SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

June 2023 Vol.66 No.6: 1392–1407
https://doi.org/10.1007/s11427-022-2277-y

DEtail-seq is an ultra-efficient and convenient method for meiotic DNA break profiling in multiple organisms

Wei Xu[1,4,5†*], Chao Liu[2†], Zhe Zhang[3†], Changbin Sun[5], Qin Li[1,4], Kuan Li[1,4], Hui Jiang[3*], Wei Li[2*] & Qianwen Sun[1,4*]

[1]*School of Life Sciences, Tsinghua University, Beijing 100084, China;*
[2]*Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou 510623, China;*
[3]*Department of Urology, Department of Andrology, Department of Reproductive Medicine Center, and Department of Human Sperm Bank, Peking University Third Hospital, Beijing 100191, China;*
[4]*Tsinghua-Peking Center for Life Sciences, Beijing 100084, China;*
[5]*Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China*

Received November 17, 2022; accepted January 2, 2023; published online January 19, 2023

Programmed DNA double-strand break (DSB) formation is a crucial step in meiotic recombination, yet techniques for high-efficiency and precise mapping of the 3′ ends of DSBs are still in their infancy. Here, we report a novel technique, named DNA End tailing and sequencing (DEtail-seq), which can directly and ultra-efficiently characterize the 3′ ends of meiotic DSBs with near single-nucleotide resolution in a variety of species, including yeast, mouse, and human. We find that the 3′ ends of meiotic DSBs are stable without significant resection in budding yeast. Meiotic DSBs are strongly enriched in *de novo* H3K4me3 peaks in the mouse genome at leptotene stage. We also profile meiotic DSBs in human and find DSB hotspots are enriched near the common fragile sites during human meiosis, especially at CCCTC-binding factor (CTCF)-associated enhancers. Therefore, DEtail-seq provides a powerful method to detect DSB ends in various species, and our results provide new insights into the distribution and regulation of meiotic DSB hotspots.

meiosis, DSB, DNA break, fragile site, enhancer, spermatogenesis

Citation: Xu, W., Liu, C., Zhang, Z., Sun, C., Li, Q., Li, K., Jiang, H., Li, W., and Sun, Q. (2023). DEtail-seq is an ultra-efficient and convenient method for meiotic DNA break profiling in multiple organisms. Sci China Life Sci 66, 1392–1407. https://doi.org/10.1007/s11427-022-2277-y

INTRODUCTION

In eukaryotes, meiosis is a fundamental process required for sexual reproduction, which involves a single round of DNA replication followed by two consecutive cell divisions (Ohkura, 2015). During meiosis, meiotic recombination results in the exchange of genetic material between homologous chromosomes, which is beneficial for genetic diversity (Keeney et al., 2014). Meiotic recombination is initiated with programmed DNA double-strand breaks (DSBs) induced by Spo11, an evolutionarily conserved transesterase-like enzyme that covalently binds to the 5′ end of DNA through a phosphotyrosyl linkage (de Massy, 2013; Keeney et al., 2014). Following break formation, a Mre11-Rad50-Xrs2/NBS1 (MRX) complex plus Sae2/CTIP generates nicks on Spo11-bound DNA to facilitate DNA end resection through Mre11 endo- and 3′ to 5′ exonuclease activity and robust 5′ to 3′ Exo1 exonuclease activity (Figure S1A in Supporting Information) (Garcia et al., 2011; Neale et al., 2005; Pan et

†Contributed equally to this work
*Corresponding authors (Wei Xu, email: xuwei01@caas.cn; Hui Jiang, email: jianghui55@163.com; Wei Li, email: leways@lzu.edu.cn; Qianwen Sun, email: sunqianwen@mail.tsinghua.edu.cn)

al., 2011). During this process, MRE11 first cleaves the 5′ strand away from the DSB through its endonuclease activity. Then, MRX complex generates short single-stranded DNA through MRE11 3′ to 5′ exonuclease activity (Symington, 2014). The DNA resection gives rise to 3′ single-stranded DNA (ssDNA) tails that serve as substrates for recombinases DMC1 and RAD51, which search for homology and invade homologous repair templates (Borde and de Massy, 2013; Mimitou et al., 2017). We now know that meiotic DSBs are not randomly located in the genome, but are preferentially enriched within specific regions called hotspots (Baudat et al., 2013).

Spo11-induced meiotic DSBs can be divided into three parts: upstream DNA ends, downstream DNA ends, and Spo11-bound oligos (de Massy, 2013) (Figure S1A in Supporting Information). Both upstream and downstream DNA ends provide 3′ overhang structures, which contain a free 3′ ssDNA end (Mimitou et al., 2017) (Figure S1A in Supporting Information). Currently, several methods have been developed to profile meiotic DSBs, including S1-seq (S1 sequencing) (Mimitou et al., 2017), END-seq (Canela et al., 2016), Spo11-oligo-seq (Pan et al., 2011), DMC1 SSDS (single-stranded DNA sequencing) (Pratto et al., 2014), and CC-seq (Gittens et al., 2019) (Table S1 in Supporting Information). However, these methods cannot directly provide information for the precise location of 3′ overhang ends in meiotic DSBs (Figure S1A in Supporting Information), confining the scope of application to different organisms, such as human. To improve the limitations of those methods, two DNA break profiling methods, TrAEL-seq (Kara et al., 2021) and GLOE-seq (Sriramachandran et al., 2020), were recently developed to map the 3′ end of DNA break. In TrAEL-seq, TdT (terminal deoxynucleotidyl transferase) is used to add 1 to 4 adenosine nucleotides onto ssDNA 3′ ends followed by adaptor ligation with T4 RNA ligase 2 truncated KQ (Figure S1B in Supporting Information) (Kara et al., 2021). In GLOE-seq, DNA 3′ ends are ligated to adaptors by T4 DNA ligase utilizing a splinter oligonucleotide with 6 random bases after dissociative or embedded genomic DNA (gDNA) denatured at 95°C (Figure S1B in Supporting Information) (Sriramachandran et al., 2020). However, there are some disadvantages of these two methods that may limit their utilization (Table S1 in Supporting Information). The efficiency of the first adaptor ligation in TrAEL-seq is not high enough (10% (Kara et al., 2021)). GLOE-seq requires gDNA denaturation before the first adaptor ligation based on a splinter with random bases and has not been used for meiotic DSB profiling (Sriramachandran et al., 2020). In addition, both of these two methods are relatively complicated and time-consuming, especially GLOE-seq which requires five days to prepare libraries (Petrosino et al., 2020) (Table S1 in Supporting Information). Thus, as the 3′ overhang structure is the distinctive feature of meiotic DSB ends (Figure S1A in Supporting Information, (Baudat et al., 2013; de Massy, 2013)), we sought to develop a new method called DNA End tailing and sequencing (DEtail-seq) to profile the 3′ ends of meiotic DSB directly and efficiently, without denaturation treatment and immunoprecipitation (Figure S1A in Supporting Information). DEtail-seq relies on a high-efficiency ssDNA ligation technique that uses Adaptase (see Methods), which we have previously used to detect ssDNA strands in RNA:DNA hybrids throughout the genome in multiple species (Li et al., 2020; Xu et al., 2020; Xu et al., 2017; Yang et al., 2019; Yuan et al., 2019). After embedding and lysis, cells in an agarose plug are lysed, so the unpaired 3′ overhangs of DSBs can be directly ligated to ssDNA adaptors (Figure 1D).

RESULTS

Adaptase is an ideal tool for DNA 3′ end mapping

Initially, we set up an efficient and convenient method for mapping 3′ tail of DNA break, especially for meiotic DSB, named DEtail-seq. Adaptase (Swift Accel-NGS, Accel-NGS 1S Plus DNA Library Kit) is a commercial enzyme complex that generates a ligation product containing a low complexity polynucleotide tail (Y-tail) and P7 truncated adaptor (Figure 1A and B). To test the efficiency of Adaptase, we ligated the adaptor to an ssDNA accepter with three random bases at the 3′ end. The result showed a very high efficiency of ligation, which was almost 100% (Figure 1C). Because of this near-complete ligation of ssDNA substrates with the random end (Figure 1C), which means each ssDNA with different end sequences was completely ligated, it is a very reasonable inference that Adaptase has no significant base bias for ssDNA ligation. Thus, with very high ligation efficiency and low bias, which was most likely due to Y-tailing and splinter-based ligation strategy, Adaptase is an ideal tool for DNA break detection.

The brief workflow of DEtail-seq is shown in Figure 1D: first, cell embedding and lysis; second, ssDNA adaptor was ligated to the 3′ end of DNA break sites catalyzed by Adaptase; third, DNA extraction, fragmentation, denaturation, extension, and the rest steps followed the instruction. All steps for library preparation, from cell embedding to purification of indexing PCR product, took at most 2 days. In DEtail-seq, antibody, transgenic line, immunoprecipitation (used in SSDS, Spo11-oligo-seq, or CC-seq) and biotin enrichment steps (used in TrAEL-seq, END-seq, and S1-seq) are not required, which strongly simplified the procedure and reduced the specimen loss (Table S1 in Supporting Information). In addition, DEtail-seq does not denature gDNA before 1st adaptor ligation (which is used in GLOE-seq), as the meiotic DSB contains a single-stranded 3′ overhang structure and can be ligated directly, and the signal-to-noise

