

# Caution for simple sequence repeat number variation in the mitochondrial DNA D-loop to determine cancer-specific variants

TOKIKO NAKAI<sup>1,2\*</sup>, AKIHISA SAKURADA<sup>3\*</sup>, TOSHIHIDE ENDO<sup>3\*</sup>, HIROKO KOBAYASHI<sup>1\*</sup>,  
SHINOBU MASUDA<sup>1</sup>, MAKOTO MAKISHIMA<sup>3</sup> and MARIKO ESUMI<sup>3</sup>

<sup>1</sup>Division of Oncologic Pathology, Department of Pathology and Microbiology; <sup>2</sup>Department of Diagnostic Pathology, Nara Medical University, Kashihara, Nara 634-8521; <sup>3</sup>Division of Biochemistry, Department of Biomedical Sciences, Nihon University School of Medicine, Tokyo 173-8610, Japan

Received May 26, 2018; Accepted November 27, 2018

DOI: 10.3892/ol.2018.9809

**Abstract.** The mitochondrial DNA (mtDNA) displacement loop (D-loop) is often altered in various cancer types, including with regard to simple sequence repeat number variation (SSRNV), which includes the C-tract and CA-tract. However, because of mitochondrial heteroplasmy and slippage errors by the Taq DNA polymerase used in polymerase chain reaction (PCR) analysis, it is difficult to precisely evaluate mtDNA D-loop SSRNV experimentally. In this study, to precisely determine cancer-specific variants in mtDNA SSRNV, various microscopic portions of cancerous tissues and normal control tissues were obtained from a patient with breast cancer, followed by laser-capture microdissection of formalin-fixed paraffin-embedded specimens. Regions containing (CA)<sub>7</sub> repeats (positions 514-523) and (C)<sub>8</sub> repeats (positions 303-315) of the mitochondria DNA D-loop were amplified and sequenced. Variant signals of mtDNA SSRs of (CA)<sub>7</sub> and (C)<sub>8</sub> were observed in normal and cancerous tissues, with the content of minor alleles (CA)<sub>6</sub> and (C)<sub>7</sub>/(C)<sub>9</sub> differing among samples. These results were confirmed by PCR using various

primers and proofreading DNA polymerases. PCR of genomic SSRs of (CA)<sub>7</sub> in the *NAALAD2* gene and (C)<sub>8</sub> in the *BMP6* gene showed a simple repeat in all samples that was different from the observed mtDNA SSRNV. The present study suggests a reliable procedure for determining cancer-specific variants in mtDNA SSRNV: Using a proofreading DNA polymerase for PCR, the background of slippage by PCR is determined by PCR of the same genomic sequence as the target. Due to the varied heteroplasmy level of mtDNA SSRNV among normal tissues, the second background of polymorphic variations should be determined by several normal tissue DNA as PCR templates. Finally, the cancer-specific variant, including its variation frequency, is determined by subtracting the two background signals from the variant signals in cancer. However, care must be taken, as normal heteroplasmy drifts observed in mtDNA SSRNV may complicate such estimations.

## Introduction

Human mitochondrial DNA (mtDNA) is a 16,569-bp circular double-stranded DNA molecule containing genes encoding 13 polypeptides that compose the oxidative phosphorylation apparatus, 2 rRNAs, and a set of 22 tRNAs that are essential for protein synthesis in mitochondria (1). In addition to these coding regions, mtDNA contains a non-coding region called the displacement loop (D-loop) and elements that regulate mtDNA replication as well as promoters for mtDNA transcription. There are many reports of D-loop alterations in various cancers (2-9). In particular, it is well known that a poly-C stretch (C-tract) located between nt 303 and nt 315 (denoted D310) within the D-loop is a hot spot for mutations (2,3,5-8). In general, simple sequence repeat number variation (SSRNV), such as the C-tract and CA-tract, is observed in both sporadic and inherited cancers and is referred to as the genetic instability characteristic of cancer due to the dysfunction of the mismatch repair system (10). Indeed, Lievre *et al* recently suggested that mitochondrial SSRNV is useful for early cancer detection (5). Chatterjee *et al* reviewed that it could not functionally impair mitochondria but could be a molecular marker, not only for early detection of cancer but also prevention, recurrence and monitoring response to therapy (7). On the other hand, it

