

ARTICLE OPEN



High-resolution analysis for urinary DNA jagged ends

Tingting Xie^{1,2,3,6}, Guangya Wang^{1,2,3,6}, Spencer C. Ding^{1,2,3}, Wing-Shan Lee^{1,2,3}, Suk Hang Cheng^{1,2,3}, Rebecca W. Y. Chan^{1,2,3}, Ze Zhou^{1,2,3}, Mary-Jane L. Ma^{1,2,3}, Diana S. C. Han^{1,2,3}, Jeremy Y. C. Teoh⁴, W. K. Jacky Lam^{1,2,3}, Peiyong Jiang^{1,2,3}, Rossa W. K. Chiu^{1,2,3}, K. C. Allen Chan^{1,2,3,5} and Y. M. Dennis Lo^{1,2,3,5}✉

Single-stranded ends of double-stranded DNA (jagged ends) are more abundant in urinary DNA than in plasma DNA. However, the lengths of jagged ends in urinary DNA remained undetermined, as a previous method used for urinary DNA jagged end sequencing analysis (Jag-seq) relied on unmethylation at CpG sites, limiting the resolution. Here, we performed high-resolution Jag-seq analysis using methylation at non-CpG cytosine sites, allowing determination of exact length of jagged ends. The urinary DNA bore longer jagged ends (~26-nt) than plasma DNA (~17-nt). The jagged end length distribution displayed 10-nt periodicities in urinary DNA, which were much less observable in plasma DNA. Amplitude of the 10-nt periodicities increased in patients with renal cell carcinoma. Heparin treatment of urine diminished the 10-nt periodicities. The urinary DNA jagged ends often extended into nucleosomal cores, suggesting potential interactions with histones. This study has thus advanced our knowledge of jagged ends in urine DNA.

npj Genomic Medicine (2022)7:14; <https://doi.org/10.1038/s41525-022-00285-1>

INTRODUCTION

A number of recent efforts have been taken to explore the biological properties and clinical utilities of urinary cell-free DNA molecules^{1,2}. Single-stranded overhangs of a double-stranded DNA fragment (i.e. jagged ends) were commonly present in plasma, which were possibly related to DNA fragmentation processes such as DNA nuclease activities^{3,4}. In the plasma DNA pool, fetal DNA and tumoral DNA molecules were generally characterized with higher jaggedness compared with background DNA mainly of hematopoietic origin⁴.

In a recent study, the overall jaggedness of urinary DNA was found to be much higher than that of plasma DNA¹. The jaggedness of urinary DNA decreased in patients with bladder cancer, compared with subjects without bladder cancer¹. However, the previous study regarding urinary DNA jagged ends was based on the detection of unmethylated CpG sites introduced during the DNA end-repair process¹. The analysis would limit the resolution of jagged end analysis of urinary DNA, because only 1% of dinucleotides in the human genome are CpG sites⁴. We reasoned that it would be informative to bring new mechanistic insights into jagged ends of urinary DNA using an approach with higher resolution by which the exact length of jagged ends could be determined. It was reported that the use of methylation signal at CH (H: A, C, or T) sites allowed the determination of exact length of jagged ends, on the basis of enhanced jagged end sequencing analysis (Jag-seq)⁴. Such methylation signals at CH sites were introduced during the DNA end-repair process with the use of methylated cytosines instead of unmethylated cytosines⁴.

In this study, we adopted a high-resolution Jag-seq approach to investigate the properties of urinary DNA jagged ends in human subjects with and without renal cell carcinoma (RCC) (Fig. 1).

RESULTS

Evaluation of modified Jag-seq based on synthetic oligonucleotides

We refined the Jag-seq method by removing the step of 3'→5' exonuclease T (Exo T) treatment, thus preserving the authenticity of urinary DNA ends. The missing complements of 5' single-stranded protruding DNA ends present in a double-stranded DNA molecule were restored using dATPs (As), dTTPs (Ts), dGTPs (Gs), and methylated dCTPs (i.e. mCs) using an end-repair process, resulting in blunt ends. Such end-repaired DNA molecules were ligated with sequencing adapters, followed by bisulfite treatment. Hence, the methylated signal at CH sites proximal to the ends of urinary DNA molecule could be used for reflecting the jaggedness of urinary DNA.

To validate this refined Jag-seq, we designed two synthetic oligonucleotides. The first synthetic oligonucleotide was a 46-bp molecule carrying a 1-nt 5' protruding G nucleotide end, with another G nucleotide immediately preceding that jagged end (Supplementary Fig. 1a). The second synthetic oligonucleotide carried 14-nt 5' protruding ends, with four consecutive C nucleotides at the recessed ends (Supplementary Fig. 1b). For data generated by an experimental protocol without Exo T (Fig. 2a), we observed that 99.9% of nucleotides at the first base relative to the 3' end were determined to be Cs, whereas only 0.2% of nucleotides at the second base relative to the 3' end were determined to be Cs. These results suggested that the 1-nt jagged end could be accurately detected according to the methylation signal introduced by the end-repair process. For the data generated by the previously published experimental protocol with Exo T⁴, a substantial proportion of nucleotides at the second base relative to the 3' end was determined to be Cs (98.2%), suggesting that a proportion of 1-nt jagged ends was changed to at least 2-nt jagged ends (Fig. 2b). From the data related to the

¹Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China. ²Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China. ³Centre for Novostics, Hong Kong Science Park, Pak Shek Kok, New Territories, Hong Kong SAR, China. ⁴S.H. Ho Urology Centre, Department of Surgery, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China. ⁵State Key Laboratory of Translational Oncology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China. ⁶These authors contributed equally: Tingting Xie, Guangya Wang. ✉email: loym@cuhk.edu.hk