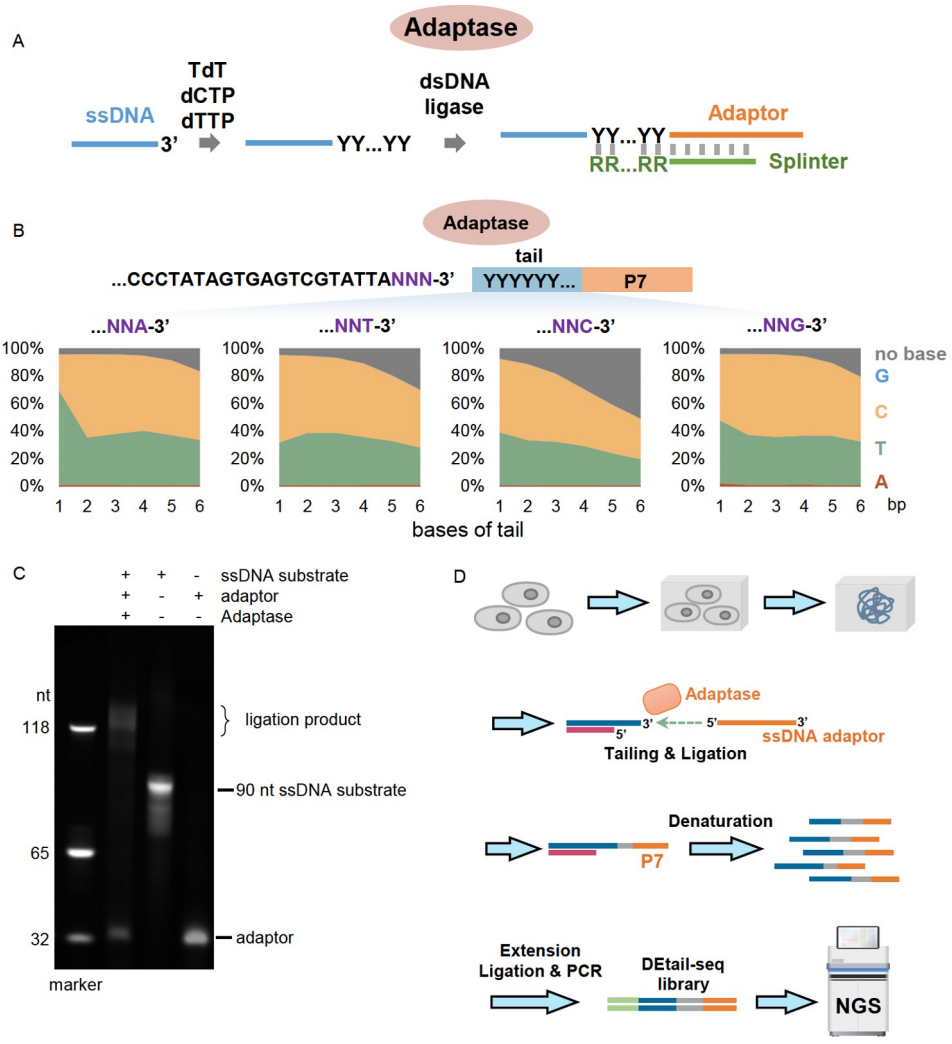

Figure 1 The characteristics and performance of Adaptase. A, As the detailed principle of Adaptase was not completely described in the instruction, based on the information from the instruction and related patent application (US 20200392551A1 and CN 104395480B), the speculated process of Adaptase-induced ssDNA ligation based on the sequencing results of library was shown. Y=C or T, R=A or G. B, The sequencing results showed the frequencies of bases of tails ligated at different types of 3′ ssDNA ends. *X*-axis 1–6 represent the first, second, to sixth bases of tail, as the nearest base to the substrate was defined as the first base. C, *In vitro* assay of the first adaptor ligation by Adaptase. An ssDNA oligonucleotide was ligated to ssDNA adaptor by Adaptase at 37°C for 1 h. Ligation products were separated on a 15% PAGE gel with ssDNA markers. D, Brief schematic of DEtail-seq detection of DSBs with 3′ overhangs. First, cells were embedded in an agarose plug and lysed. Second, the unpaired 3′ overhangs of DSBs were ligated to ssDNA adaptors by Adaptase. Third, after denaturation and sonication, the DEtail-seq library was prepared following an ssDNA library protocol.

ratio will also be lower without treatment at high temperature. Adaptase gives DEtail-seq low bias for ssDNA end ligation (Figure 1), different from other ligation strategies using splinter with random bases (Gansauge et al., 2017) (used in GLOE-seq). However, the resolution of DEtail-seq is speculated to be influenced by the Y-tail, which might confuse the tails and adjacent Y bases downstream of the 3′ DSB ends.

DEtail-seq detects the 3′ ends of DSBs ultra-efficiently and precisely

To verify the reliability and sensitivity of the DEtail-seq method, we generated several types of artificial DNA breaks by digesting pUC19 plasmids with different restriction enzymes (Figure 2A). *Pst*I digestion can generate 4 nt 3′ DNA overhangs (Figure 2A), which should be detectable by DEtail-seq. The cleavage site of *Pst*I in the pUC19 plasmid used in this study is immediately upstream of a G base (Figure 2A). After *Pst*I digestion and DEtail-seq library construction, the library was subjected to both Sanger sequencing and next-generation sequencing (NGS). The results from Sanger sequencing clearly showed the expected features of the library, which included five components: P5 adaptor (green), *Pst*I upstream DNA (grey), an A base at the 3′ end that is produced by *Pst*I (red), a tail that is randomly generated during library construction (blue), and P7 adaptor (orange) (Figure 2B). Analysis of NGS data also showed that reads

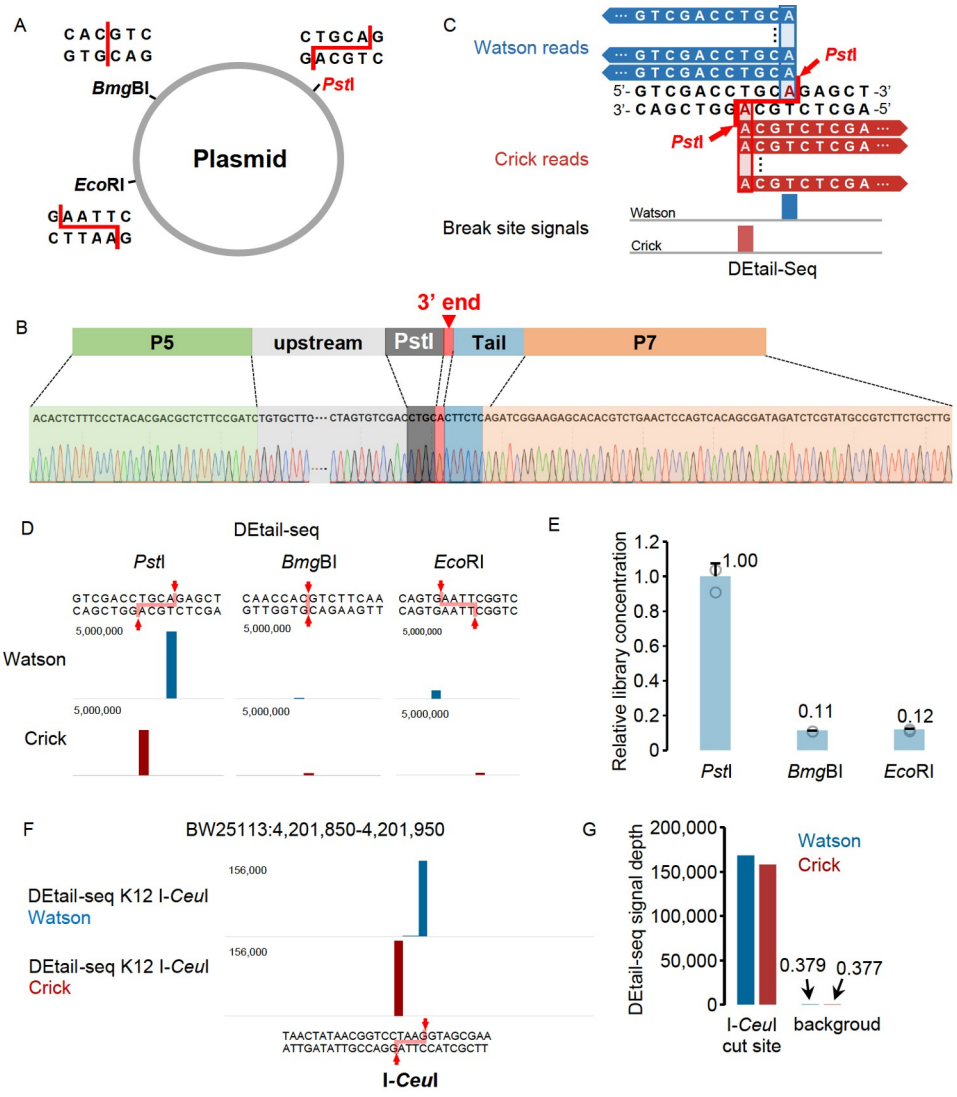

Figure 2 DEtail-seq profiles ends of 3′ overhangs of DSB immediately upstream of R bases with single-nucleotide resolution. A, Sketch showing the plasmid and restriction enzyme cleavage sites used in this study. B, Structure of the DEtail-seq library of artificial DSB from restriction enzyme digestion (upper), demonstrated by Sanger sequencing result (lower). Green, P5 adaptor. Grey, nucleotides in the upstream of 3′ overhangs. Black, *Pst*I recognizing sequence. Red, the last nucleotide at the 3′ overhang. Blue, library tail with Y sequences generated by Adaptase. Orange, P7 adaptor. C, Genome browser profiles of artificial DSB generated by *PstI*. Upper, sequence of the plasmid around the *PstI* cleavage site (arrow) and DEtail-seq Watson (blue) or Crick (red) reads. Lower, processed DEtail-seq signal showing the nucleotides at the 3′ overhangs of DSB. D, Snapshots showing the Watson and Crick DEtail-seq signals at *Pst*I, *Bmg*BI, or *Eco*RI cleavage site. Libraries were generated using plasmids treated with enzyme cocktails (*Pst*I, *Bmg*BI, and *Eco*RI). *Y*-axis, count of reads. E, The relative concentration of DEtail-seq libraries performed on the plasmids treated by *Pst*I, *Bmg*BI, or *Eco*RI respectively, measured by qPCR. F, Snapshot showing the Watson and Crick DEtail-seq signals at one of the I-*Ceu*I cleavage sites in the *E. coli* genome. Libraries were generated using *E. coli* K-12 strain treated with I-*Ceu*I enzyme. *Y*-axis, count of reads. G, Histogram showing the average depth of DEtail-seq signal at all seven I-*Ceu*I cleavage sites or background. Depth was calculated by dividing the count of the reads by the length. The background is the other region of the *E. coli* genome in the flank of I-*Ceu*I cleavage sites.

aligned to the DNA Watson strand or the Crick strand started with an "A" base, consistent with the *Pst*I digestion results (Figure 2C). These data show that DEtail-seq can characterize DNA 3′ end immediately upstream of G or A bases at single-nucleotide resolution. In addition to DSBs with 3′ overhangs, we also used *Bmg*BI and *Eco*RI to generate DSBs with blunt and 5′ overhangs (Figure 2A). As shown by the results in Figure 2D and E, the signal on the *Bmg*BI or *Eco*RI cleavage site was much lower than that on the *Pst*I cleavage site (approximately 10-fold lower), suggesting the ligation efficiency of Adaptase on the paired 3′ end was much weaker than that of the unpaired 3′ end. Notably, the single-nucleotide resolution of DEtail-seq data was also shown at *Bmg*BI and *Eco*RI cleavage sites, which were also immediately upstream of the R bases (Figure 2D). Together, these results indicate that DEtail-seq can detect free DNA 3′ overhangs directly in a high-efficiency manner at single-nucleotide resolution as long as the end is followed by an R

base.