---

**Correspondence to:** Professor Mariko Esumi or Professor Makoto Makishima, Division of Biochemistry, Department of Biomedical Sciences, Nihon University School of Medicine, 30-1 Ohayaguchikami-cho, Tokyo 173-8610, Japan  
E-mail: esumi.mariko@nihon-u.ac.jp  
E-mail: makishima.makoto@nihon-u.ac.jp

\*Contributed equally

**Abbreviations:** BC, breast cancer; *BMP6*, bone morphogenetic protein 6; D-loop, displacement loop; FFPE, formalin-fixed and paraffin embedded; *NAALAD2*, *N-acetylated α-linked acidic dipeptidase 2*; NB, normal breast epithelia; NLN, normal lymph node; mtDNA, mitochondrial DNA; SSR, simple sequence repeat; SSRNV, simple sequence repeat number variation

**Key words:** genetic instability, heteroplasmy, mismatch repair, mtDNA D-loop, proofreading, repeat number variation, SSR, slippage error, Taq DNA polymerase

has currently been used as a marker for cancer clonality to determine the monoclonal or multicentric origin of cancer (9). However, it is difficult to precisely evaluate cancer-specific SSRNVs of the mtDNA D-loop experimentally because of the following two issues. i) mtDNA can exist in a state of heteroplasmy, even in normal cells, which involves a mixed population of variant and wild-type mtDNAs. Furthermore, the frequency of heteroplasmy varies among tissues within an individual (11,12). Therefore, it is difficult to discriminate variants in cancer from the background heteroplasmy of SSRNVs. ii) Taq DNA polymerase, which is used for amplification of sequencing templates, frequently generates replication errors due to slippage because of the general lack of proofreading capacity (13,14); such artifactual SSRNVs are not distinguishable from the naturally occurring SSRNVs.

In this study, to discriminate between native SSRNVs *in vivo* (including cancer variants and normal polymorphisms) and artifactual variants *in vitro*, we used three validation methods based on causative points of artifactual variants in a case of breast cancer. We found that SSRNV results are reproducible i) by PCR using different primers and ii) by PCR using different proofreading DNA polymerases, and iii) we confirmed the different SSRNV patterns in mtDNA D-loops in different tissues from the same individual. The C-tract and CA-tract in the genome sequence that are the same as a major type of those in the mtDNA D-loop did not vary in all tissue samples. Finally, we found different patterns of heteroplasmy for the mtDNA D-loop in multiple normal tissues and cancer lesions from a single patient, and we herein propose a new validation method for SSRNV.

## Materials and methods

**Samples.** Tissue specimens were obtained from a female patient with breast cancer. Four areas of breast cancer-containing tissues were formalin-fixed and paraffin embedded (FFPE); normal breast (NB) epithelia and two non-tumorous/non-mammary tissues (NLN, Nskin), were obtained from FFPE tissues of the same curative resected specimens for use as normal control tissues of the same patient. Four breast cancers (BC1-BC4) were histopathologically different and obtained by laser capture microdissection. Ten- $\mu$ m-thick sections of NLN were directly subjected to DNA extraction. Other FFPE tissues were subjected to laser capture microdissection using an LMD6000 microscope (Leica Biosystems, Wetzlar, Germany) to precisely obtain cancer cells and normal epithelial cells. DNA was extracted from these FFPE tissues as described previously (15,16); briefly, a RecoverAll Total Nucleic Acid Isolation kit for FFPE (Thermo Fisher Scientific, Inc., Waltham, MA, USA) was used according to the manufacturer's protocol with some modifications such as heat treatment of the lysate at 95°C for 30 min and final purification of DNA by ethanol precipitation. As a control, a fresh, frozen tissue of normal liver was obtained from a non-tumorous portion of curative resected liver cancer, metachronous metastasis of colorectal carcinoma; a patient was 75-year-old female who had been affected with stage II (T3N0M0) colorectal carcinoma. DNA was extracted as described previously (15,16).