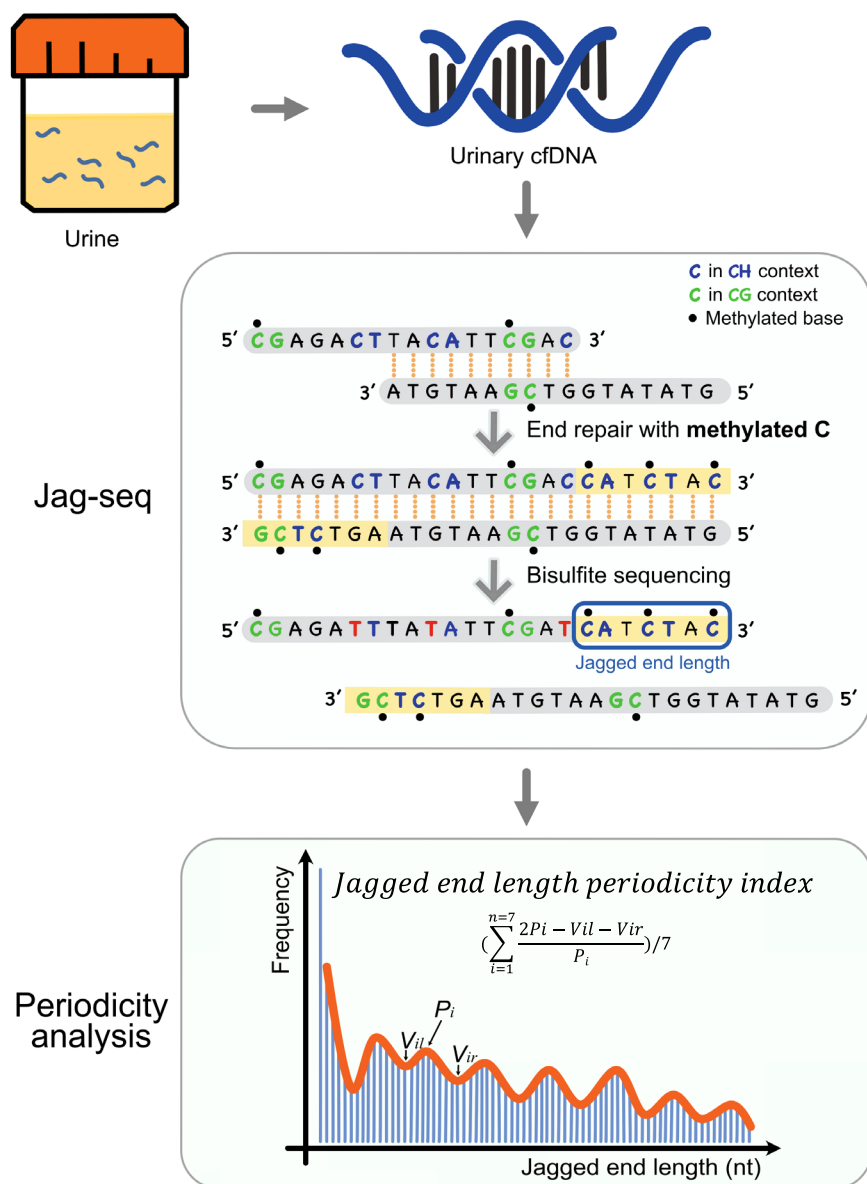


Fig. 1 Illustration for jagged end analysis in this study. The 5' protruding ends of urinary double-stranded molecules were filled up with dATP (A), dGTP (G), dTTP (T), and methylated dCTP (5mC) with the use of Hemo KlenTaq DNA polymerase that lacked 3'→5' exonuclease activity and strand displacement. The methylation signals were introduced through such end-repair process. Based on bisulfite sequencing, the methylation signals at the CH sites near the 3' end of a DNA strand indicates the presence of jagged ends. The exact jagged end length of a DNA molecule can be deduced with the two consecutive Cs in the strand with the presence of jagged ends, in which the first C is detected to be unmethylated by bisulfite treatment, but the second C be methylated. Jagged end length periodicity index was measured with series of peaks (P_i) and troughs (V_i) and used for renal cell carcinoma detection.

second synthetic oligonucleotide (Fig. 2d), there were 15.9% and 2.6% of nucleotides identified as Cs at the positions of 1-nt and 2-nt relative to the jagged-end starting point in data prepared by the protocol with Exo T, whereas the counterparts were nearly close to 0% (Fig. 2c). Such data suggested that Exo T would commonly introduce an extra of 1-nt jagged nucleotide in a double-stranded DNA end.

Difference in jaggedness between plasma and urinary DNA

We applied the modified Jag-seq approach to analyze urinary DNA and plasma DNA from healthy control subjects. CH methylation levels proximal to the 3' end of a DNA molecule (i.e. read2), which was termed Jagged index-methylated (JI-M) values, were used to reflect the jaggedness of cell-free DNA. As shown in Fig. 3a, the JI-

M value of urinary DNA (median: 48.78; range: 38.08–56.47) was much higher than that in plasma DNA (median: 17.33; range: 14.22–20.62) (p value: 4×10^{-5} , Mann–Whitney U test). The value of JI-M varied according to the sizes of molecules, showing wave-like patterns in both urinary and plasma DNA (Fig. 3b). For urinary DNA molecules, the JI-M value rapidly increased ranging from 25 to 80. The profile of JI-M over fragment sizes displayed a small peak at ~120 bp and a major peak at ~250 bp, followed by a second major peak at ~410 bp. We speculated that the cell-free DNA molecules with ~120 bp in size might not be associated with an intact nucleosome structure, whereas the cell-free DNA molecules at ~250 bp and ~410 bp might be associated with intact nucleosome structures. These two populations of cell-free DNA were previously reported to have different DNA nuclease accessibilities in plasma⁵. For plasma DNA, the JI-M value had a similar trend across different