Next, we quantified DSB levels using DEtail-seq by introducing artificial DSB ends into *E. coli* gDNA. We used I-*Ceu*I digested *E. coli* gDNA as substrates for DEtail-seq experiment, as only seven cleavage sites of I-*Ceu*I in the *E. coli* genome (~1.5 DSB per Mb). We found that DEtail-seq showed high accuracy and sensitivity: first, it accurately locates all seven cleavage sites, and second, the signal depth at the cleavage sites far exceeds the background (~40,000 times) (Figure 2F and G; Figure S1C in Supporting Information). Then, we quantified DSB levels using DEtail-seq in budding yeast. Lysed samples from wild-type budding yeast (diploid SK1 strain) were treated with *Asi*SI, a restriction enzyme that generates 3′ overhangs with an adjacent downstream C base (Figure S2A in Supporting Information). Samples were then processed by DEtail-seq library preparation, sequencing, and data analysis. We found that DEtail-seq signals were strongly enriched at most of the 38 *Asi*SI cleavage sites in the yeast genome, and that the density and aligned reads were similar at each *Asi*SI cleavage site, while the background from DEtail-seq was undetectably low, supporting the reliability and sensitivity of DEtail-seq in mapping DNA breaks (Figure S2B and C in Supporting Information). As speculated above, a +1 bp offset of the DEtail-seq signal is found at each *Asi*SI cleavage site because the adjacent downstream base is C (Figure S2D in Supporting Information). Although DEtail-seq could not expeditiously detect DSBs with blunt or 5′ overhangs, and slight offsets of the end location signals are inevitable at a part of break sites which downstream adjacent bases are Y, it still be a suitable method to precisely detect meiotic DSBs with typical 3′ overhang at near single-nucleotide resolution (Figure S1A in Supporting Information).

Profile meiotic DSBs in budding yeast

Meiotic DSBs were strongly induced at four hours in synchronous meiosis and enhanced when Dmc1 was knocked out, but blocked when Spo11 was depleted (Mimitou et al., 2017). To evaluate the performance of DEtail-seq compared with other methods for yeast meiotic DSB detection (Mimitou et al., 2017; Pan et al., 2011), we performed DEtail-seq on wild-type SK1 yeast cells and Dmc1- or Spo11-knockout lines before meiosis (0h), and during meiosis when meiotic DSBs were strongly induced (4h). In line with the known progression of meiotic DSBs (Mimitou et al., 2017), WT 0h exhibited very weak or undetectable signals, demonstrating that the background from DEtail-seq was undetectably low (Figure 3A–C). At four hours after entering meiosis, the DNA break signals in WT cells were strongly enriched at Spo11 cleavage sites (Figure 3A–C). The S1-seq signals (Figure 3A), which indicate DSB resection endpoints (Mimitou et al., 2017), were distributed outside the signals from DEtail-seq meiotic DSB resection (Figure 3A). The enrichment of DEtail-seq for Spo11 binding sites disappeared in *spo11Δ* 4h, and the level of DEtail-seq signals in *dmc1Δ* 4h was significantly higher than that in WT 4h (Figure 3A–C; Figure S3A–C in Supporting Information), consistent with previous results (Mimitou et al., 2017). Therefore, our findings suggest that DEtail-seq could precisely profile meiotic DSBs in budding yeast.

As meiotic DSBs have been previously profiled by Spo11-oligo-seq in yeast (Pan et al., 2011), we next compared these published data with our DEtail-seq data. As shown in Figure S1A in Supporting Information, two meiotic DSB ends are generated by Spo11, and the density of signals from these two DNA ends should be equivalent to each other, which could be used as an ideal evaluation criterion for data analysis. After reanalysis, we found that around Spo11-oligo-seq Watson hotspots, Spo11-oligo-seq signals from the Crick strand were significantly lower than those from the Watson strand (Figure 3C and D; Figure S3B in Supporting Information), and vice versa (Figure S3B and C in Supporting Information). However, the DEtail-seq data showed that the levels of the break signals from the Watson strand and the Crick strand around Spo11-oligo-seq Watson hotspots (Figure 3C and D; Figure S3B in Supporting Information) or Crick hotspots (Figure S3B and C in Supporting Information) were almost equivalent to each other, which is presumably due to the high efficiency of DEtail-seq. These results suggest that DEtail-seq could detect meiotic DSBs with low bias and high sensitivity.

Two other methods, TrAEL-seq and CC-seq, were also used to profile Spo11 cleavage sites in budding yeast (Gittens et al., 2019; Kara et al., 2021). Thus, we then compared the meiotic DSB data generated by using TrAEL-seq, CC-seq, and DEtail-seq, with the Spo11-oligo-seq hot spots chosen as the referential cleavage sites. The meta-analysis results showed that all signals from these three methods were enriched around Spo11 cleavage sites (Figure 3D). The offset range of TrAEL-seq, CC-seq, or DEtail-seq was −2 to +2 bp, 0 to +2 bp, or −2 to +1 bp, respectively (Figure 3D). This result demonstrated that all these three methods, including DEtail-seq, can be used to map Spo11 cleavage sites with extra high resolution.

Similar to previous results that have shown meiotic DSBs to be enriched in nucleosome-depleted regions (NDR) including promoters (Pan et al., 2011), we found that the DEtail-seq signal was also enriched upstream of the TSS (transcription start site) and downstream of the TTS (transcription termination site) in WT 4h (Figure 3E). Interestingly, the enrichment of the DEtail-seq signal upstream of the TSSs was much higher than that downstream of the TTSs (Figure 3E). Next, we wondered if there were any correlations between gene expression and meiotic DSB formation. Thus, we analyzed the levels of DSB signals within the

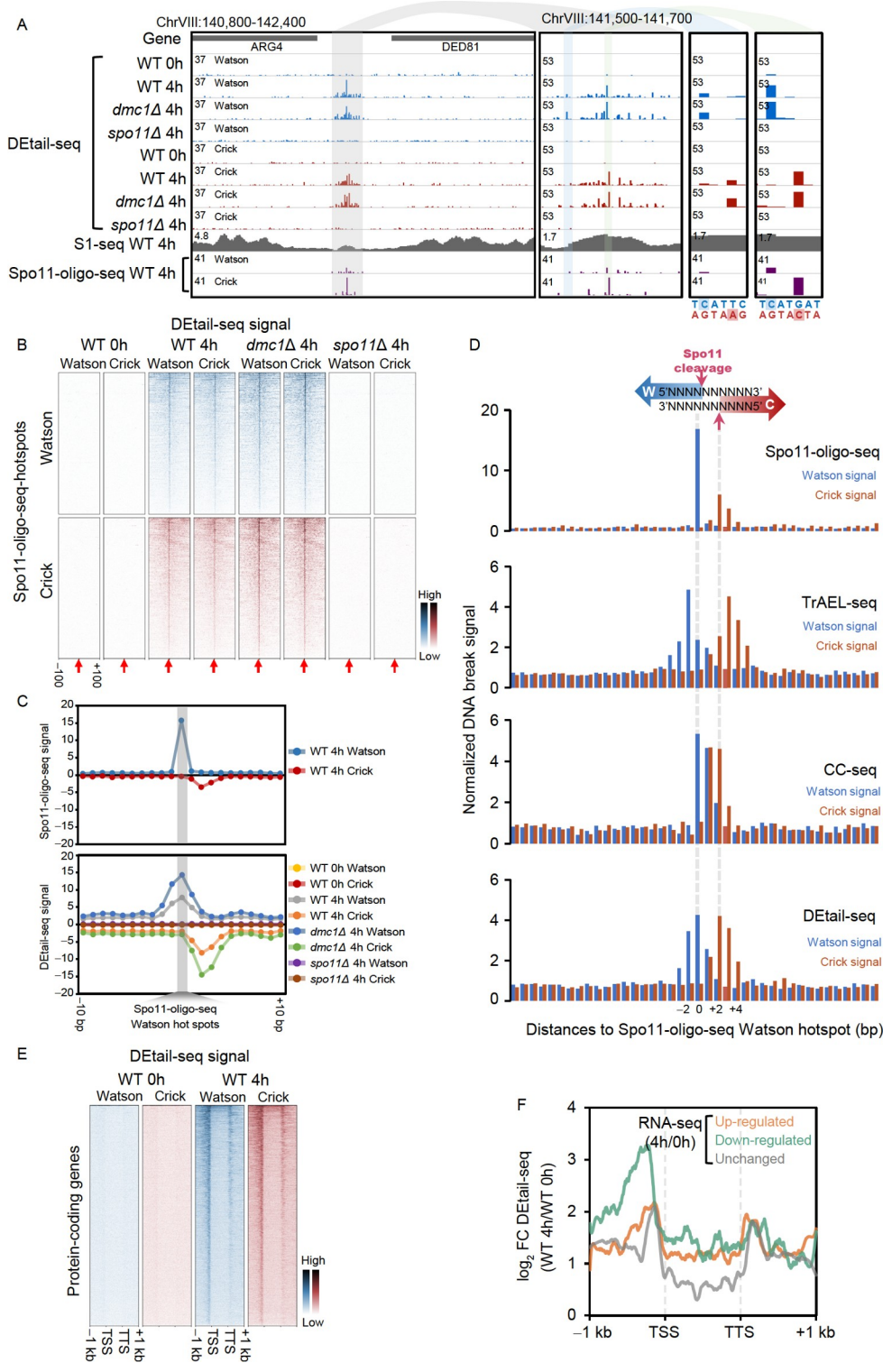

Figure 3 DEtail-seq profiled meiotic DNA breaks in budding yeast. A, Representative snapshots of Watson (blue) and Crick (red) DEtail-seq signals in WT 0h , WT 4h, *dmc1Δ* 4h, and *spo11Δ* 4h, and S1-seq (grey, GSM2263208 from GSE85253) and Spo11-oligo-seq (purple, GSE26449) signals in WT 4h. *Y*-axis, normalized DNA break signal (see Methods). B, Heatmaps showing DEtail-seq Watson/Crick signals in WT 0h, WT 4h, *dmc1Δ* 4h, and *spo11Δ* 4h around Spo11-oligo-seq Watson/Crick hotspots (GSE26449). C, Metaplots showing Spo11-oligo-seq Watson/Crick signals in WT 4h (upper) and DEtail-seq Watson/Crick signals in WT 0h, WT 4h, *dmc1Δ* 4h, and *spo11Δ* 4h (lower) around Spo11-oligo-seq Watson/Crick hotspots (GSE26449). *Y*-axis, normalized DNA break signal. Signal on Crick strand was shown as negative value. D, Metaplots showing Watson and Crick signals of Meiotic DSBs profiled by Spo11-oligo-seq, TrAEL-seq (GSE154811), CC-seq (GSE137685), and DEtail-seq around Spo11-oligo-seq Watson hotspots (GSE26449). *Y*-axis, normalized DNA break signal. E, Heatmaps of DEtail-seq Watson/Crick signals in WT 0h and WT 4h around protein-coding genes. F, Metaplot of $\log_2$ FC (fold change) of DEtail-seq signals (WT 4h/WT 0h) around genes with different expression changes (RNA-seq WT 4h/WT 0h, from GSE34082). Up-regulated genes, RNA-seq $\log_{10}$FC>1. Down-regulated genes, $\log_{10}$FC<−1. Unchanged genes, −0.1<$\log_{10}$FC<0.1.

