habit + maternal exercise habit + conception type.  
<sup>b</sup>Paternal DNM number ~ COS protocols + ET strategies + fertilization methods + subtype of FSH + trigger strategies + OAT + ovarian factors + tubal factors + endometriosis + paternal age + maternal age + paternal smoking status + paternal drinking behavior + paternal exercise habit + maternal exercise habit + conception type.  
<sup>c</sup>Maternal DNM number ~ COS protocols + ET strategies + fertilization methods + subtype of FSH + trigger strategies + OAT + ovarian factors + tubal factors + endometriosis + paternal age + maternal age + paternal smoking status + paternal drinking behavior + paternal exercise habit + maternal exercise habit + conception type.

0.047; Table 3). The model without the correction for parental lifestyles showed similar results (Supplementary information, Table S9).

## DISCUSSION

By using WGS data in families from a prospective cohort, the present study showed that the application of ART was associated with the risk of CHD through accumulating functional gDNMs. However, the majority of the mutations were originated from the fathers rather than the mothers. The evidence suggested that ART procedures may not be the major reason for the increased number of DNMs that contribute to the risk of CHD, even though specific operations were associated with the increased trends of maternal gDNMs.

The present study found that the mutation class with the greatest increase in ART offspring was paternal C>T mutations at CpG sites, which are commonly generated by cytosine deamination damage during DNA replication. However, the sperm genome cannot replicate after meiosis until fertilization.<sup>27</sup> It is possible, therefore, that the increase in paternal gDNMs may not be related to subsequent ART procedures but to paternal spermatogenesis abnormalities in fathers who underwent ART. Importantly, paternal gDNMs in the ARTP group are frequently located in the functional sites of genes and are associated with an increased risk of CHDs, suggesting the association between paternal genome and offspring diseases. Similarly, a recent study revealed the connection between paternal age, gDNMs, and offspring diseases,<sup>28</sup> proposing the hypothesis that the number of mutations in the sperm increases with paternal age and the chance that a child would carry a deleterious mutation that could lead to diseases (e.g., autism or schizophrenia) increases significantly.<sup>23</sup> Thus, careful evaluation of the sperm genome of fathers might be important to prevent disease-causing mutations from passing to their children through ART. Previous studies have widely reported an increase in sperm aneuploidy levels in infertile men.<sup>29,30</sup> In line with these studies, our results suggested that paternal DNMs increased in the infertile father, which partially explained the increased paternal DNMs in the ARTP group. However, we should not ignore other mechanisms besides parental genetic causes. For example, a recent study found that the association between CHD incidence and ART pregnancies is largely mediated by twinning, which provided additional potential mechanisms.<sup>31</sup> Since there is only one CHD case diagnosed in the families with twins in this study, a cohort with increased sample size is warranted for the evaluation of the mediation effect of twinning and DNMs in the association between ART and CHD. It is also interesting that some paternal behaviors, such as drinking and exercise, show the trends to influence the number of paternal gDNMs, providing the evidence for the association between environmental factors and the number of DNMs,<sup>32</sup> though the associations observed in the current sample size require further validation. Recent proteomic analyses suggested that plasma protein patterns could be greatly altered by lifestyles, such as smoking, drinking, and exercise.<sup>33</sup> Thus, the influence of parental lifestyles could be complicated and warranted further investigation in the cohorts with multi-omics data followed by solid experiments. Taken together, we emphasize the importance of the fathers' prepregnancy factors, which may help prevent the accumulation of paternal DNMs.

Similarly, maternal DNMs also increased in the ARTP group, but we did not find any characterized mutation class. Thus, ART procedures may not generate maternal DNMs through a different mechanism. Specific ART procedures, such as ovarian stimulation and trigger protocols, indeed contributed to the number of maternal DNMs. Both processes stimulated the oocyte meiosis resumption<sup>34,35</sup> and maturation.<sup>36</sup> Consistently, a recent study found that meiotic recombination could play a role in the formation of maternal DNMs.<sup>20</sup> However, we did not find evidence

of the association between maternal DNMs and CHD. Although the non-coding functional mutations were important mediators between ART and CHD, maternal DNMs accounted for only a small proportion of these mutations. Thus, ART procedures may not directly contribute to the risk of CHD through the mechanisms of gDNM accumulation.