The curative surgical operations were performed at Nihon University School of Medicine Itabashi Hospital (Itabashi-ku, Tokyo, Japan) in 2013-2014. The use of tumor and adjacent

non-tumor tissues from the resected specimens to this study was approved by the Ethics Committee of Nihon University School of Medicine (approval nos. 147 and 115). Informed consents were obtained from the patients prior to the start of the study.

**Polymerase chain reaction (PCR).** PCR was performed in 20- $\mu$ l reactions using Takara EX Taq DNA polymerase Hot Start Version (Takara Bio Inc., Shiga, Japan) and KOD-Plus-Neo (Toyobo, Osaka, Japan) according to the manufacturer's instructions. The former is a mixture of recombinant Taq DNA polymerase and 3'-5' exonuclease, and the latter is an enzyme possessing 3'-5' exonuclease activity; this enzyme is derived from the KOD1 strain of *Thermococcus kodakaraensis* and as described in the manufacturer's instructions, has 80-fold fidelity compared to recombinant Taq DNA polymerase. The thermal cycling protocol included 35 to 37 cycles of 98°C for 10 sec and 60°C for 30 sec. The PCR and sequencing primers are shown in Table I. The BLASTN search as described by Simone *et al.* (17) resulted in no nuclear sequence of mitochondrial origin relevant to these PCR target sequences.

**Sanger sequencing.** PCR products were treated with Illustra ExoProStar 1-Step kit (GE Healthcare Life Sciences, Little Chalfont, UK) and subjected to a sequencing reaction using the BigDye Terminator version 1.1 cycle sequencing kit (Thermo Fisher Scientific, Inc.). The reaction products were purified using the BigDye Xterminator kit (Thermo Fisher Scientific, Inc.). Nucleotide sequences were determined by subjecting the purified products to an ABI PRISM 310 Genetic Analyzer (Thermo Fisher Scientific, Inc.) and analyzed using Sequencing Analysis version 5.4 software. The difference in minor allele content of SSRNV was semi-quantitatively determined around the nucleotide position clearly indicating a variant signal in the electropherograms by considering signal ratio to the major allele.

## Results

**(CA) repeat number variation.** SSRNV of (CA)<sub>n</sub> between nt positions 514 and 523 (rCRS NC\_012920.1) was determined by PCR of a 237-bp fragment (F7/R5) (Fig. 1; Table I). The major genotypes of normal tissues, NB and NLN were (CA)<sub>7</sub> in this patient (Fig. 2A). A small amount of the (CA)<sub>6</sub> allele was also observed in NB and NLN, whereas the 3 tumor samples, BC1, BC3 and BC4, exhibited a significant amount of the (CA)<sub>6</sub> allele. The signal of the (CA)<sub>6</sub> allele was as high as that of (CA)<sub>7</sub> in BC3 but was lower in BC1 and BC4 (Fig. 2A). We performed three experiments to examine whether these polymorphisms were due to artifactual slippage by Taq DNA polymerase in PCR. First, a PCR primer was altered to confirm the reproducibility of the amplified sequence. The PCR product using F5/R5 primers (Fig. 1; Table I) revealed the same sequencing electropherogram for all 5 samples, including the varied signal of the (CA)<sub>6</sub> allele (Fig. 2B). For the second experiment, Taq DNA polymerase (from EX Taq polymerase) was exchanged for another proofreading polymerase, KOD DNA polymerase, to confirm the varied signal of the (CA)<sub>6</sub> allele. Both enzymes have proofreading activity, but the KOD DNA polymerase derived from the KOD1 strain of *T. kodakaraensis* naturally contains

Table I. Primer sequences used in the present study.