context of gene expression during meiosis, using published RNA-seq data of WT 4h and WT 0h (Brar et al., 2012). The result showed that the level of meiotic DSB upstream of TSSs of down-regulated genes was much higher than that of unchanged genes or up-regulated genes (Figure 3F). No dramatic differences were observed in DSBs located downstream of the TTSs when comparing the up-, down-regulated and unchanged genes in WT 4h and WT 0h (Figure 3F). Genes in the budding yeast genome are arranged frequently in head-to-tail orientation (Dujon, 1996). To exclude the possibility of the higher DEtail-seq signals around TSSs of down-regulated genes generated from the nearby TTSs of upstream genes, we analyzed the distances between TSSs and their nearest upstream TTSs. The statistical result showed that the distances between TSSs of down-regulated genes and their nearest TTSs were slightly longer than those of unchanged genes or up-regulated genes (Figure S3D in Supporting Information), which excluded the possibility of nearby TTSs contributing higher signals. A similar pattern was also found in the Spo11-oligo-seq results (Figure S3E in Supporting Information). These results suggested a negative correlation between meiotic DSB formation and gene expression.

Many studies focus on the 5′ end resection process (Mimitou and Symington, 2009), while few studies directly delineate the 3′ end of meiotic break. Although southern blotting can show that meiotic DSB 3′ overhangs are stable (Zhu et al., 2008), there is no direct report showing the NGS results of 3′ end location. As DEtail-seq relies on an ssDNA ligation technique with the 3′ end, the method provides an efficient way to directly examine the 3′ end of the DNA break. When characterizing the precise 3′ overhang locations at the genome-wide level during meiosis, we found that DEtail-seq signals from both WT 4h and *dmc1Δ* 4h were dramatically enriched for Spo11 hotspots without significant location shift (Figure 3A–D; Figure S3A–C in Supporting Information). These results disclose that, unlike 5′ end resection, the 3′ overhangs of meiotic DSBs were not processed during or post-break formation, providing direct evidence to further confirm that the 3′ overhangs of meiotic DSBs were stable (White and Haber, 1990; Zhu et al., 2008).

Taken together, our results suggest that DEtail-seq can precisely characterize the 3′ ends of Spo11-induced meiotic DSBs in yeast with low bias, high resolution, and high sensitivity. Furthermore, the integrality of 3′ overhangs during meiotic DSBs was revealed by DEtail-seq data (Figure 3A–D), filling the gap of the feature of 3′ overhangs during meiotic DSB processing.

DSB hotspot distribution in the mouse genome during meiosis

Next, we used DEtail-seq to investigate the profiles of mammalian meiotic DNA breaks in mouse germ cells at leptotene/zygotene stage (L/Z) and pachytene stage (Pac) with Sertoli cells (SC) as the control (Figure S4A in Supporting Information). First, we compared the DEtail-seq data with published END-seq data on *Asi*SI-induced artificial DSBs (Canela et al., 2016). Although 216 *Asi*SI cleavage sites could be commonly detected by both methods, 30 sites were detected by END-seq only, while 179 sites were uniquely detected by DEtail-seq (no END-seq read was detected on 129 of them) (Figure 4A; Figures S5A and S4B in Supporting Information). This result confirms the higher sensitivity of DEtail-seq.

To further validate the DEtail-seq data, we performed meta-analysis of DSB breaks obtained from DEtail-seq data at the L/Z stage together with published Spo11-oligo-seq data in mouse testis (Lange et al., 2016). The heatmaps showed dramatic enrichment of Watson and Crick break sites in Spo11-oligo-seq signals (Figure S4C in Supporting Information), demonstrating that DEtail-seq had successfully profiled SPO11-induced meiotic DSBs in mice. PRDM9 is a histone methyltransferase that can be used to determine the locations of meiotic recombination hotspots (Baudat et al., 2010). Therefore, we further analyzed the co-localization between DEtail-seq mapped meiotic DSBs and PRDM9 binding sites (Hinch et al., 2020). The results showed dramatic enrichment of DEtail-seq signal around PRDM9 binding sites at the L/Z stage, while such enrichment was not found in SC (Figure 4B; Figure S5B in Supporting Information), indicating the background from DEtail-seq was low. A strong correlation has also been reported between meiotic DSB and H3K4me3 modifications in mice (Lange et al., 2016). In line with this, DEtail-seq data found DNA breaks at the L/Z stage were enriched at H3K4me3 peaks at the leptotene stage (Chen et al., 2020) (Figure 4C), further confirming that DEtail-seq is an ideal method for meiotic DSB profiling in mammals.

We then characterized the break sites for each sample. Plots showed a number of DNA break sites (2,608 Watson hotspots and 2,540 Crick hotspots) were uniquely detected in L/Z, compared with those in SC or Pac (Figure 4D). This result demonstrated that the meiotic DNA breaks were strongly induced at the L/Z stage.

As meiotic DSBs were found to be prevalent upstream of the TSS in budding yeast (Figure 3E and F), we wondered whether a similar pattern would be detected in mice. Metaplots showed that DNA breaks in L/Z and SC were equally enriched in whole transcription regions, from TSS to TTS (Figure S5C in Supporting Information). However, in Pac, additional enrichments were found around the TSS of protein-coding genes (Figure S5C in Supporting Information). To find out whether the enrichment around TSS in Pac was due to single-strand breaks (SSBs) or DSBs, we analyzed the distribution of breaks in Watson and Crick strands. The result

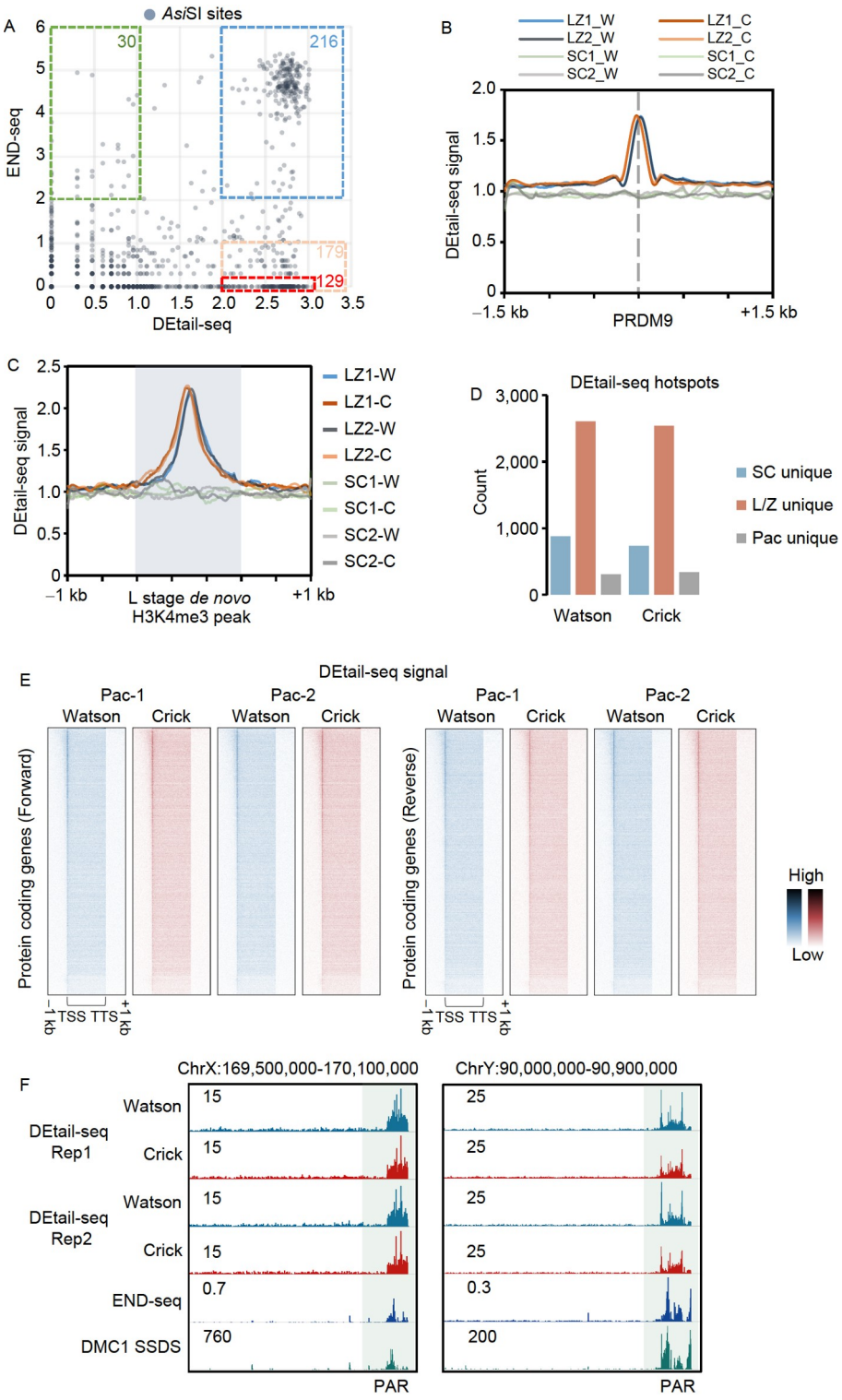

Figure 4 DEtail-seq profiled DNA breaks in mouse during meiosis. A, Reads ($\log_{10}$) from END-seq (SRX1867541 from SRP076804) and DEtail-seq data mapped at *Asi*SI cleavage sites. Blue frame, common positive sites, defined as $\log_{10}$Reads>2 in both END-seq and DEtail-seq. Green frame, END-seq unique positive sites, defined as $\log_{10}$Reads>2 in END-seq and <1 in DEtail-seq. Light red frame, DEtail-seq unique positive sites, defined as $\log_{10}$Reads>2 in DEtail-seq and <1 in END-seq. Dark red frame, DEtail-seq strong unique positive sites, defined as $\log_{10}$Reads>2 in DEtail-seq and =0 in END-seq. B, Metaplot of normalized DEtail-seq signals (W, Watson; C, Crick) from L/Z or SC around PRDM9 binding sites. C, Metaplot of Watson or Crick DEtail-seq signals (RPGC) in two biological replicates of SC or L/Z on leptotene stage *de novo* H3K4me3 peaks (GSE132446). *Y*-axis, normalized DEtail-seq signal (see Methods). D, Plot showing unique DEtail-seq Watson (left) or Crick (right) hotspots during SC, L/Z, or Pac stage. E, Heatmaps showing DEtail-seq signals (RPGC) in Pac around forward (or positive, left panel) and reverse (or negative, right panel) protein-coding genes in the mouse genome. F, Snapshot showing the DEtail-seq, END-seq (GSM4122777) (Paiano et al., 2020), and DMC1 SSDS (GSM3351206) (Boekhout et al., 2019) signal during L/Z stage in PAR regions.

showed that enrichments of DEtail-seq signals were equally distributed in both strands (Figure 4E), suggesting that DSBs could be detected around the TSS. As meiotic DSBs in mammalian chromosome X and Y were enriched in PAR (pseudoautosomal region) (Brick et al., 2018), we wanted to know if this phenomenon could be reflected via DEtail-seq. The results showed that DEtail-seq signal was strongly enriched in PAR, consistent with END-seq (Paiano et al., 2020) and SSDS results (Boekhout et al., 2019) (Figure 4F). These results suggested that both conservation and specificity of DNA breaks exist between budding yeast and mouse during meiosis, and the DNA breaks enriched around TSS in the mouse genome during pachytene stage might be due to increased chromatin accessibility or RNA transcription process.