The primary strength of our study is the use of a well-designed birth cohort. First, our cohort recruited each family as a unit, and then prospectively and longitudinally collected comprehensive health information as well as biospecimens from parents and offspring, which enabled the study on the association between pre-pregnant exposures, parental diseases, ART procedures, and gDNMs. Second, the CNBC cohort was applied with the Cloud Based Cohort Management System, which provided the paperless questionnaire survey and intelligent follow-up arrangement. Thus, the efficiency of data collection, QC, and cohort participant management were greatly improved, promising the follow-up with high quality as well as high response rate. In addition, we applied identical follow-up schedules to ART and spontaneous births by investigators and physicians without knowing their conception mode to avoid bias during data collection and CHD diagnosis. With complete information on nuclear families and high-quality follow-up, we were able to elucidate the association between pre-pregnant exposures, parental diseases, ART procedures, gDNMs, and CHD in the offspring step by step. Third, our study emphasized the contribution of the ignored paternal genome in ART families to the risk of congenital diseases, and thus offered new insight into the evaluation of the safety of ART from the aspects of genomic landscapes.

This study has several limitations. First, we cannot provide a full explanation for the elevation of DNM burden in the ART children because of the limited sample size, though we found several important factors associated with DNMs in addition to paternal age. The completion of CNBC with a much larger sample size will allow a comprehensive evaluation of the effect of many other environmental factors (e.g., air pollution and other exposure factors before and during pregnancies) on gDNMs and their association with CHD as well as other birth defects with low incidence.<sup>37–40</sup> It should be mentioned that our cohort may not be suitable for the evaluation of the effect of maternal smoking and drinking behaviors because of the low rate in our population. Second, a relatively low rate of origin inferences from haplotypes of DNMs by short-read sequencing also limited the power of the study. Advances in long-read sequencing<sup>41</sup> in the cohort with a larger sample size will allow for more accurate origin inference of DNMs and enable a more comprehensive analysis to determine additional factors responsible for the increase of gDNMs in the ARTP group, and eventually guide the prevention of gDNMs and CHD in the clinics.

In summary, the increased gDNMs in offspring conceived by ART were primarily originated from fathers, indicating that ART itself may not be a major reason for the accumulation of gDNMs. These results emphasized the importance of evaluating the germline status of the fathers in ART families and provided new insight into the prevention of congenital diseases in the ART offspring.

## MATERIALS AND METHODS

### Study design and participants

All participating couples who had live-born infants conceived with and without ART between December 2014 and March 2018 were enrolled in two hospitals (i.e., the Women's Hospital of Nanjing Medical University and Suzhou Municipal Hospital). Reproductive endocrinologists, obstetricians, or professional health workers explained the objectives, procedures, potential benefits, and confidentiality issues of the study to all eligible couples. Written informed consent was obtained from those who agreed to

participate. All methods and protocols for information collection were approved by the institutional review board of Nanjing Medical University.

This prospective cohort was one of the sub-cohort of CNBC, and only families with live-born infants were studied (Supplementary information, Fig. S1). We included 181 families from the ARTP group and 231 families from the SP group. In the ARTP group, 19 were excluded from the sequencing experiments (18 were lost to follow-up and 1 was in vitro fertilized with donated semen). In the SP group, 21 families were excluded because they were lost to follow-up. Because only two SP families with twins participated in the cohort, they were also excluded from the study to avoid their influence on the following regression model. In two ARTP families with twins, blood was collected from only one child of the twins (Supplementary information, Fig. S1). Finally, we conducted WGS on 1152 individuals from 370 families, including 412 offspring.

Baseline information was collected with standardized and structured questionnaires before pregnancy, and blood specimens of the parents were collected on the day of recruitment. Clinical information, including infertility diagnoses, specific ART procedures at each cycle, and health status during pregnancy were extracted from medical records. All cohort children were offered free child health examinations from their infancy to early childhood as scheduled. Fingertip blood samples were collected at the children's 1st year clinic visits.