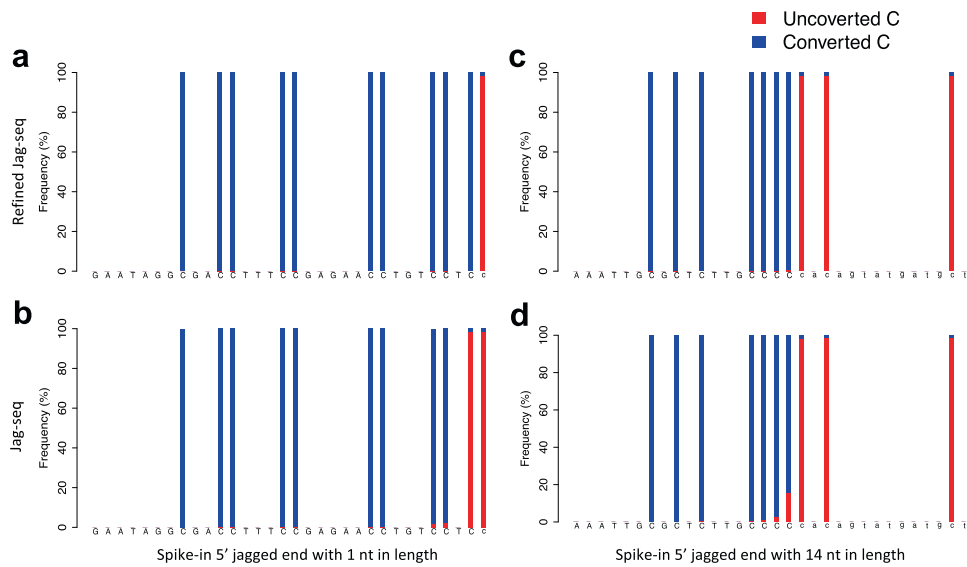


Fig. 2 Validation for spike-in molecules with 5' protruding jagged ends. **a–d** Sequencing base compositions for spike-in sequence with known jagged ends. The partial sequence of the spiked-in sequences with a 1-nt and a 14-nt jagged end are indicated on the x-axis. The nucleotides denoted in the uppercase letters indicate that the sequences are in double-stranded form. The nucleotides in lowercase letters indicate that the sequences are newly filled during the end-repair process. Vertical bars with blue color and red color represent the frequencies of sequenced T and C, respectively. Sequencing results from the refined (**a, c**) and original version (**b, d**) of Jag-seq are shown.

fragment sizes but with lower JI-M values. These results were generally consistent with the previous study¹.

We could further determine the exact length of jagged ends for a subset of DNA molecules, on the basis of the CC-tag strategy published previously⁴. This method was to employ the methylation pattern for two consecutive cytosines in a molecule for which the first cytosine was unmethylated but the second cytosine was methylated. Such a methylation pattern enabled the determination of the starting point of a jagged end. The average jagged end length in urinary DNA (median: 27.3 nt; range: 22.5–28.6 nt) was much longer than plasma DNA (median: 16.4 nt; range: 13.2–19.0 nt) (p value: 4×10^{-5} , Mann–Whitney U test) (Fig. 3c, d and Supplementary Fig. 2a, b).

Jagged end length periodicity of urinary cfDNA

We studied the relative frequency of jagged end length ranging from 0 to 74 nucleotides (nt). Generally, the longer jagged ends coincided with the lower abundance in urinary DNA (Fig. 4a). Interestingly, we observed that there were 10-nt periodic patterns of jagged end length of urinary cfDNA, which were much less observable in plasma DNA (Supplementary Fig. 3).

To quantitatively analyze the periodicity of jagged end length, we used the relative difference between a series of peak values and the paired trough values using the following formula:

$$\text{Jagged end length periodicity index} = \frac{\sum_{i=1}^{n-7} \frac{2P_i - V_{i-7} - V_{i+7}}{P_i}}{7} \quad (1)$$

Where the P_i is the frequency of jagged end length at a particular peak i , V_{i-7} is the frequency of jagged end length at the left valley relative to the peak i and V_{i+7} is the frequency of jagged end length at the right valley relative to the peak i .

A higher jagged end length periodicity index indicated that there was a stronger 10-nt periodic patterns present in urinary DNA jagged ends. Such jagged end length periodicity index was shown to be higher in urinary DNA of patients with RCC (median: 1.06; range: 0.81–1.38), compared with control subjects (median: 0.72; range: 0.37–1.09) (p value: 0.01, Mann–Whitney U test) (Fig. 4b). The receiver operating characteristic curve (ROC) analysis

showed that the area under the curve (AUC) was 0.86 in differentiating patients with and without RCC, based on the periodicity index of urinary DNA jagged end length (Fig. 4c). These results implied that the periodicity of urinary DNA jagged end length might be used as a biomarker for RCC detection.

Figure 4d shows that sharp oscillations of jagged end length periodicity index were observed in those urinary DNA molecules below 120 bp in both RCC patients and healthy controls, whereas such oscillations were attenuated for those molecules above 120 bp.

Effects of heparin treatment on jagged end length periodicity

The strong 10-nt periodicities present in the distribution of jagged end lengths were reminiscent of the intrinsic characteristics of helical periodicity close to 10 bp per turn on the nucleosome. We conjectured that the urinary DNA molecules would be in part associated with histone proteins. As heparin with a high affinity to all histones could disrupt the interaction between DNA and histone proteins^{6,7}, we used heparin to treat the urinary DNA to study whether the 10-nt periodicities of jagged end lengths would be dependent on histone proteins or not.

We performed EDTA treatment and heparin treatment on urine samples for 0 h, 0.5 h, and 1 h of in vitro incubation under room temperature, followed by Jag-seq. The 10-nt periodicities were remarkably reduced for the urine samples treated with heparin even at 0 h, compared with those treated by EDTA (Fig. 5a). With the prolonged periods of heparin incubation, the periodicity index was gradually decreased (Fig. 5b). Both the JI-M value and the average jagged end length showed decreased trends with heparin treatment (Supplementary Fig. 4a, b). These data suggested that the urinary DNA would be in part associated with histone proteins.