DSB hotspot distribution in the human genome during meiosis

Unlike in yeast or mouse, meiotic DSB sites have not been fully characterized in human, most possibly due to the unavailability of human SPO11 antibody. Currently, the only published genome-wide data for human meiotic DSBs has been indirectly generated from DMC1 SSDS (single-stranded DNA sequencing) (Pratto et al., 2014), which detects ssDNA overhangs around DSB sites, but cannot detect DSB sites directly (Figure S1A in Supporting Information). To fill this gap, we performed DEtail-seq using testicular samples from two obstructive azoospermia (OA) patients who underwent vasoepididymostomy (see Methods). Data analysis revealed approximately 200,000 hotspots from each replicate (Table S2 in Supporting Information), which is approximately 6 to 10 times than those identified by SSDS (Pratto et al., 2014), indicating a higher sensitivity and resolution of DEtail-seq.

Given the limitation of the purity of leptotene and zygotene cells collected from human testis and that many dead germ cells could also be observed in normal spermatogenesis (Liu et al., 2017; Shaha et al., 2010; Weinbauer et al., 2001; Young et al., 2001), some non-meiotic DSBs may be observed from the testicular DEtail-seq results. To exclude these background noises, the conservation of DEtail-seq hotspots between two patients was analyzed, which were termed shared hotspots (Figure S6A in Supporting Information). We found most hotspots were shared by two patients (Figure S6A in Supporting Information). To detect whether the shared hotspot between two patients could reflect the meiotic DSBs, we analyzed the overlap of the shared hotspots obtained from DEtail-seq and previously published human DMC1 SSDS data (Pratto et al., 2014). The results showed that DMC1 SSDS signals were enriched around break sites obtained from DEtail-seq (Figure 5A; Figure S7A in Supporting Information), while such enrichment was not found in unique DEtail-seq signals in two patients (Figure 5A), supporting the reliability of DEtail-seq for human meiotic DSB detection.

Since the correlation between meiotic DSB and H3K4me3 was found in mouse, we wondered if this correlation could potentially occur in human. Meta-analysis showed that DEtail-seq signal was strongly enriched at H3K4me3 peaks (Bae and Lesch, 2020) (Figure S6B and C in Supporting Information), which is similar to that in the mouse. We also investigated the correlation between meiotic DNA breaks and other histone-active markers, such as H3K4me1 and H3K27ac, which have been reported to be associated with meiotic DSBs in mouse (Chen et al., 2020). The results showed that DNA breaks were slightly enriched at H3K4me1 and H3K27ac peaks (Figure S6B, E, and F in Supporting Information). By analyzing the genome distribution, we found that DNA breaks were enriched in the intergenic regions (Figure S6D in Supporting Information), in line with the enrichment of active marks H3K4me3, H3K4me1, and H3K27ac. As H3K4me1 and H3K27ac are associated with enhancers (Creyghton et al., 2010), we wondered whether enhancer regions in general contained more DNA breaks than other intergenic regions. The results showed that DNA breaks were strongly enriched at 397 enhancer regions that were marked by intergenic H3K27ac and H3K4me1 peaks (Figure 5B; Figure S7B in Supporting Information). Heatmaps and snapshots showed that DEtail-seq signals were enriched at the borders of enhancer regions (Figure S7B and C in Supporting Information). A similar pattern was also found in the SSDS results (Figure S7B in Supporting Information), indicating that meiotic DSBs in human are associated with some enhancer regions.

Common fragile sites (CFSs) are specific chromosomal loci that exhibit increased tendencies for chromosome breakage under replication stress. To investigate the relationship between CFSs and meiotic DNA breaks, we analyzed chromosome localizations of DEtail-seq signals and mitotic DNA synthesis sites (MDSs), which were defined by mitotic DNA synthesis sequencing (MiDASeq) and covered most CFSs (Ji et al., 2020). To quantify the correlation between MDSs and DNA breaks profiled by DEtail-seq, meta-analysis and permutation tests were performed. Metaplots and snapshots showed dramatic enrichments of MDSs at DEtail-seq hotspots (Figure 5C; Figure S8A in Supporting Information). These results demonstrated a positive correlation between meiotic and mitotic DNA breaks, suggesting that a common mechanism might regulate these two different types of breaks.

Loop anchors bound by cohesin and CCCTC-binding factor (CTCF) have been reported to serve as fragile sites that generate DSBs (Canela et al., 2017), and a class of CTCF-bound anchor points overlaps meiotic recombination hotspots near the core PRDM9-binding motif (Kaiser and Semple, 2018). Thus, we investigated the correlation

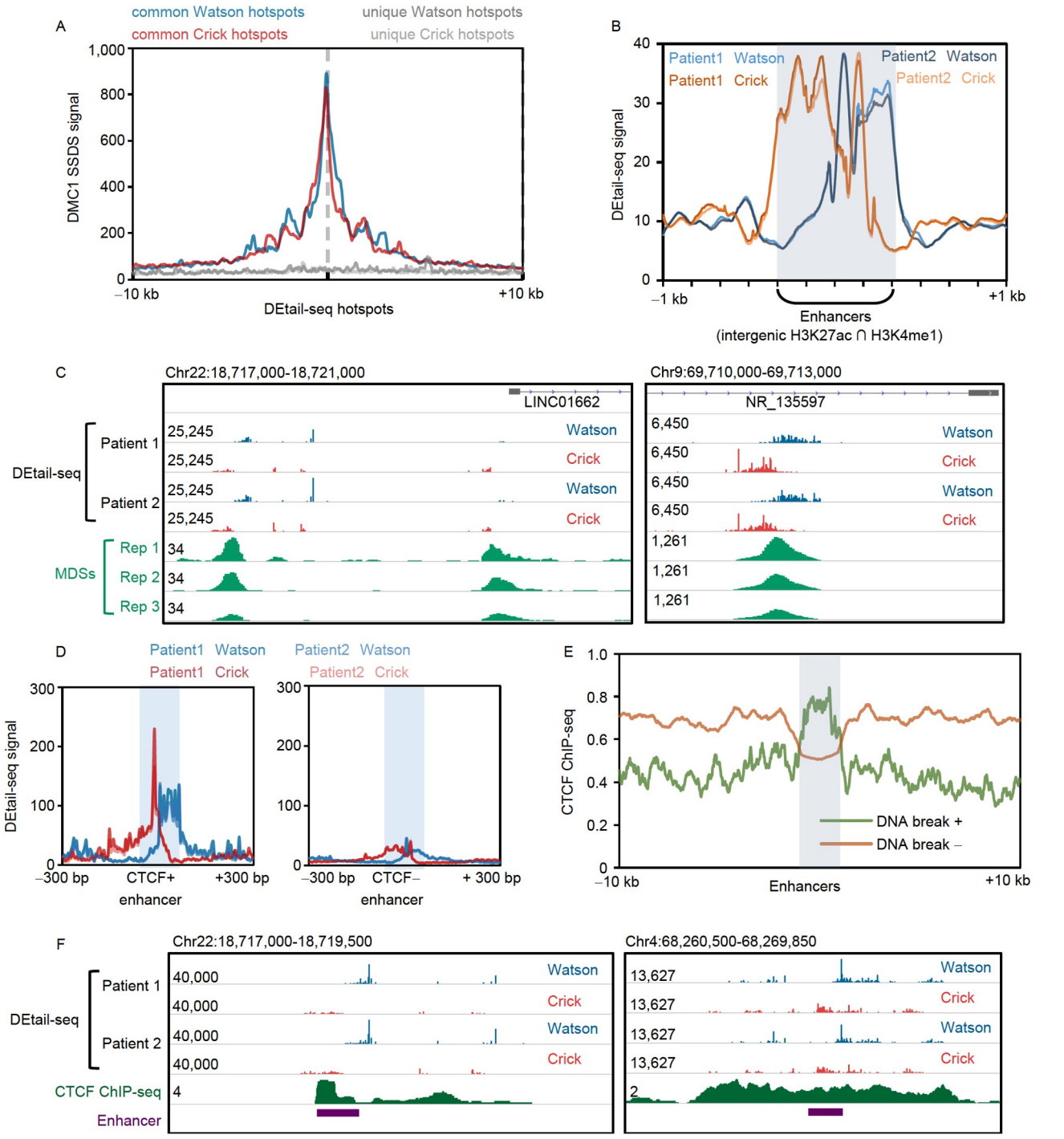

Figure 5 DEtail-seq profiled DNA breaks in human during meiosis. A, Metaplot of DMC1-SSDS signals (GSM1447325 from GSE59836) around shared or unique Watson/Crick DEtail-seq hotspots. Shared means the hotspots shared between two patients, while unique means hotspots with the shortest distance between two patients are longer than 100 bp. B, Metaplot of Watson or Crick DEtail-seq signals on enhancers (intergenic intersection of H3K4me1 and H3K27ac peaks, from GSE145225) in human spermatogenic cells. *Y*-axis, normalized DEtail-seq signal. C, IGV snapshots showing the DEtail-seq and MDS signals in the human genome. *Y*-axis, normalized DEtail-seq or MDS signal. D, Metaplot of Watson or Crick DEtail-seq signals on enhancers overlapping without CTCF peaks (GSE105739, CTCF− enhancer), or with CTCF peaks (CTCF+ enhancer). *Y*-axis, normalized DEtail-seq signal. E, Metaplot of CTCF ChIP-seq signal on enhancers with or without DEtail-seq hotspots, represented by "DNA break +" or "DNA break −" respectively. F, IGV snapshots showing the DEtail-seq and CTCF ChIP-seq signals around enhancers. *Y*-axis, normalized DEtail-seq or ChIP-seq signal.