Once reaching 1 year of age, we followed up with all participants and their infants received systematic physical examination at child health care clinics. All CHD cases were diagnosed by echocardiography and confirmed by two pediatric cardiologists.

#### Variables

ART was the main exposure in the study. Parental lifestyles (smoking (yes/no), drinking (yes/no) and exercise habit (yes/no)), infertility diagnoses in the ART group (oligoasthenoteratozoospermia, ovulatory dysfunction, tubal factors, endometriosis), and ART procedures were studied in the study. Participants who reported drinking at least once a month were defined as having drinking behavior. Those who reported exercising at least three times a week were defined as having exercise habits. Ovulatory dysfunction is typical of producing eggs abnormally and can be further sorted into polycystic ovarian syndrome (PCOS), diminished ovarian reserve (DOR), premature ovarian insufficiency (POI), and others. This diagnosis was evaluated by three senior chief physicians independently. Parental infertility diagnoses in the ART group were directly included in the multivariable model in Table 3.

ART procedures included: (1) controlled ovarian stimulation (COS) protocols: COS protocols aim to stimulate the ovaries to produce the available oocytes in a controlled manner, thus the protocols may influence the number and quality of oocytes. We classified COS protocols used in the clinics into four major classes: antagonist, short agonist, long agonist, and microstimulation. The dosage and type of drugs used were different across protocols. The detailed dosage of the drug (i.e., gonadotrophin-releasing hormone agonist (GnRHa), antagonist, gonadotropin (Gn)) and the selection of FSH, luteinizing hormone (LH), growth hormone, and trigger drug were also studied. hCG trigger with a dosage over 6000 IU was defined as high hCG usage. (2) Fertilization methods: in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI). In IVF, the oocyte and sperm are left in a petri dish to fertilize on their own. In ICSI, one sperm is directly injected into an oocyte. The process included the manipulation of both sperm and egg. (3) Embryo transfer (ET) strategy: ET is a step to place embryos into the uterus of a female. A fresh ET cycle: embryos were selected and transferred immediately after the oocyte retrieval, fertilization and 2–7 days of embryo culture. In contrast, when embryos are frozen and thawed for transfer when the proper condition is available, this is referred to as a frozen ET cycle.

The COS protocols, ET strategy, fertilization methods, and trigger strategies were directly included in the multivariable model in Table 3.

#### WGS

The genomic DNAs of the offspring and their parents were extracted from blood samples and sequenced with the Illumina HiSeq X Ten System, and these DNA samples were prepared for sequencing according to the manufacturer's instructions (Illumina). Briefly, DNA was extracted and sheared into fragments that were then purified by gel electrophoresis. DNA fragments were ligated with adapter oligonucleotides to form paired-end DNA libraries. To enrich the libraries for sequencing with 5' and 3' adapters, we used ligation-mediated PCR amplification with primers complementary to the adapter sequences. The DNA libraries were sequenced to generate 150 bp pair-end reads and targeted ~30× coverage for each sample.

#### Alignment and variant calling

The FastQC package (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) was used to assess the quality-score distribution of the sequencing reads. Read sequences were mapped to the human reference genome (GRCh37) using the Burrows-Wheeler Aligner (BWA-MEM v0.7.15-r1140)<sup>42</sup> with the default parameters. The quality metrics (including coverage, median insert size, and percentage of chimeric reads) and contamination estimation of "Bam" files were obtained by Picard (v1.70) (<http://broadinstitute.github.io/picard>) and VerifyBamID2 (v1.04) (<https://github.com/Griffan/VerifyBamID>), respectively. Duplicates were marked and discarded using Picard (v1.70). Before variant calling, we applied Canvas<sup>43</sup> for the identification of large copy number changes (≥10 Mb) in all trios.