Gene	Repeat	Primer	Sequence (5'-3')	PCR (Product bp)
mtDNA D-loop	(CA) <sub>7</sub>	F5	TCCCCCGCTTCTGGCCACA	F5R5 (293)
		F7	GAACCCTAACACCAGCCT	F7R5 (237)
		R5	GCTTTGAGGAGGTAAGCTAC	
	(C) <sub>8</sub>	F3	ACCCTATGTCGCAGTATCTG	F3R5 (492)
		R5	GCTTTGAGGAGGTAAGCTAC	
		R10	AGGCTGGTGTTAGGGTTC	
NAALAD2	(CA) <sub>7</sub>	F	AGTCCACAATGTTTTTCACCA	FR (244)
		R	CGCTGGGACTATTCTTTTAGA	
BMP6	(C) <sub>8</sub>	F	TCATTTGAACTGGCTAGGGT	FR (135)
		R	CTCTCATGGAAGTACACATGGT	

PCR, polymerase chain reaction; mtDNA, mitochondrial DNA; D-loop, displacement loop; NAALAD2, *N-acetylated  $\alpha$ -linked acidic dipeptidase 2*; BMP6, *bone morphogenetic protein 6*; F, forward; R, reverse.

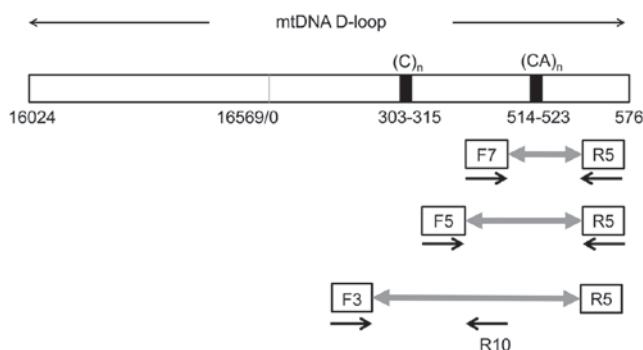


Figure 1. Simple sequence repeat regions of the mtDNA D-loop. Poly-C tract, termed D310 (nt 303 to 315), and (CA)<sub>n</sub> (nt 514 to 523) were amplified with primer pairs F7/R5, F5/R5, and F3/R10. Sequencing primers are shown by one-sided arrows.

3'-5' exonuclease activity as well. KOD PCR produced a similar electropherogram (Fig. 2C) as that of EX Taq PCR (Fig. 1). Although the signal of the (CA)<sub>6</sub> allele in NLN was slightly decreased when using KOD, a minor (CA)<sub>6</sub> allele was clearly present in NLN. Normal liver DNA (NL16) from a different patient showed a (CA)<sub>4</sub> allele, and no polymorphic electropherogram was observed for the PCR using KOD. Thus, the polymorphic (CA) repeat of the mtDNA D-loop in our study of a breast cancer patient was confirmed. NLN and NB exhibit mtDNA heteroplasmy and contain (CA)<sub>7</sub>>>(CA)<sub>6</sub>; the content of the (CA)<sub>6</sub> allele increased as follows: NLN=NB<BC1=BC4<BC3.

The third experiment was performed to disprove the potential sample-dependent manner of artifactual amplification and to confirm the variable content of a minor (CA)<sub>6</sub> allele among the samples. The (CA)<sub>7</sub> repeat-containing genomic sequence was amplified in the same way, and the amplicons were compared to those of the samples. The *N-acetylated  $\alpha$ -linked acidic dipeptidase 2* (NAALAD2) gene displays (CA)<sub>7</sub> at nt position chr11: 89906355-89906369 (hg19), with no repeat number polymorphisms reported to date. A minor (CA)<sub>6</sub> signal was observed at the same level in all electropherograms of

5 samples, including NL16 DNA (fresh-frozen tissue DNA but not FFPE DNA) (Fig. 2D). This minor allele was due to artifactual slippage by EX Taq DNA polymerase not *in vivo* slippage. Thus, different frequencies of a minor (CA)<sub>6</sub> allele of the mtDNA D-loop in different BC samples were confirmed but were not due to artifacts of *in vitro* PCR.