between CTCF binding sites, common fragile sites and/or DNA breaks during meiosis. By conjoint analysis with published CTCF ChIP-seq data from ENCODE (The Encyclopedia of DNA Elements) (Consortium, 2012), our permutation tests revealed that DEtail-seq hotspots were indeed enriched within CTCF binding regions (Figure 5D;

Figure S8B in Supporting Information). Therefore, chromosome loop anchors associated with genome organization may also drive DSB formation. Given that CTCF is highly associated with enhancers, the enriched signals of DEtail-seq at the borders of enhancer regions may also be related to CFSs. To examine this, we sorted enhancers into two groups, "CTCF+" (with CTCF binding), and "CTCF−" (without CTCF binding). We found a weaker enrichment of the DEtail-seq signal in CTCF− enhancers and a stronger enrichment in CTCF+ enhancers (Figure 5D; Figure S8C in Supporting Information). Next, meta-analysis showed a dramatic enrichment of CTCF in enhancers with DNA breaks and a decline of CTCF in enhancers without DNA breaks (Figure 5E and F; Figure S8B in Supporting Information). These results demonstrate a positive correlation between CFSs, CTCF, enhancers, and DNA breaks during meiosis in human.

Conservation of DSB distribution in mammals during meiosis

Considering the high similarity between the human and mouse genomes, we were then interested in similarities and differences on their meiotic DSB distributions. Given that co-localization of DEtail-seq signal and CFSs was found in human, we wondered if this co-localization was conserved in other mammals. Thus, we used CFSs in mouse B cells that have previously been identified through HU-EdU-seq (Tubbs et al., 2018) to investigate the localization of DEtail-seq signals and CFSs in mouse. Our metaplots and snapshots showed a strong enrichment of CFSs signals around L/Z *de novo* break sites or Pac break sites profiled by DEtail-seq (Figure S9A and B in Supporting Information), which is consistent with what we found in human. Further analysis also showed the co-localization between CTCF (Luo et al., 2020) and DEtail-seq-profiled DNA break sites in mouse (Figure S9C in Supporting Information), consistent with those in human (Figure 5D–F). These results suggest the conservation of some common mechanisms that affect meiotic and mitotic DNA breaks in mammals.

As shown in Figure 5B, a strong enrichment of DSBs was identified around human enhancers during meiosis. Thus, we investigated whether this correlation was also conserved in other mammals. The results from a similar analysis in mouse showed that DSBs in Pac were also enriched around enhancer regions, although much weaker than enrichment in L/Z (Figure S9D in Supporting Information). These results imply that the distribution pattern of some DSB hotspots is diverse between mouse and human.

DISCUSSION

In the last couple of decades, great achievements have been made in DNA breaks and repair (d'Adda di Fagagna, 2008; Panier and Durocher, 2013; Scully et al., 2019). As a crucial step for successful meiosis, meiotic DSBs are induced at specific regions across the genome by Spo11 and other co-factors (Keeney, 2008; Keeney et al., 2014). To fully understand meiotic DSBs, several methods have been developed for meiotic DSB mapping, such as S1-seq (Mimitou et al., 2017), END-seq (Canela et al., 2016), DMC1-SSDS (Pratto et al., 2014), Spo11-oligo-seq (Pan et al., 2011), TrAEL-seq (Kara et al., 2021), and CC-seq (Gittens et al., 2019). Here, we developed DEtail-seq to profile meiotic DNA breaks. DEtail-seq has the ability to map the 3′ ends of DSBs directly, thus it has a high potential to be used not only in meiotic DSB research but also in many other fields to detect DSB or SSB sites. It is especially suitable for studying DSBs or other DNA structures that bear 3′ overhangs (Table S1 in Supporting Information).

By using DEtail-seq, we profiled Spo11-induced meiotic DSBs with near single-nucleotide resolution and provided an ideal method to study meiotic DSBs in various organisms, circumventing the limitation of needing an anti-Spo11 antibody or transgenic line to generate tagged Spo11 proteins. DEtail-seq provides high-quality data, as the two ends of one Spo11 hotspot can be equally detected (Figure 3C). DEtail-seq also showed higher sensitivity compared to END-seq (Figure 4A). DSBs induced by Spo11 could be characterized at near single-nucleotide resolution (Figure 3A and C; Figure S3A and C in Supporting Information), which confirms the reliability and detail of the DEtail-seq data, making the method applicable for genome-wide meiotic DSB mapping. In addition, compared with another 3′ end profiling method TrAEL-seq, DEtail-seq has a similar resolution and much simpler procedure, making DEtail-seq more usable, which will strongly support researches on DNA breaks in a variety of organisms. Although background noise could not be completely excluded in DEtail-seq, the I-*Ceu*I assay in *E. coli* showed that the background noise in DEtail-seq was very low (Figure 2G). Therefore, DEtail-seq is a low-noise DNA break profiling method with high reliability and sensitivity.

In mouse, during meiosis, RNA expression is suppressed at the L/Z stage and then reactivated at the Pac stage (Chen et al., 2018). From DEtail-seq data, we found that mouse germ cells at the Pac stage showed significant enrichment of DSBs around the TSS compared with germ cells at the L/Z stage, which might be due to the activation of gene expression during the Pac stage. This further supported that DEtail-seq could be used to reveal a wide range of DNA breaks.

Furthermore, we performed DEtail-seq using testicular samples and directly profiled the high-resolution DSB map during human meiosis, filling the gap in meiotic DSB study in human. Due to the limitation of the purity of leptotene and zygotene cells collected from human testis, and the fact that

many dead germ cells could also be observed in normal spermatogenesis (Liu et al., 2017; Shaha et al., 2010; Weinbauer et al., 2001; Young et al., 2001), we could not completely exclude non-meiotic DSBs. On the other hand, the different genetic backgrounds of these tested patients may also be associated with the distinguished distribution of the meiotic DSB hotspots (Pratto et al., 2014). However, the shared DEtail-seq signals between two patients were highly associated with DMC1 SSDS data (Figure S6A in Supporting Information), indicating that DEtail-seq results from human testicular samples contain meiotic DSBs.

By analyzing DEtail-seq data, we found strong co-localization between meiotic DSBs and CFSs in human. The strong enrichment of CFS signals around L/Z *de novo* break sites profiled by DEtail-seq could also be detected in mouse (Figure 5C). Further analysis showed that the co-localization between CTCF and meiotic DNA break sites profiled by DEtail-seq in mouse (Figure S9C in Supporting Information) was consistent with that in human (Figures 5D and 4E). DSB hotspot localization has been proposed to be governed by higher-order chromosome architecture (Borde and de Massy, 2013). In mitotic cells, active TOP2B is generally bound on both sides of a CTCF loop boundary and could be responsible for deletions or insertions that disrupt boundary elements (Canela et al., 2017). During meiosis, DSBs are known to occur primarily within the chromatin loop sequences that are tethered to the chromosome axis by some as yet unidentified recombination-promoting factors, which has been termed as the "tethered-loop/axis complex" model (Blat et al., 2002). In budding yeast, Spp1, a member of the Set1 Complex, can recognize H3K4me2/3 near gene promoters in chromatin loops and interact transiently with Spo11 accessory protein Mer2 on the axis to promote meiotic DSB formation (Borde and de Massy, 2013). These results suggest the conservation in regulating meiotic and mitotic DNA breaks in mammals, which might be due to common regulatory mechanism(s), such as the overall structure of chromosomes.

Surprisingly, we found that DEtail-seq signals were enriched around enhancers and CTCF binding sites during human meiosis (Figure 5B–E). This suggested a possibility that meiotic DNA breaks might play an unexpected role in the construction or function of enhancers by regulating CTCF binding. Interestingly, this observation was not clear in mouse meiosis, suggesting that this correlation might be diverse in different species.

MATERIALS AND METHODS

Yeast sporulation

Yeast sporulation was performed as previously described (Wen et al., 2016). In brief, yeast cells (SK1 strains) were cultured for 24 h at 30°C in liquid YPD medium. Then, cells were diluted in liquid YPA medium to A_{600}=0.3, and cultured at 30°C for 10 h. Cells were harvested and washed with sterilized water for three times. Synchronized cells were resuspended in sporulation medium (2% potassium acetate, SPM) with A_{600}=1.9 to induce meiosis.

Isolation of mammalian spermatogenic cells

Spermatogenic cells were isolated as previously described (Xu et al., 2016). Testes from adult C57BL/6 male mice were decapsulated and torn into small pieces. The seminiferous tubules were then treated with 1 mg mL^{-1} collagenase (Sigma-Aldrich, USA) and 1 mg mL^{-1} hyaluronidase (Sigma-Aldrich) in 8 mL PBS for 6 min at 37°C with gentle shaking. Pipetted spermatogenic cells were further incubated for 5 min with gentle shaking at 37°C. Spermatogenic cells were washed with PBS and treated with 0.25% Trypsin and 1 mg mL^{-1} DNase I in 15 mL PBS with gentle shaking for 5 min at 37°C. After washing with 0.5% BSA in PBS, spermatogenic cells were filtered through a nylon cell strainer (40 μm) and separated by sedimentation velocity in 2%–4% BSA gradient.

Human adult testis sample preparation

Adult human testis samples were collected from two OA patients who were enrolled from Reproductive Center of Peking University Third Hospital and underwent vasoepididymostomy (Table S3 in Supporting Information). We excluded patients with chromosomal abnormalities, Y-chromosome microdeletions, varicocele, and other known factors associated to male infertility. All donors were informed written consent for this study. The ethical approval for this study was provided by the Reproductive Study Ethics Committee of Peking University Third Hospital (2017SZ-035).

Human adult testes were washed twice in 1× HBSS and torn into small pieces. After digestion with 1 mg mL^{-1} collagenase and 1 mg mL^{-1} hyaluronidase in PBS for 6 min with gentle shaking at 37°C, spermatogenic cells were pipetted and incubated with gentle shaking for 5 min at 37°C. The digestion was stopped by using 10% FBS (Gibco, USA). Spermatogenic cells were filtered through 70 μm (Thermo Fisher Scientific, USA) and 40 μm (Thermo Fisher Scientific) strainers. After pelleting and washing twice with DMEM, spermatogenic cells were placed into culture dishes with F12-DMEM containing 15% FBS (Gibco). After incubation at 34°C and 5% CO_2 for 6 h, spermatogenic cells were harvested by centrifugation at 600×*g* for 5 min. Cells were washed twice with 1× PBS and resuspended in 1× PBS ready for DEtail-seq.