Relationship between jagged ends and nucleosomal structures

We recently presented evidence that the fragmentation patterns of cfDNA were associated with nucleosomal structures in urine^{1,7} and plasma^{8–10}. High-resolution Jag-seq could provide an

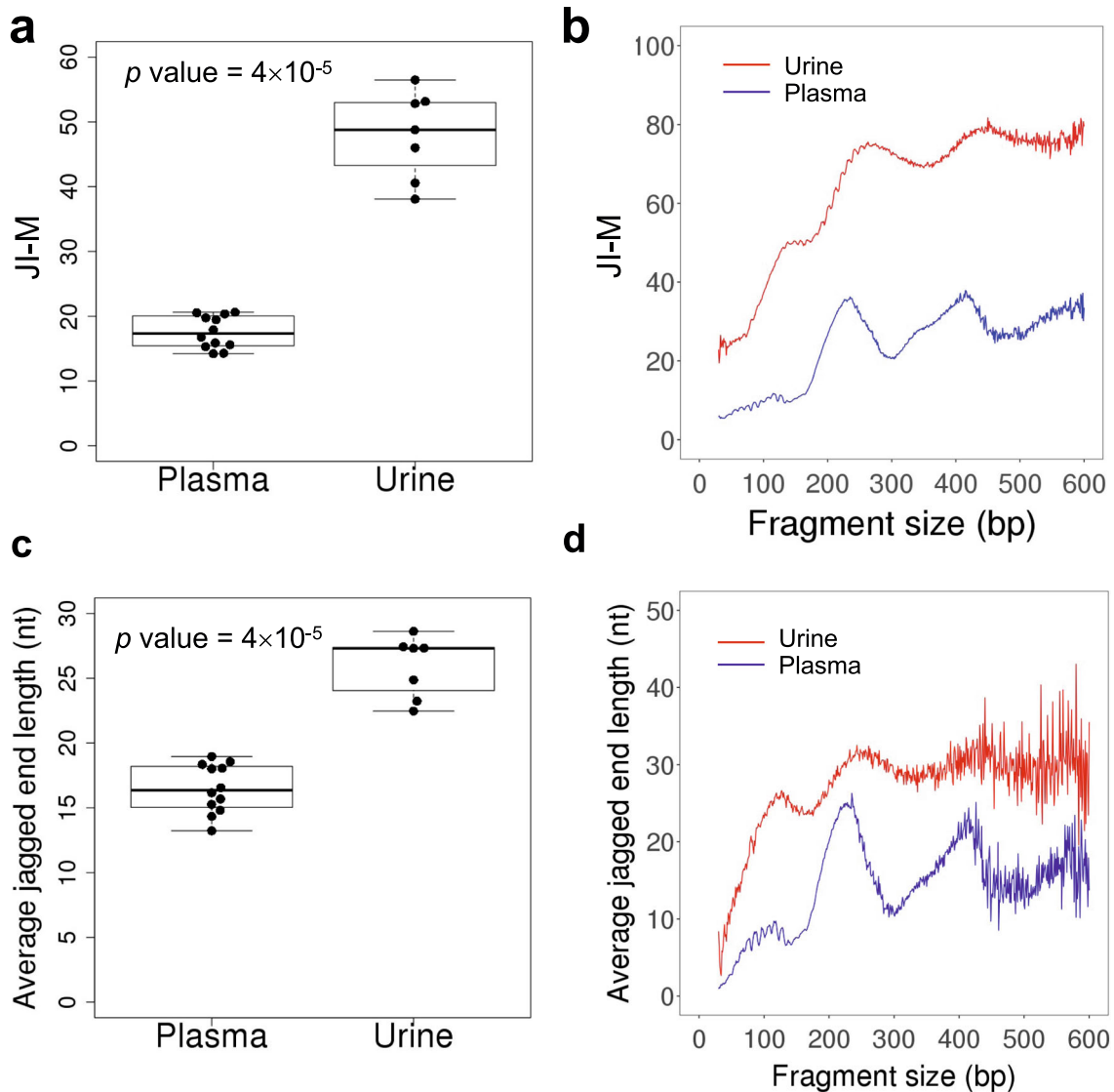


Fig. 3 Comparison of jagged ends between plasma and urinary DNA. **a, b** CH methylation level in read2 (JI-M) across the different fragment sizes. **c, d** Average jagged end length across the different fragment sizes. The central line indicates the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range.

opportunity to study the 3' recessed ends and the corresponding 5' protruding ends in the complementary strand for a cell-free DNA molecule (Fig. 6a). For a double-stranded DNA molecule with 5' overhangs, there were two pairs of 3' recessed ends and 5' protruding ends, namely part A (corresponding to a lower value in the genome coordinate) and part B (corresponding to a higher value in the genome coordinate) (Fig. 6a).

We sought to investigate whether the 3' recessed end density and the 5' protruding end density would be related to nucleosomal structures. We applied the analysis in urinary DNA molecules falling within the 1-kb upstream and downstream relative to CTCF binding sites and calculated the end density around CTCF binding sites. The nucleosome tracks around CTCF binding sites were determined by the cell-free DNA coverages (gray lines in Fig. 6b). The preferred cutting sites during cfDNA generation were indicated by the peaks of end density¹¹.

Figure 6b shows the results of plasma DNA (upper panel) and urinary DNA (lower panel) end density derived from part A. The peaks of 5' protruding end density and 3' recessed end density for plasma DNA were both approximately aligned to the linker

regions, indicating that generation of plasma jagged ends might be derived from the cutting of the linker regions (Fig. 6b). A small shift between the 3' recessed and 5' protruding end density curves was observed in plasma DNA, likely because of the presence of the jagged ends (Fig. 6b). In contrast, in urine samples, the peaks of 3' recessed end density and 5' protruding end density were phased in different locations. The large shift between 3' recessed and 5' protruding end density curves indicated that the DNA degradation from the 3' end might be more severe in urinary DNA, even extending into the nucleosomal cores. These data highlighted the potential interaction between urinary DNA and nucleosomal cores. The patterns derived from part B were consistent with that from part A (Supplementary Fig. 5).