Single-nucleotide and short insertion/deletion variants were jointly called using the Genome Analysis Toolkit (GATK v3.8.1)<sup>44</sup> Best Practices for germline SNVs and indels. Briefly, samples were called individually using local realignment by HaplotypeCaller in the gVCF mode with default read filtering parameters. Individual gVCF files were jointly genotyped for high-confidence alleles using GenotypeGVCFs for all autosomes and X chromosome of genomes. We performed relatedness and sex examination based on a temporary list of high-confidence variants: biallelic, high-call rate (> 0.99), common SNVs (allele frequency > 1%). We estimated pairwise relatedness by KING (v2.2.4)<sup>45</sup> and inferred the chromosomal sex of the samples based on the inbreeding coefficient (F) for common variants on X chromosome, excluding pseudoautosomal regions.

#### QC

**Sample QC.** We excluded individuals with (1) average coverage < 15×; (2) median insert size < 250 bp; (3) excess chimeric reads:  $pct\_chimera > 0.05$ ; (4) high contamination:  $freemix > 0.05$ ; (5) unmatched relatedness; and (6) inconsistent chromosomal sex and gender information. Families that failed to satisfy the "trios" design after sample QC were excluded from subsequent analyses. We excluded five families with members who failed in sample QC, and eventually included 1137 individuals from 365 families with 202 offspring born from ARTP and 205 from SP. The general information and detailed ART procedures for these families are listed in Supplementary information, Table S1. The detailed information of all offspring and their parents in our study and read-level QC metrics (i.e., PF HQ ERROR RATE and STRAND BALANCE) are listed in Supplementary information, Table S10.

**Callable region.** Because not all positions in the genome could be sequenced and this fraction varied across the trios, we defined the callable region as those that had a minimum depth ≥12 for each individual and the callable base number as the total base of the intersect callable region for each of the trios. We found that there

was no significant difference in the callable base number between the ARTP and SP groups (ARTP mean: 26.11 Gb, SP mean: 26.14 Gb, Student's *t*-test,  $P = 0.69$ ). The callable base number was used as the offset in the subsequent Poisson regression model.

**Variant QC.** We only included point DNMs in this study. GATK variant quality score recalibration (VQSR) was used to filter variants. The SNP VQSR model was trained using HapMap 3.3 and 1000 Genomes Omni 2.5 SNP sites, and a 99.6% sensitivity threshold was applied to filter variants. We excluded variants if they (1) failed to pass VQSR filtering criteria; (2) had an inbreeding coefficient  $< -0.3$ ; (3) were located in segmental duplication or simple tandem repeats downloaded from UCSC genome browser (<http://genome.ucsc.edu/>); and (4) were out of the callable regions.

#### Identification and filtering of gDNMs

We restricted our DNM analysis to the SNVs in the autosomes and called DNMs using DeNovoGear (v1.1.1)<sup>46</sup> and the GATK Genotype Refinement workflow<sup>44</sup> for each offspring-parents trio. Consistent DNMs called from two pipelines were further filtered by the following criteria: (1) minor allele frequency (MAF)  $< 0.01$  in the total gnomAD control samples; (2) MAF  $< 0.01$  in a custom-curated WGS database with 6600 Chinese individuals; (3) MAF = 0 in all parents included in this cohort; (4) the offspring must be an alternative allele carrier; (5) minimum depth  $\geq 12$  for parents and offspring; (6) no more than one read supporting the alternative allele in the parent; (7) maximum allelic fraction for the parent of 0.05; and (8) minimum allelic fraction for the offspring of 0.20.

We evaluated our pipeline based on public data from 70 individuals of three-generation families.<sup>9</sup> We collected a total of 4176 DNMs in this study. After we excluded (1) indel variants, (2) variants in X chromosome, (3) variants in segmental duplication/simple repeats, (4) variants with MAF  $\geq 0.01$  in the total gnomAD control samples, and (5) variants with MAF  $\geq 0.01$  in a custom-curated WGS database with 6600 Chinese individuals, we curated a dataset including 3807 DNMs as a set of true DNMs. We identified 3433 DNMs by our protocols and a detection rate was 90.2%. To assess the accuracy of our DNM calls, we randomly selected 92 DNMs (0.5% of total DNMs, including 15 paternal DNMs and 4 maternal DNMs) and sequenced both parents and offspring by Sanger sequencing. Of these, 86 DNMs were successfully validated (True positive: 93.47%) and all DNMs with clear parental origin were validated. We noticed that all 6 DNMs that failed in the validation had the problem of strand bias and homologous mapping, thus we applied a custom script to filter the called DNMs. After that, all 6 failed DNMs and 1 validated DNM were excluded. After applying the filtering to all the DNM mutations, we evaluated the sharing DNMs in a monozygotic twin family. We identified 59 and 60 DNMs for two offspring, respectively, 56 of which (94.9% and 93.3%) were shared.