Characterization of Adaptase

In a 20 μL reaction system, 5 pmol 90 nt ssDNA with 8 random nucleotide end (sequence: NNNNNNNNTGGACGAAGACTGACGCTCCTGGATATCTAGACGATATCGATACGATCAGTCGAGCCACCCGAGACTCAGTGAANNNNNNNN) was ligated to the adaptor by Adaptase at 37°C for 1 h. Then, electrophoresis was performed in 15% PAGE gel for ligation efficiency imaging.

For the tailing sequencing experiment, ssDNA with 3 random nucleotide 3′ ends (sequence: CGCAAAGCATCCTCTAAACTTGGCGGAGAGCCCAGCGGACTCCAGCTCGTCGTACGCCACGGAGCCAGTGAGCTCATGGATCCGCGACCCATTTGCTGTCCACCAGTCATGCTAGCCATATGGCTGCCGCGCGGCACCAGGCCGCTGCTGTGATGATGATGATGATGGCTGCTGCCCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTANNN) was used as substrate for library construction according to the manual of Accel-NGS 1S plus DNA library kit (Swift BioSciences, USA) and sequenced on the Illumina NovaSeq 6000 platform (150 bp pair-end reads). Read 2 sequences were used for tailing characterization analysis with text-processing utility in Linux system.

DEtail-seq

Cell embedding and lysis were performed by following the steps described previously (Canela et al., 2016; Mimitou et al., 2017). Briefly, for budding yeast, cells were resuspended in 150 μL 50 mmol L^{-1}, pH 8.0 EDTA buffer. 498 μL melted LMP agarose (1.5% in 50 mmol L^{-1} pH 8.0 EDTA) was incubated with 102 μL Solution 1 (0.1 mol L^{-1} sodium citrate, 1 mol L^{-1} sorbitol, 60 mmol L^{-1} pH 7.0 EDTA, 1 mg mL^{-1} zymolyase 100T and 5% β-mercaptoethanol). The mixture was kept at 40°C and then added to the cells. Resuspended cells were pipetted into plug molds and chilled at 4°C for 30–40 min. The agarose plugs were then cut into small pieces (one agarose plug to 16–24 pieces, 2–3 mg/piece), and all small plugs were soaked in 1 mL of Solution 2 (0.01 mol L^{-1} pH 7.5 Tris-HCl, 0.45 mol L^{-1} pH 8.0 EDTA, 10 μg mL^{-1} RNase A, 7.5% β-mercaptoethanol) and incubated at 37°C for 12 h. Plugs were quickly washed three times with 1 mL TE buffer, 1 min each time, and then washed five times with TE buffer incubated at 37°C, 6 min each time. Then plugs were soaked in 1 ml Solution 3 (0.01 mol L^{-1} pH 7.5 Tris-HCl, 0.25 mol L^{-1} pH 8.0 EDTA, 1% SDS, 1 mg mL^{-1} proteinase K) and incubated at 54°C for 12 h. Plugs were finally washed six times with TE buffer at 37°C for 6 min each time. For artificial DSB by restriction enzyme digestion, small plugs were incubated with 10 U *Asi*SI in 100 μL 1× CutSmart Buffer at 37°C for 12 h. After digestion, the plugs were washed five times with incubation in 1 mL TE buffer at 37°C for 6 min each time.

For mammalian cells, cells were collected and resuspended in 50 μL PBS buffer. 50 μL LMP agarose (2% in PBS) was pre-warmed at 40°C, followed by mixing with the resuspended cells and pipetting into plug molds. After solidification by chilling at 4°C for 30–40 min, the agarose plug was cut into small pieces (one agarose plug to 16–24 pieces, 2–3 mg/piece), and they were added into 1 mL of Solution 3 in 2 mL tubes and incubated at 54°C for 12 h. Plugs were quickly washed three times with 1 mL TE buffer, and then washed five more times with incubation in 1 mL TE buffer at 37°C for 6 min each time. Plugs were added to 1 mL TE with 10 μg mL^{-1} RNase A (final) and incubated at 37°C overnight. Plugs were quickly washed three times with 1 mL TE buffer, and then washed five more times with 1 mL TE buffer incubated at 37°C for 6 min each time. Spike-in was performed by incubating small plugs with 10 U *Asi*SI in 100 μL 1× CutSmart Buffer at 37°C for 12 h. After digestion, the plugs were washed five times with incubation in 1 mL TE buffer at 37°C for 6 min each time.

One small plug was used for library preparation. The small plug was washed in 30 μL 1× Adaptase buffer (Swift BioSciences, Accel-NGS 1S plus DNA library kit; related patent application, US 20200392551A1 and CN 104395480B) three times at room temperature, 15 min each time. The plug was incubated in 30 μL 1× Adaptase buffer with Adaptase at 37°C for 12 h. The reaction solution was removed, and the plug was heated quickly to 98°C for 2 min to deactivate the enzymes. DNA was recovered from the agarose gel by using a DNA purification kit (Magen Biotechnology, Guangzhou, China). The DNA was fragmented to ~250 bp by using the Focused-ultrasonicator (Covaris, LLC. USA). The next steps, including extension, ligation, and indexing PCR, were performed following the protocol of the Accel-NGS 1S plus DNA library kit, completing the DEtail-seq library preparation.

For plasmids, cell embedding and lysis were omitted instead of using direct phenol chloroform extraction. After a 2 h digestion with *Pst*I, *Bmg*BI, or *Eco*RI (10 U each) at 37°C., digested plasmid was extracted with phenol chloroform, and dissolved in 50 μL TE buffer. The plasmid was ligated with Adaptase at 37°C for 1 h, and then deactivated at 98°C for 2 min. DNA ligated with P7 adaptor was fragmented to ~250 bp by using the Focused-ultrasonicator (Covaris S220), and the rest of the steps were performed by following the instruction of the 1S plus DNA library kit.

All sequencing was performed on the Illumina NovaSeq 6000 (150 bp pair-end reads).

DEtail-seq data analysis

First, duplicated reads which had the same sequences for both forward and reverse reads were removed. Our own

scripts were used to remove duplicated reads (https://github.com/PEHGP/DEtail-seq/blob/master/RemoveSamReads.py). Reads were then aligned to the reference genome with Bowtie2 (Langmead and Salzberg, 2012; Langmead et al., 2009) using "–local" settings. For visualization and splitting strands, the aligned read files (BAM) including second read in pairs were converted to first single base bigWig file with 1 bp bins using bamCoverage from deepTools (Ramírez et al., 2014) with "–binSize 1 –Offset 1 –samFlagInclude 128 –filterRNAstrand forward/reverse". The option "–filterRNAstrand forward" means Crick strand and "–filterRNAstrand reverse" means Watson strand. Normalization was performed by deepTools based on RPGC (reads per genomic content, 1× normalization). DEtail-seq hotspots were defined by using our own script (https://github.com/PEHGP/DEtail-seq/blob/master/Hotspotcalling.py). Briefly, the lambda of Poisson distribution was calculated using non-zero base mean value in Bigwig. And the p value of each non-zero base was calculated according to the Poisson. Then we used the Benjamin Hochberg method to correct the p value and to obtain the q value. Finally, we filtered the q value to obtain hotspots, and the hotspots in mitochondrial genome were discarded. The detailed pipeline can be viewed on GitHub at https://github.com/PEHGP/DEtail-seq.

Meta-analysis was performed by using computeMatrix from deepTools (Ramírez et al., 2014). BEDTools (Quinlan and Hall, 2010) and regioneR (Gel et al., 2016) were used to assess the associations between different types of genomic regions. IGV (Integrative Genomics Viewer) was used for snapshots of sequencing signals (Robinson et al., 2011). Seqtk tool (http://github.com/lh3/seqtk) was used for statistical analysis of bases at hotspots.

Re-analysis of published data

For END-seq, FASTQ files were downloaded using fastq-dump through accession numbers SRR3705559 (*Asi*SI *in vitro*) (Canela et al., 2016), SRR10286908 (WT), SRR10286909 (WT), and SRR10286910 (WT) (Paiano et al., 2020). Alignment was performed with Bowtie2 using the mm10 genome as the reference with default parameters. BigWig file conversions and hotspot callings used the same method as DEtail-seq.

For Spo11-oligo-seq, FASTQ files were downloaded using fastq-dump through accession numbers SRR094608, SRR094609, SRR094610, and SRR094611 (Pan et al., 2011). The four FASTQ files were merged together and aligned to the sacCer3 genome with Bowtie2 using default parameters. BigWig file conversions and hotspot callings used the same method as DEtail-seq without "–samFlagInclude 128".

For CC-seq, FASTQ files were downloaded using fastq-dump through accession number SRR8602148 (Gittens et al., 2019). Reads were aligned to the sacCer3 genome with Bowtie2 using default parameters. BigWig file conversions and hotspot callings used the same method as DEtail-seq, but different parameters, instead of "–samFlagInclude 128" with "–samFlagInclude 64".

For TrAEL-seq, FASTQ files were downloaded using fastq-dump through accession number SRR12279038 (Kara et al., 2021). Reads were aligned to the sacCer3 genome with Bowtie2 using default parameters. BigWig file conversions and hotspot callings used the same method as DEtail-seq.

Data availability

All the data of this study are available from the corresponding author upon any reasonable request. The GEO accession numbers for the DEtail-seq data used in this paper are GSE154226, GSE154227 and GSE154289. The detailed information of software used in this study is listed in Table S4 in Supporting Information.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Natural Science Foundation of China (91740105, 31822028, 32071437, 31900302) and Central Public-interest Scientific Institution Basal Research Fund (Y2022QC33). We greatly appreciate all the Sun Lab members for useful discussions. The Sun Lab is supported by Tsinghua-Peking Joint Center for Life Sciences.*

References

Bae, S., and Lesch, B.J. (2020). H3K4me1 distribution predicts transcription state and poising at promoters. Front Cell Dev Biol 8, 289.

Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., and de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327, 836–840.

Baudat, F., Imai, Y., and de Massy, B. (2013). Meiotic recombination in mammals: localization and regulation. Nat Rev Genet 14, 794–806.

Blat, Y., Protacio, R.U., Hunter, N., and Kleckner, N. (2002). Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. Cell 111, 791–802.