We also studied the jagged end length periodicity index within all the nucleosomal core regions (143 bp) and among the linker regions (20 bp), respectively. The nucleosomal cores were defined by the 143-bp regions surrounding the peak signals of nucleosome tracks present in Fig. 6c. The linker regions were defined by the 20-bp regions at each side of a nucleosomal core. The jagged end length periodicity index was found to be higher at the

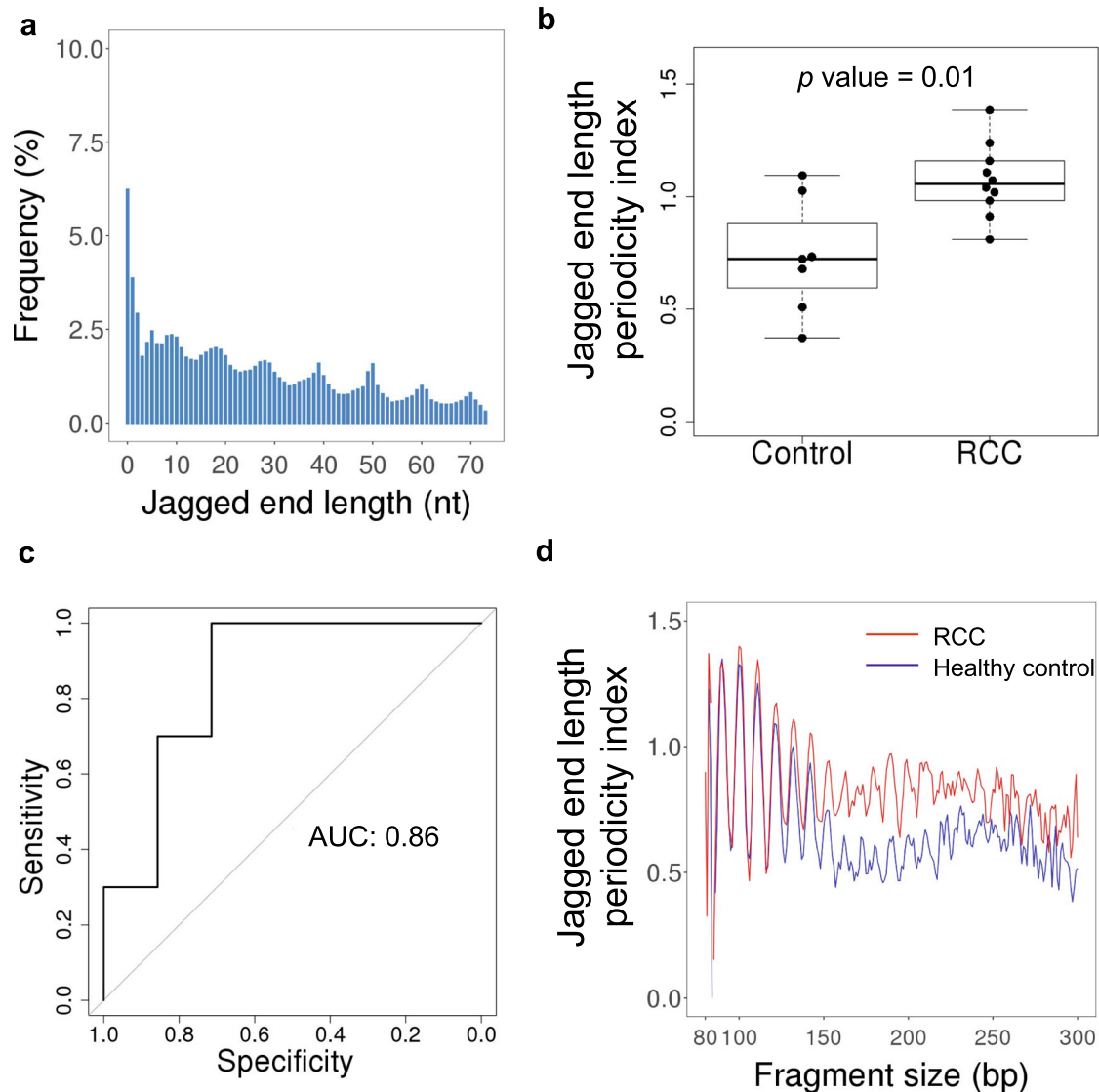


Fig. 4 The characteristics of jagged end length of urinary DNA. **a** Jagged end length distribution measured by the CC-tag strategy. **b** Overall jagged end length periodicity index in RCC patients and healthy controls. The central line indicates the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. **c** The area under the ROC curve (AUC) of differentiating between patients with and without RCC using jagged end length periodicity index. **d** Jagged end length periodicity index across different fragment sizes in healthy controls and patients with RCC.

nucleosomal core regions (0.77) compared with the linker regions (0.68) in urine (Fig. 6c). These results further highlighted that the generation of 10-nt periodicity of jagged end length might be associated with histone proteins in urine.

DISCUSSION

In this work, we demonstrated a high-fidelity experimental protocol for analyzing jagged ends of urinary DNA at single-base resolution. We revealed that the exonuclease T would artificially modify a proportion of jagged ends with an extra 1 nt or 2 nt, when it was used in DNA end-repair process for removing 3' overhangs of double-stranded DNA to generate blunt-ends. The key refinement over a previously published protocol for jagged end analysis was removal of the exonuclease T (Exo T) step. Even though exonuclease T reportedly has a high activity of 3'->5' exonuclease on single-stranded DNA, our data clearly suggested

that exonuclease T would also bear a low activity of 3'->5' exonuclease on double-stranded regions.