To diminish the influence of post-zygotic mutations, we further excluded mutations with allelic fraction lower than 0.35. We evaluated the threshold in a publicly available database (CEPH/Utah pedigrees study),<sup>9</sup> which included DNMs validated by three-generation pedigrees. We downloaded genotype data from dbGaP with accession number phs001872.v1.p1. The validated DNMs were obtained from the GitHub repository from a previous study (<https://github.com/quinlan-lab/ceph-dnm-manuscript>). The threshold excluded the majority (75%, 303/404) of post-zygotic DNMs at a cost of 7.54% (287/3807) germline mutations. We listed variant-level QC metrics of all DNMs in Supplementary information, Table S11.

The standardized DNM number of each parental origin was estimated by dividing the number of phased DNMs by the proportion phased in each trio and the callable base number, then the number was standardized to the whole genome length ( $2.7 \times 10^9$ ). The DNM number of specific mutation type was directly standardized to the whole genome length ( $2.7 \times 10^9$ ).

We evaluated the association between the standardized DNM number and other QC metrics that might influence the mutation rate. DNM number was not significantly associated with PF HQ ERROR RATE of offspring and their parents (offspring: Pearson's correlation coefficient  $r = -0.09$ ,  $P = 0.061$ ; father:  $r = 0.00$ ,  $P = 0.925$ ; mother:  $r = -0.07$ ,  $P = 0.150$ ; Supplementary information, Fig. S6). There was no significant difference between ARTP and SP groups for PF HQ ERROR RATE of offspring and their parents (Student's *t*-test,  $P_{\text{offspring}} = 0.052$ ,  $P_{\text{father}} = 0.112$ , and  $P_{\text{mother}} = 0.491$ ; Supplementary information, Fig. S7). All data were strand balance (STRAND BALANCE = 0.500). In addition, MQ, FS, and GQ of variants identified in ARTP and SP groups were comparable (Student's *t*-test,  $P_{\text{MQ}} = 0.479$ ,  $P_{\text{FS}} = 0.760$ , and  $P_{\text{GQ}} = 0.552$ ; Supplementary information, Table S11).

We defined singleton variants<sup>25</sup> of the parents if the SNV (1) passed the general QC for variants; (2) MAF = 0 in the total gnomAD control samples; (3) MAF = 0 in a custom-curated WGS database of 6600 Chinese; (4) was successfully transmitted to their offspring. Then, the singleton loads of the parents were divided by the callable base number of each of the trios and rescaled to the Mb level.

#### Phasing and origin classification of gDNMs

We applied a haplotype assembly strategy similar to that used in the previous studies<sup>9,20–23</sup> to trace the origin of the gDNM allele. Briefly, we considered that the human autosomes were diploid and that a gDNM affected only one allele from either the father or the mother, and thus the other informative variants (uniquely attributable to a single parent) on the same haplotype could help distinguish the origin of the allele of the gDNM. Besides, we also applied a pedigree-based haplotype inference method<sup>47</sup> to increase the inference rate of informative variants. For 2295 mutations with origin inferred from both informative variants themselves and pedigree-based haplotype, we compared the inference results. The consistent rate was 99.9% (2293/2295). In our cohort, we obtained the origin information for 4879 DNMs out of a total of 15,862 (30.76%). Among these gDNMs, 3785 (77.58%) were of paternal origin, and 1094 (22.42%) were of maternal origin. The percentages are comparable to those of phased mutations reported in recently published studies.<sup>9,20–23</sup>