**(C) repeat number variation.** Another polymorphic poly-C tract termed D310 (nt 303 to 315, rCRS NC\_012920.1) was examined in a similar way using EX Taq DNA polymerase followed by Sanger sequencing (Fig. 1; Table I). The patient had (C)<sub>8</sub> as a major allele, but a polymorphic (C)<sub>7</sub> variant was observed in normal tissues, NB, NLN and Nskin (Fig. 3A). Nskin showed the second minor variant (C)<sub>9</sub>. In contrast, NB and BC2 contained a significant amount of the (C)<sub>7</sub> allele, with the highest level. BC3 showed a minor amount of the (C)<sub>7</sub> allele at the lowest level, and the order of the (C)<sub>7</sub> allele content was BC3<NLN=Nskin<<NB=BC2 (Fig. 3A). Both BC1 and BC4 also exhibited low (C)<sub>7</sub> allele frequencies, similar to NLN (data not shown).

To ascertain the variable content of the (C)<sub>7</sub> allele in different tissues from a single patient, the same (C)<sub>8</sub>T (C)<sub>6</sub> sequence-containing genomic sequence was amplified and sequenced. The *bone morphogenetic protein 6* (BMP6) gene contains a (C)<sub>8</sub>T (C)<sub>6</sub> sequence at nt position chr6:7801139-7801153 (hg19), with no repeat number polymorphisms reported. Although some slight noise was observed in all 4 samples, the signals were clearly different from slippage signals (Fig. 3B). Thus, the different pattern and frequency of D310 variations were validated in normal and cancerous regions of tissue from the examined breast cancer patient. Differential heteroplasmy of the mtDNA D-loop was observed in different normal tissues, and enhancement of heteroplasmy and homoplasmy was observed in various cancerous regions of a single patient.

## Discussion

We clearly demonstrate in this study that SSRNV of the mtDNA D-loop is present in normal tissues and that the variant pattern