With a high-fidelity experimental protocol for jagged end analysis, we found that the urinary DNA molecules bore much higher jaggedness than plasma DNA, which was consistent with a previous report¹. For example, the median of average jagged end length among urinary DNA samples was 27.3 nt, whereas the counterpart among plasma DNA samples was 16.4 nt. Such consistency between studies suggested that the jagged ends were the intrinsic properties of cell-free DNA.

In addition, the jagged end analysis based on end-repair process with methylated cytosine would allow for the deduction of exact jagged end length for a subset of DNA molecules⁴. Jagged end lengths in urinary cfDNA displayed 10-nt periodic patterns which were not observable in plasma DNA. These results suggested that the mechanism in the production of jagged ends might be different between urinary DNA and plasma DNA. There are indeed a number of factors that are different between urine

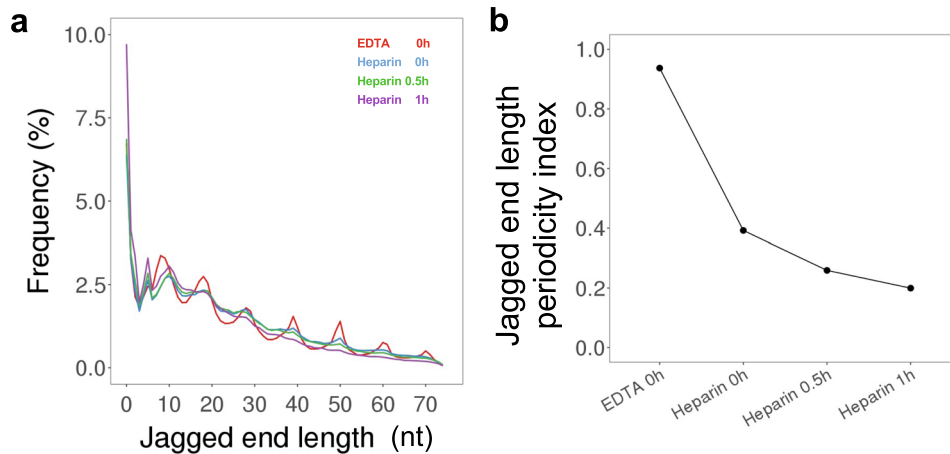


Fig. 5 Effects of heparin treatment on jagged ends of urinary cfDNA. **a** Jagged end length distribution in urinary cfDNA with heparin incubation treatment. The lines with red, blue, green, and purple colors represent jagged end length distribution at EDTA 0 h, heparin 0 h, heparin 0.5 h, and heparin 1 h treatment, respectively. **b** Jagged end length periodicity index with EDTA and heparin treatment.

and plasma. For example, the DNASE1 activity was reported to be much higher in urine than in plasma¹².

Interestingly, the amplitude of jagged end length periodicity enabled us to differentiate subjects suffering from RCC from healthy individuals. The periodicity index of jagged end length was significantly higher in patients with RCC. We achieved an AUC of 0.86 between healthy controls and patients with RCC using the value of jagged end length periodicity index. Thus, the periodicity index of jagged end length could serve as an emergent biomarker for RCC.

The 10-nt periodicities of jagged end length disappeared in the urinary DNA once heparin was added into the urine. As heparin could disrupt the interaction between DNA and histone proteins^{6,7,13}, we thus speculated that urinary DNA would in part be associated with histone proteins, rendering such 10-nt periodicities of jagged end length present in urinary DNA. Seminucleosomal structures (H₃H₄)₂ were reported to be energetically favorable during nucleosome disassembly¹⁴. The absence of 10-nt periodicities of jagged end length in plasma DNA might be due to the fact that the major DNA nuclease involving plasma DNA fragmentation was DNASE1L3 instead of DNASE1^{5,15}. We further demonstrated that the 3' recessed ends appeared to more often extend to nucleosomal cores in comparison with plasma DNA molecules. The jagged end length periodicity index was higher in nucleosomal cores than in linker regions. These results implied another evidence that the presence of urinary DNA jagged ends might in part interact with histone proteins. However, the determination of 3' recessed end relied on the two consecutive cytosines in a fragment, thus only reflecting the information regarding a subset of cell-free DNA molecules. It would be interesting to verify this observation in another study using other methods such as ligation-based jagged end detection¹⁶.

In summary, urinary DNA was characterized with higher jaggedness in comparison with plasma DNA. The high-fidelity experimental protocol for jagged ends analysis revealed the 10-nt periodicities of jagged end length that were not realized before. The generation of jagged ends of urinary DNA might be in part associated with histone proteins. Such 10-nt periodicities of jagged end length had potential for RCC detection.

METHODS

Human sample collection and processing

Urine samples of kidney cancer were recruited from the Department of Surgery of the Prince of Wales Hospital, Hong Kong. Healthy samples were

recruited from the Department of Chemical Pathology of the Prince of Wales Hospital, Hong Kong. All subjects involved in this study gave written informed consent, and the study was approved by The Joint Chinese University of Hong Kong–Hospital Authority New Territories East Cluster Clinical Research Ethics Committee, under the Declaration of Helsinki. Collection and storage of fresh urine were followed as previously described¹⁷. Urinary DNA was extracted from 10 mL–30 mL cell-free urine component with the Wizard Plus Minipreps DNA Purification System; (Promega). 15 mL of 6 mol/L guanidine thiocyanate (Sigma–Aldrich) and 1 mL of resin (Wizard Plus Minipreps DNA Purification System; Promega) were added in each 10 mL of the processed urine. The mixture was incubated and shaken gently at room temperature for 2 h. Then, the DNA was purified and eluted into 100 µL RNase-free water using the minicolumns and according to the manufacturer's instruction of the purification system. The quantity of urinary DNA obtained ranged from ~10 ng to ~100 ng according to Qubit assays.