#### Mutation spectrum

Similar to the approach used in the previous studies,<sup>9,20–23</sup> we grouped DNMs into six classes according to nucleotide substitution. Besides, we considered dichotomized C>T substitutions depending on the CpG context, i.e., C>T on CpG sites and C>T on non-CpG sites, resulting in a total of seven classes. Then, we calculated the enrichment of the mutational classes between different (DNM origin or ARTP/SP) groups. For each mutational class, we performed Fisher's exact test on a  $2 \times 2$  table, tabulating the numbers of DNMs by group and by the mutational class. For the comparison between paternal and maternal DNMs, which have been reported by previous studies, we presented original Odds ratios and *P* values. To account for multiple tests of the comparison between ARTP and SP groups, we applied Bonferroni correction by dividing the obtained *P* values by the number of statistical tests.

#### Variant annotation and functional classification

We used SnpEff (v4.3.1, <http://snpeff.sourceforge.net/>) to annotate all sequence variants and their potential mutational effects on the associated genes, and used vcfanno (v0.3.2)<sup>48</sup> to annotate CADD score v1.4<sup>42</sup> for each mutation. Mutations annotated as UTR variants and up- and downstream gene variants were reannotated as other mutations near genes. Then, we grouped all mutations into three major classes: (1) functional coding DNMs (Functional); LoFs, missense mutations with CADD scores  $> 15$  (Missense High,



MH), (2) other DNMs near genes (Gene near): missense mutations with  $CADD \leq 15$  (Missense Low, ML), synonymous mutations, intron mutations and other mutations of nearby genes; and (3) intergenic mutations (Intergenic). Similar to the analysis for the mutational spectrum, we calculated the enrichment of the functional classes for the ARTP and SP groups by using Fisher's exact test. We also annotated the DNMs with ClinVar database (version 20191021, <https://www.ncbi.nlm.nih.gov/clinvar/>) and none of the variants was annotated as high-confidence (with two or more stars) pathogenic or likely pathogenic (P/LP) variants.

We also applied the HeartENN model, which predicts molecular effect differences between any two alleles for every regulatory feature by using convolutional neural networks, to predict the potentially pathogenic non-coding DNMs.<sup>26</sup>

### Statistical analysis

The Student's *t*-test was used to compare the differences of continuous baseline characteristics between the ARTP and SP groups. Fisher's exact test was used to compare the differences of categorical baseline characteristics between ARTP and SP groups and investigate the enrichment of the mutational classes between different (DNM origin or ARTP/SP) groups. We fit linear mixed models (LMM) to evaluate the association of ART (or other exposures) with standardized gDNM numbers (or DNMs with specific origin), accounting for the twin information in the cohort. We fit logistic regression models to estimate the association between standardized gDNM numbers (or DNMs with a specific origin) and the risk of CHD. Structural equation modeling was conducted using the piecewiseSEM package<sup>49</sup> according to the models constructed by nlme<sup>50</sup> and lme4 packages.<sup>51</sup> Two-sided *P* values < 0.05 were considered significant. All analyses were conducted in R (v3.6.0).

### ACKNOWLEDGEMENTS

This work was funded by the National Key R&D Program of China (2018YFC1004200, 2016YFC1000200), and the National Natural Science Foundation of China (81830100). We are grateful to all the families participating in this study, and the whole CNBC cohort team. We thank the Nanjing Jiangbei New Area Biopharmaceutical Public Service Platform Co., Ltd, National Health & Medical Big Data Center (Nanjing), and National Human Genetic Resources Sharing Service Platform Jiangsu Innovation Center (YCZYPT[2018]05).

### AUTHOR CONTRIBUTIONS

Z.H. and H.S. initiated, conceived, and supervised the study. Z.H. and C.W. performed bioinformatics/statistical analysis and prepared the manuscript with C.C. H. Lv, X. Ling, H. Li, and J. Du were involved in study design, the conduct of the CNBC cohort study, long-term follow-up with T.C., Q.X., Y.Z., K.Z., B.X., X.H., X. Liu, M.P., S.T., L.H., C. Liu, M.W., C. Lu, W.W., D.W., M.C., and Y. L. H. Lv, F.D., J. Du, Y.Z., M.P., S.T., Y.J., and T.J. organized clinical information. J. Dai conducted the sequencing experiments with K.Z., B.X., X.H., X. Liu, L.H., C. Liu, and M.W. X.G., R.H., J.L., H.M., G.J., Y.X., and J.S. proofread the manuscript.

### ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41422-021-00521-w>.

**Competing interests:** The authors declare no competing interests.

### REFERENCES

- Kushnir, V. A., Barad, D. H., Albertini, D. F., Darmon, S. K. & Gleicher, N. Systematic review of worldwide trends in assisted reproductive technology 2004–2013. *Reprod. Biol. Endocrinol.* **15**, 6 (2017).
- Yang, X., Li, Y., Li, C. & Zhang, W. Current overview of pregnancy complications and live-birth outcome of assisted reproductive technology in mainland China. *Fertil. Steril.* **101**, 385–391 (2014).

- Hansen, M., Kurinczuk, J. J., Milne, E., de Klerk, N. & Bower, C. Assisted reproductive technology and birth defects: a systematic review and meta-analysis. *Hum. Reprod. Update* **19**, 330–353 (2013).
- Romundstad, L. B. et al. Effects of technology or maternal factors on perinatal outcome after assisted fertilisation: a population-based cohort study. *Lancet* **372**, 737–743 (2008).
- Niakan, K. K., Han, J., Pedersen, R. A., Simon, C. & Pera, R. A. Human pre-implantation embryo development. *Development* **139**, 829–841 (2012).
- Declercq, E. et al. Perinatal outcomes associated with assisted reproductive technology: the Massachusetts Outcomes Study of Assisted Reproductive Technologies (MOSART). *Fertil. Steril.* **103**, 888–895 (2015).
- Boulet, S. L. et al. Assisted reproductive technology and birth defects among liveborn infants in Florida, Massachusetts, and Michigan, 2000–2010. *JAMA Pediatr.* **170**, e154934 (2016).
- Goldmann, J. M., Veltman, J. A. & Gilissen, C. De novo mutations reflect development and aging of the human germline. *Trends Genet.* **35**, 828–839 (2019).
- Sasani, T. A. et al. Large, three-generation human families reveal post-zygotic mosaicism and variability in germline mutation accumulation. *Elife* **8**, e46922 (2019).
- Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature* **542**, 433–438 (2017).
- Iossifov, I. et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**, 216–221 (2014).
- Epi, K. C. et al. De novo mutations in epileptic encephalopathies. *Nature* **501**, 217–221 (2013).
- Short, P. J. et al. De novo mutations in regulatory elements in neurodevelopmental disorders. *Nature* **555**, 611–616 (2018).
- de Ligt, J. et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* **367**, 1921–1929 (2012).
- Homsy, J. et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* **350**, 1262–1266 (2015).
- Zaidi, S. et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* **498**, 220–223 (2013).
- Sifrim, A. et al. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat. Genet.* **48**, 1060–1065 (2016).
- Jin, S. C. et al. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nat. Genet.* **49**, 1593–1601 (2017).
- Menken, J., Trussell, J. & Larsen, U. Age and infertility. *Science* **233**, 1389–1394 (1986).
- Goldmann, J. M. et al. Germline de novo mutation clusters arise during oocyte aging in genomic regions with high double-strand-break incidence. *Nat. Genet.* **50**, 487–492 (2018).
- Jonsson, H. et al. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).
- Goldmann, J. M. et al. Parent-of-origin-specific signatures of de novo mutations. *Nat. Genet.* **48**, 935–939 (2016).
- Kong, A. et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
- Bergh, C., Pinborg, A. & Wennerholm, U. B. Parental age and child outcomes. *Fertil. Steril.* **111**, 1036–1046 (2019).
- Carlson, J. et al. Extremely rare variants reveal patterns of germline mutation rate heterogeneity in humans. *Nat. Commun.* **9**, 3753 (2018).
- Richter, F. et al. Genomic analyses implicate noncoding de novo variants in congenital heart disease. *Nat. Genet.* **52**, 769–777 (2020).
- Gunes, S., Al-Sadaan, M. & Agarwal, A. Spermatogenesis, DNA damage and DNA repair mechanisms in male infertility. *Reprod. Biomed. Online* **31**, 309–319 (2015).
- McGrath, J. J. et al. A comprehensive assessment of parental age and psychiatric disorders. *JAMA Psychiatry* **71**, 301–309 (2014).
- Hassold, T. & Hunt, P. To err (meiotically) is human: the genesis of human aneuploidy. *Nat. Rev. Genet.* **2**, 280–291 (2001).
- Harton, G. L. & Tempest, H. G. Chromosomal disorders and male infertility. *Asian J. Androl.* **14**, 32 (2012).
- Wen, S. W. et al. Associations of assisted reproductive technology and twin pregnancy with risk of congenital heart defects. *JAMA Pediatr.* **174**, 446–454 (2020).
- Kessler, M. D. et al. De novo mutations across 1,465 diverse genomes reveal mutational insights and reductions in the Amish founder population. *Proc. Natl. Acad. Sci. USA* **117**, 2560–2569 (2020).
- Williams, S. A. et al. Plasma protein patterns as comprehensive indicators of health. *Nat. Med.* **25**, 1851–1857 (2019).
- Shimada, M. & Terada, T. FSH and LH induce progesterone production and progesterone receptor synthesis in cumulus cells: a requirement for meiotic resumption in porcine oocytes. *Mol. Hum. Reprod.* **8**, 612–618 (2002).
- Coticchio, G. et al. Mouse oocyte meiotic resumption and polar body extrusion in vitro are differentially influenced by FSH, epididymal growth factor and meiosis-activating sterol. *Hum. Reprod.* **19**, 2913–2918 (2004).