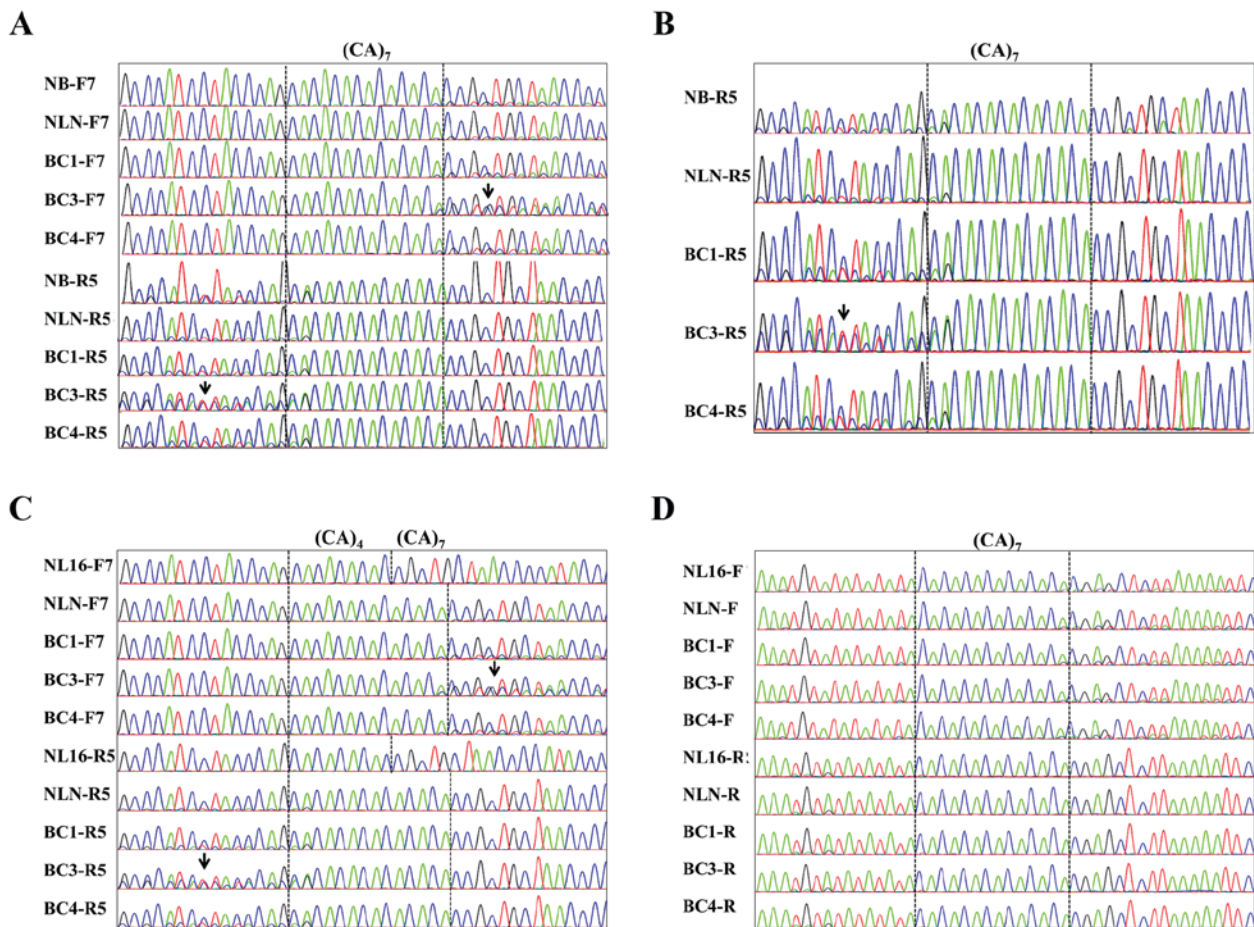


Figure 2. Sequencing electropherogram of  $(CA)_7$ . PCR of the mtDNA D-loop  $(CA)_7$  was performed using EX Taq DNA polymerase and F7/R5 primer pairs (A) or F5/R5 primer pairs (B). The sequencing electropherogram using the F5 primer is the same as Fig. 2B (data not shown). (C) PCR was performed using a different proofreading DNA polymerase KOD and F7/R5 primer pairs. (D) PCR of genomic  $(CA)_7$  of the *NAALAD2* gene was performed using EX Taq DNA polymerase. An arrow indicates a representative position used to evaluate a variant allele. PCR, polymerase chain reaction; mtDNA, mitochondrial DNA; NB, normal breast epithelia; NLN, normal lymph node; BC, breast cancer; NL16, normal liver; F, forward; R, reverse. NL16 is a different patient with a  $(CA)_4$  repeat.

and frequency vary in both normal and cancerous tissues. In particular, to avoid the artifactual slippage generated by Taq DNA polymerase during pre-amplification of the SSR region, we used two different proofreading DNA polymerases for PCR in this study. In fact, these enzymes reproducibly resulted in the same variant signal patterns of mtDNA SSRs of  $(CA)_7$  and  $(C)_8$  in individual samples. When amplifying genomic SSRs of  $(CA)_7$  in the *NAALAD2* gene (Fig. 2D) and  $(C)_8$  in the *BMP6* gene (Fig. 3B) using the same proofreading DNA polymerase as for the mtDNA D-loop, a slight or no slippage signal was observed, and cancer-specific mtDNA variant signals were confirmed. In general, most variant signals in cancer are so weak that it is difficult to determine if there is a real variant. Moreover, the heteroplasmy issue of mtDNA in normal tissues complicates determination of the presence of a cancer-specific variant. Howell and Smejkal also validated mtDNA SSRNV in Leber hereditary optic neuropathy using a proofreading enzyme: Pfu polymerase (18). However, it remains to be examined whether background slippage actually occurred *via* amplification of the same SSR sequence in the genome, as in the present study. Based on our results, we suggest a reliable procedure for determining cancer-specific variants in mtDNA SSRNV: i) use of a proofreading DNA polymerase for PCR; ii) assessment of background slippage by PCR using the same genomic sequence as