Library preparation

Hemo KlenTaq (NEB) and dATP (A), dGTP (G), dTTP (T), and methylated dCTP (5mC) were used for filling in 5' protruding ends of double-stranded DNA molecules, forming blunt-end molecules. 3' A-tailing on blunt ends was processed by Hemo KlenTaq. The end-repaired DNA was subjected to DNA purification using a MinElute Reaction Cleanup kit (QIAGEN). We further applied polynucleotide kinase (PNK) (NEB) for 5' phosphorylation of purified DNA, followed by methylated adapter ligation by T4 DNA ligase (NEB), obtaining the resultant DNA libraries ready for sequencing. All methylated adapters were TruSeq single indexed adapter (Illumina). The molecules with 3' protruding ends were not able to be ligated with sequencing adapters, thus would not be present in the sequencing results.

We treated the adapter-ligated DNA with two rounds of bisulfite conversion by EpiTect Plus DNA Bisulfite kit (Qiagen). We amplified bisulfite-converted DNA molecules with 12 cycles using KAPA HiFi HotStart Uracil + ReadyMix (Roche). PCR primers included P5 PCR Primer (5' AA TGATACGGCGACCACCGAG) and P7 PCR Primer (5' CAAGCAGAAGACGGC ATACGAG).

We examined the quality of DNA library and the size of DNA by an Agilent 4200 TapeStation with HSD1000 ScreenTape (Agilent). The DNA libraries were sequenced on a NextSeq platform (Illumina) using the paired-end mode (75-bp × 2).

Sequencing data analysis

The adapter sequences and low-quality bases were trimmed off from the raw sequencing reads using Trim Galore¹⁸. The pre-processed reads were aligned to the human genome (hg38) using Bismark¹⁹. The median number of aligned paired-end sequencing reads was 24.1 million (interquartile range: 16.0 – 28.1 million). The jagged ends detection was performed as described in our previous study⁴. The Mann–Whitney test was used to compare the difference between groups. *P* value of <0.05 was considered statistically significant.

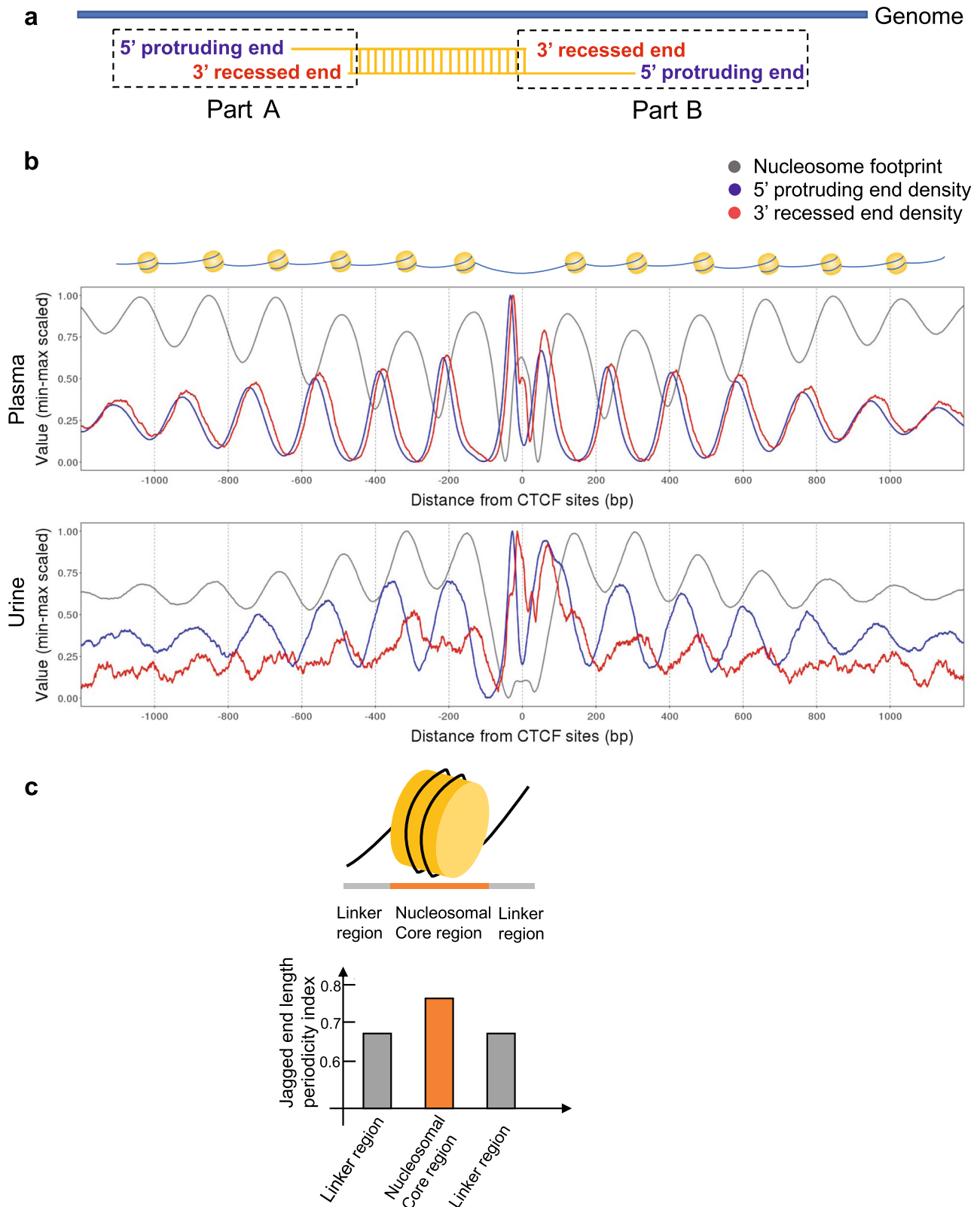


Fig. 6 Relationship between jagged ends and nucleosome tracks. a Illustration for the definition of 5' protruding end and 3' recessed end in part A and part B. **b** 5' protruding end density and 3' recessed end density surrounding CTCF binding sites for plasma and urinary DNA molecules in part A. **c** Jagged end length periodicity index in nucleosomal linker regions and nucleosomal core regions.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

The sequence data generated and analyzed have been deposited in the European Genome-Phenome Archive (EGA), www.ebi.ac.uk/ega, hosted by the European Bioinformatics Institute (EBI), www.ebi.ac.uk (accession no. <https://ega-archive.org/studies/EGAS00001005603>).