36. Schachter, M. et al. Can pregnancy rate be improved in gonadotropin-releasing hormone (GnRH) antagonist cycles by administering GnRH agonist before oocyte retrieval? A prospective, randomized study. *Fertil. Steril.* **90**, 1087–1093 (2008).
37. Zhang, X. et al. Changes in maternal age and prevalence of congenital anomalies during the enactment of China's universal two-child policy (2013-2017) in Zhejiang Province, China: an observational study. *PLoS Med.* **17**, e1003047 (2020).
38. Yan, J. et al. Birth defects after assisted reproductive technologies in China: analysis of 15,405 offspring in seven centers (2004 to 2008). *Fertil. Steril.* **95**, 458–460 (2011).
39. Davies, M. J. et al. Reproductive technologies and the risk of birth defects. *N. Engl. J. Med.* **366**, 1803–1813 (2012).
40. Zhu, J. L., Basso, O., Obel, C., Bille, C. & Olsen, J. Infertility, infertility treatment, and congenital malformations: Danish national birth cohort. *BMJ* **333**, 679 (2006).
41. Logsdon, G. A., Vollger, M. R. & Eichler, E. E. Long-read human genome sequencing and its applications. *Nat. Rev. Genet.* **21**, 597–614 (2020).
42. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019).
43. Roller, E., Ivakhno, S., Lee, S., Royce, T. & Tanner, S. Canvas: versatile and scalable detection of copy number variants. *Bioinformatics* **32**, 2375–2377 (2016).
44. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
45. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
46. Ramu, A. et al. DeNovoGear: de novo indel and point mutation discovery and phasing. *Nat. Methods* **10**, 985–987 (2013).
47. Patterson, M. et al. WhatsHap: weighted haplotype assembly for future-generation sequencing reads. *J. Comput. Biol.* **22**, 498–509 (2015).
48. Pedersen, B. S., Layer, R. M. & Quinlan, A. R. Vcfanno: fast, flexible annotation of genetic variants. *Genome Biol.* **17**, 118 (2016).
49. Lefcheck, J. S. piecewiseSEM: Piecewise structural equation modelling in R for ecology, evolution, and systematics. *Methods Ecol. Evol.* **7**, 573–579 (2016).
50. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, R.C. nlme: linear and non-linear mixed effects models. *R package version 3*, 111 (2013).
51. Bates, D., Sarkar, D., Bates, M. D. & Matrix, L. The lme4 package. *R package version 2*, 74 (2007).