the target; iii) determination of the second background of polymorphic variations using DNA from several normal tissues as PCR templates; iv) and evaluation of the cancer-specific variant, including the variation frequency, by subtracting the two background signals from the variant signals in cancer. According to this procedure, the present case of breast cancer is summarized and considered as follows: When four cancer lesions were compared with three normal tissues, the minor allele of  $(CA)_6$  was found to be increased in three of the cancer tissue samples (BC1, BC3, BC4), with the highest level in BC3 (Fig. 2). BC2 exhibited the low level, similar to three normal tissues (data not shown). In contrast, the minor allele of  $(C)_7$  was increased only in BC2 (Fig. 3A). Compared to NB epithelial cells, the minor allele was decreased in the other three cancer samples (BC1, BC3, BC4) (Fig. 3A) (data not shown). These results suggest that cancer cell mtDNA is variable in individual cancer lesions, and interestingly that the minor allele frequencies are inversely correlated between the two SSR regions.

Taq DNA polymerase slippage has often been documented in microsatellite analysis for genetic instability in cancer (19-21). For example, Shinde *et al* thoroughly investigated Taq DNA polymerase slippage variant rates by PCR and quasi-likelihood analysis (14). Slippage is dependent on the repeat number: The more the repeat number increases, the higher the slippage rate

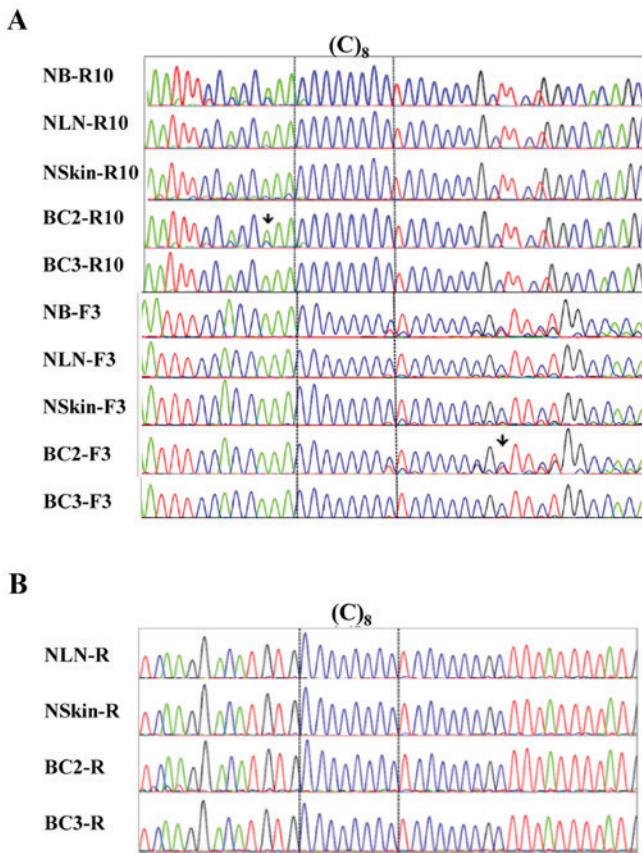


Figure 3. Sequencing electropherogram of  $(C)_8$ . (A) PCR of the mtDNA D-loop  $(C)_8$  was performed using EX Taq DNA polymerase and F3/R5 primer pairs. Inner reverse primer R10 and PCR forward primer F3 were used for sequencing. (B) PCR of genomic  $(C)_8$  of the *BMP6* gene was performed using EX Taq DNA polymerase. The sequencing electropherogram using the F primer is the same as Fig. 3B (data not shown). An arrow indicates a representative position used to evaluate a variant allele. PCR, polymerase chain reaction; NB, normal breast epithelia; NLN, normal lymph node; NSkin, normal skin; BC, breast cancer; F, forward; R, reverse.