Received: 20 September 2021; Accepted: 22 December 2021;

Published online: 23 February 2022

REFERENCES

- Zhou, Z. et al. Jagged ends of urinary cell-free DNA: characterization and feasibility assessment in bladder cancer detection. *Clin. Chem.* **67**, 621–630 (2021).
- Cheng, T. H. T. et al. Noninvasive detection of bladder cancer by shallow-depth genome-wide bisulfite sequencing of urinary cell-free DNA for methylation and copy number profiling. *Clin. Chem.* **65**, 927–936 (2019).
- Lo, Y. M. D., Han, D. S. C., Jiang, P. & Chiu, R. W. K. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. *Science* **372**, eaaw3616 (2021).
- Jiang, P. et al. Detection and characterization of jagged ends of double-stranded DNA in plasma. *Genome Res.* **30**, 1141–1153 (2020).
- Han, D. S. C. et al. The biology of cell-free DNA fragmentation and the roles of DNASE1, DNASE1L3, and DFFB. *Am. J. Hum. Genet.* **106**, 202–214 (2020).
- Villeponteau, B. Heparin increases chromatin accessibility by binding the trypsin-sensitive basic residues in histones. *Biochem. J.* **288**, 953–958 (1992).
- Markus, H. et al. Analysis of recurrently protected genomic regions in cell-free DNA found in urine. *Sci. Transl. Med.* **13**, 52 (2021).
- Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M. & Shendure, J. Cell-free DNA comprises an *in vivo* nucleosome footprint that informs its tissues-of-origin. *Cell* **164**, 57–68 (2016).
- Sun, K. et al. Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing. *Proc. Natl Acad. Sci. USA* **115**, E5106–E5114 (2018).
- Straver, R., Oudejans, C. B. M., Sistermans, E. A. & Reinders, M. J. T. Calculating the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles. *Prenat. Diagn.* **36**, 614–621 (2016).
- Sun, K. et al. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. *Genome Res.* **29**, 418–427 (2019).
- Nadano, D., Yasuda, T. & Kishi, K. Measurement of deoxyribonuclease I activity in human tissues and body fluids by a single radial enzyme-diffusion method. *Clin. Chem.* **39**, 448–452 (1993).
- Brotherton, T. W., Jagannadham, M. V. & Ginder, G. D. Heparin binds to intact mononucleosomes and induces a novel unfolded structure. *Biochemistry* **28**, 3518–3525 (1989).
- Thåström, A., Bingham, L. M. & Widom, J. Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. *J. Mol. Biol.* **338**, 695–709 (2004).
- Serpas, L. et al. Dnase1l3 deletion causes aberrations in length and end-motif frequencies in plasma DNA. *Proc. Natl Acad. Sci. USA* **116**, 641–649 (2019).
- Harkins, K. M. et al. A novel NGS library preparation method to characterize native termini of fragmented DNA. *Nucleic Acids Res.* **48**, 47 (2020).
- Hung, E. C. W. et al. Presence of donor-derived DNA and cells in the urine of sex-mismatched hematopoietic stem cell transplant recipients: implication for the transrenal hypothesis. *Clin. Chem.* **55**, 715–722 (2009).
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).
- Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

ACKNOWLEDGEMENTS

This work was supported by the Research Grants Council of the Hong Kong SAR Government under the Theme-based Research Scheme (T12-401/16-W), a collaborative research agreement from Grail, the Innovation and Technology Commission, and the Vice Chancellor's One-Off Discretionary Fund of The Chinese University of Hong Kong (VCF2014021). Y.M.D.L. is supported by an endowed chair from the Li Ka Shing Foundation.

AUTHOR CONTRIBUTIONS

T.X., P.J., and Y.M.D.L. designed research. T.X., S.C.D., S.H.C., and R.W.Y.C. performed research. T.X., G.W., W.L., Z.Z., and P.J. performed data analysis. M.L.M., J.Y.C.T., W.K.J.L., and K.C.A.C. recruited patients and analyzed clinical data. T.X., G.W., P.J., and Y.M.D.L. wrote the manuscript. D.S.C.H., P.J., R.W.K.C., K.C.A.C., and Y.M.D.L. reviewed and interpreted the data.

COMPETING INTERESTS

K.C.A.C., R.W.K.C., and Y.M.D.L. hold equities in DRA, Take2, and Grail/Illumina. K.C.A.C., R.W.K.C., and Y.M.D.L. were consultants to Grail. K.C.A.C., R.W.K.C., and Y.M.D.L. had received research funding from Grail. Y.M.D.L. is a scientific cofounder of and serves on the scientific advisory board of Grail. R.W.K.C. is a consultant to Illumina. W.K.J.L. and P.J. hold equities in Grail/Illumina. P.J. hold equities in DRA. P.J. is a consultant to Take2. P.J. is a Director of KingMed Future. T.X., G.W., S.C.D., P.J., S.H.C., Z.Z., K.C.A.C., R.W.K.C., and Y.M.D.L. have filed patent applications based on the data generated from this work. Patent royalties are received from Grail, Illumina, Sequenom, DRA, Take2, and Xcelom. Y.M.D.L. is an Associate Editor for npj Genomic Medicine.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41525-022-00285-1>.

Correspondence and requests for materials should be addressed to Y. M. Dennis Lo.

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

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



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

© The Author(s) 2022