Boekhout, M., Karasu, M.E., Wang, J., Acquaviva, L., Pratto, F., Brick, K., Eng, D.Y., Xu, J., Camerini-Otero, R.D., Patel, D.J., et al. (2019). REC114 partner ANKRD31 controls number, timing, and location of meiotic DNA breaks. Mol Cell 74, 1053–1068.e8.

Borde, V., and de Massy, B. (2013). Programmed induction of DNA double strand breaks during meiosis: setting up communication between DNA and the chromosome structure. Curr Opin Genet Dev 23, 147–155.

Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T., and Weissman, J.S. (2012). High-resolution view of the yeast meiotic program revealed by ribosome profiling. Science 335, 552–557.

Brick, K., Thibault-Sennett, S., Smagulova, F., Lam, K.W.G., Pu, Y., Pratto, F., Camerini-Otero, R.D., and Petukhova, G.V. (2018). Extensive sex differences at the initiation of genetic recombination. Nature 561, 338–342.

Canela, A., Maman, Y., Jung, S., Wong, N., Callen, E., Day, A., Kieffer-

Kwon, K.R., Pekowska, A., Zhang, H., Rao, S.S.P., et al. (2017). Genome organization drives chromosome fragility. Cell 170, 507–521. e18.
Canela, A., Sridharan, S., Sciascia, N., Tubbs, A., Meltzer, P., Sleckman, B. P., and Nussenzweig, A. (2016). DNA breaks and end resection measured genome-wide by end sequencing. Mol Cell 63, 898–911.
Chen, Y., Lyu, R., Rong, B., Zheng, Y., Lin, Z., Dai, R., Zhang, X., Xie, N., Wang, S., Tang, F., et al. (2020). Refined spatial temporal epigenomic profiling reveals intrinsic connection between PRDM9-mediated H3K4me3 and the fate of double-stranded breaks. Cell Res 30, 256–268.
Chen, Y., Zheng, Y., Gao, Y., Lin, Z., Yang, S., Wang, T., Wang, Q., Xie, N., Hua, R., Liu, M., et al. (2018). Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. Cell Res 28, 879–896.
Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.
Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA 107, 21931–21936.
d'Adda di Fagagna, F. (2008). Living on a break: cellular senescence as a DNA-damage response. Nat Rev Cancer 8, 512–522.
de Massy, B. (2013). Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. Annu Rev Genet 47, 563–599.
Dujon, B. (1996). The yeast genome project: what did we learn? Trends Genet 12, 263–270.
Gansauge, M.T., Gerber, T., Glocke, I., Korlević, P., Lippik, L., Nagel, S., Riehl, L.M., Schmidt, A., and Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using *T4* DNA ligase. Nucleic Acids Res 45, e79.
Garcia, V., Phelps, S.E.L., Gray, S., and Neale, M.J. (2011). Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. Nature 479, 241–244.
Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A., and Malinverni, R. (2016). regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics 32, 289–291.
Gittens, W.H., Johnson, D.J., Allison, R.M., Cooper, T.J., Thomas, H., and Neale, M.J. (2019). A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome. Nat Commun 10, 4846.
Hinch, A.G., Becker, P.W., Li, T., Moralli, D., Zhang, G., Bycroft, C., Green, C., Keeney, S., Shi, Q., Davies, B., et al. (2020). The configuration of RPA, RAD51, and DMC1 binding in meiosis reveals the nature of critical recombination intermediates. Mol Cell 79, 689–701.e10.
Ji, F., Liao, H., Pan, S., Ouyang, L., Jia, F., Fu, Z., Zhang, F., Geng, X., Wang, X., Li, T., et al. (2020). Genome-wide high-resolution mapping of mitotic DNA synthesis sites and common fragile sites by direct sequencing. Cell Res 30, 1009–1023.
Kaiser, V.B., and Semple, C.A. (2018). Chromatin loop anchors are associated with genome instability in cancer and recombination hotspots in the germline. Genome Biol 19, 101.
Kara, N., Krueger, F., Rugg-Gunn, P., and Houseley, J. (2021). Genome-wide analysis of DNA replication and DNA double-strand breaks using TrAEL-seq. PLoS Biol 19, e3000886.
Keeney, S. (2008). Spo11 and the formation of DNA double-strand breaks in meiosis. In: Egel, R., and Lankenau, D.H., eds. Recombination and Meiosis. Genome Dynamics and Stability. Berlin, Heidelberg: Springer. 81–123.
Keeney, S., Lange, J., and Mohibullah, N. (2014). Self-organization of meiotic recombination initiation: general principles and molecular pathways. Annu Rev Genet 48, 187–214.
Lange, J., Yamada, S., Tischfield, S.E., Pan, J., Kim, S., Zhu, X., Socci, N. D., Jasin, M., and Keeney, S. (2016). The landscape of mouse meiotic double-strand break formation, processing, and repair. Cell 167, 695–708.e16.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359.
Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25.
Li, Y., Song, Y., Xu, W., Li, Q., Wang, X., Li, K., Wang, J., Liu, Z., Velychko, S., Ye, R., et al. (2020). R-loops coordinate with SOX2 in regulating reprogramming to pluripotency. Sci Adv 6, eaba0777.
Liu, T., Wang, L., Chen, H., Huang, Y., Yang, P., Ahmed, N., Wang, T., Liu, Y., and Chen, Q. (2017). Molecular and cellular mechanisms of apoptosis during dissociated spermatogenesis. Front Physiol 8.
Luo, Z., Wang, X., Jiang, H., Wang, R., Chen, J., Chen, Y., Xu, Q., Cao, J., Gong, X., Wu, J., et al. (2020). Reorganized 3D genome structures support transcriptional regulation in mouse spermatogenesis. iScience 23, 101034.
Mimitou, E.P., and Symington, L.S. (2009). DNA end resection: many nucleases make light work. DNA Repair 8, 983–995.
Mimitou, E.P., Yamada, S., and Keeney, S. (2017). A global view of meiotic double-strand break end resection. Science 355, 40–45.
Neale, M.J., Pan, J., and Keeney, S. (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. Nature 436, 1053–1057.
Ohkura, H. (2015). Meiosis: an overview of key differences from mitosis. Cold Spring Harb Perspect Biol 7, a015859.
Paiano, J., Wu, W., Yamada, S., Sciascia, N., Callen, E., Paola Cotrim, A., Deshpande, R.A., Maman, Y., Day, A., Paull, T.T., et al. (2020). ATM and PRDM9 regulate SPO11-bound recombination intermediates during meiosis. Nat Commun 11, 857.
Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H.G., Tischfield, S.E., Zhu, X., Neale, M.J., Jasin, M., Socci, N.D., et al. (2011). A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. Cell 144, 719–731.
Panier, S., and Durocher, D. (2013). Push back to respond better: regulatory inhibition of the DNA double-strand break response. Nat Rev Mol Cell Biol 14, 661–672.
Petrosino, G., Zilio, N., Sriramachandran, A.M., and Ulrich, H.D. (2020). Preparation and analysis of GLOE-seq libraries for genome-wide mapping of DNA replication patterns, single-strand breaks, and lesions. STAR Protoc 1, 100076.
Pratto, F., Brick, K., Khil, P., Smagulova, F., Petukhova, G.V., and Camerini-Otero, R.D. (2014). Recombination initiation maps of individual human genomes. Science 346, 1256442.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.
Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42, W187–W191.
Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol 29, 24–26.
Scully, R., Panday, A., Elango, R., and Willis, N.A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. Nat Rev Mol Cell Biol 20, 698–714.
Shaha, C., Tripathi, R., and Mishra, D.P. (2010). Male germ cell apoptosis: regulation and biology. Phil Trans R Soc B 365, 1501–1515.
Sriramachandran, A.M., Petrosino, G., Méndez-Lago, M., Schäfer, A.J., Batista-Nascimento, L.S., Zilio, N., and Ulrich, H.D. (2020). Genome-wide nucleotide-resolution mapping of DNA Replication patterns, single-strand breaks, and lesions by GLOE-Seq. Mol Cell 78, 975–985.e7.
Symington, L.S. (2014). End resection at double-strand breaks: mechanism and regulation. Cold Spring Harb Perspect Biol 6, a016436.
Tubbs, A., Sridharan, S., van Wietmarschen, N., Maman, Y., Callen, E., Stanlie, A., Wu, W., Wu, X., Day, A., Wong, N., et al. (2018). Dual roles of poly(dA:dT) tracts in replication initiation and fork collapse. Cell 174, 1127–1142.e19.

Weinbauer, G.F., Aslam, H., Krishnamurthy, H., Brinkworth, M.H., Einspanier, A., and Hodges, J.K. (2001). Quantitative analysis of spermatogenesis and apoptosis in the common marmoset (*Callithrix jacchus*) reveals high rates of spermatogonial turnover and high spermatogenic efficiency. Biol Reprod 64, 120–126.

Wen, F.P., Guo, Y.S., Hu, Y., Liu, W.X., Wang, Q., Wang, Y.T., Yu, H.Y., Tang, C.M., Yang, J., Zhou, T., et al. (2016). Distinct temporal requirements for autophagy and the proteasome in yeast meiosis. Autophagy 12, 671–688.

White, C.I., and Haber, J.E. (1990). Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. EMBO J 9, 663–673.

Xu, W., Li, K., Li, S., Hou, Q., Zhang, Y., Liu, K., and Sun, Q. (2020). The R-loop atlas of *Arabidopsis* development and responses to environmental stimuli. Plant Cell 32, 888–903.

Xu, W., Xu, H., Li, K., Fan, Y., Liu, Y., Yang, X., and Sun, Q. (2017). The R-loop is a common chromatin feature of the *Arabidopsis* genome. Nat Plants 3, 704–714.

Xu, Z., Song, Z., Li, G., Tu, H., Liu, W., Liu, Y., Wang, P., Wang, Y., Cui, X., Liu, C., et al. (2016). H2B ubiquitination regulates meiotic recombination by promoting chromatin relaxation. Nucleic Acids Res 44, 9681–9697.

Yang, X., Liu, Q.L., Xu, W., Zhang, Y.C., Yang, Y., Ju, L.F., Chen, J., Chen, Y.S., Li, K., Ren, J., et al. (2019). m^6A promotes R-loop formation to facilitate transcription termination. Cell Res 29, 1035–1038.

Young, K.A., Ball, G.F., and Nelson, R.J. (2001). Photoperiod-induced testicular apoptosis in European starlings (*Sturnus vulgaris*). Biol Reprod 64, 706–713.

Yuan, W., Zhou, J., Tong, J., Zhuo, W., Wang, L., Li, Y., Sun, Q., and Qian, W. (2019). ALBA protein complex reads genic R-loops to maintain genome stability in *Arabidopsis*. Sci Adv 5, eaav9040.

Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell 134, 981–994.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-022-2277-y. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.