becomes. The threshold for detecting slippage products is 4-5 repeats of  $(CA)_n$  tracts and 8 repeats of  $(A)_n$  tracts, and the ratio of contractions to expansions is 14-fold for  $(CA)_n$  tracts and 5-fold for  $(A)_n$  tracts. Thus, a decrease in repeat number is preferred to an increase in slippage events of any type of microsatellite repeat. Although similar slippage events occur during DNA replication *in vivo*, most replication errors are repaired by the mismatch repair system in both the nucleus and mitochondria (22-24). DNA polymerase  $\gamma$  carries out mtDNA replication with proofreading exonuclease activity and functions in the mismatch repair system (22,24). However, polymerase  $\gamma$  is prone to replication errors when copying homopolymeric sequences longer than 4 nt (22). Given the markedly increased number of mtDNA molecules compared to that of the nuclear genome, such variations lead to the coexistence of polymorphic mtDNA molecules. Therefore, heteroplasmy of mtDNA is likely to be present even in normal cells (25). Deep sequencing also revealed more universal heteroplasmy than expected in normal cells (25-27). Nonetheless, malfunction of the mismatch repair system occurs in cancer cells, which results in mitochondrial genome instability, *i.e.*, an increase in the error frequency of mtDNA SSRNV; however, this is not distinguishable from a shift in heteroplasmy levels via segregation of heteroplasmic mtDNA during cancer

cell division. Accordingly, it may also be due to a heteroplasmic shift and not necessarily to mitochondrial genome instability that the proportion of major and minor mtDNA SSRNVs differ in different cancerous tissues, as shown in the present study.

Thus, based on the difference in heteroplasmy level of mtDNA in normal somatic cells, it is difficult to estimate mtDNA genome instability in cancer. The mtDNA D310 SSRNV has been used for assessing mtDNA genome instability in multiple cancers; compared to normal tissues, cancer variants are homoplasmic or heteroplasmic (3,4,7-9,28). Because homoplasmy and various degrees of mtDNA heteroplasmy exist in normal cells within a given individual and because cancer can originate from any normal cell, a cancer-specific variant is difficult to delineate. To determine mitochondrial genome instability in cancer, it will be helpful to apply multiple single-cell next-generation sequencing to analyze whether normal and cancerous tissues with mtDNA heteroplasmy consist of a mixture of cells with different homoplasmic mtDNAs or a mixture of cells with heteroplasmic mtDNA. When we analyzed mtDNA D310 by deep sequencing using a database of single-cell sequencing of a breast cancer cell line and a matched normal B-lymphoblast cell line from a patient (29), three cells each of the two cell lines showed  $(C)_7/(C)_8$  heteroplasmy. The ratio of the minor  $(C)_8$  allele was 0.2, 0.15, and 0.15 in three B-lymphoblasts and 0.05, 0.15, and 0.11 in three cancer cells. These results suggest that mtDNA heteroplasmy of both normal and cancer cells occurs at the single-cell level but does not involve a mixture of homoplasmic cells; the proportion of heteroplasmy level is relatively constant among cell lines, but that of cancer cells can be reduced during cell division, even in a single clone. Considering the results of the present study and single-cell analyses, mtDNA SSRNV is not suitable for estimating genetic instability in cancer. Although the mtDNA D310 variant is currently used to determine clonal assessment of clinical multiple cancers, either synchronous or metachronous (9), it appears to be helpful but not definitive.

In conclusion, the present study proposes a reliable procedure for determining mtDNA SSRNV and cancer-specific variants. Using a proofreading DNA polymerase for PCR, the background of slippage by PCR is determined via PCR of the same genomic sequence as the target. Due to the varied heteroplasmy level of mtDNA SSRNV among normal tissues, the second background of polymorphic variations should be determined by using DNA extracted from several normal tissues as PCR templates. Finally, the cancer-specific variant, including its variation frequency, is determined by subtracting the two background signals from the variant signals in cancer; regardless, normal heteroplasmy drifts appear to complicate such estimation. To further evaluate the procedure and to confirm the cancer heterogeneity, there remains to study in multiple cases of cancer.

#### Acknowledgements

Not applicable.

#### Funding

The present study was supported in part by Nihon University Multidisciplinary Research Grant (grant no. M14-012) from Nihon University received by ME.



## Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to inclusion in another article, which is being prepared for submission, but are available from the corresponding author on reasonable request.

## Authors' contributions

TN, AS, TE and HK conducted the experimental work. AS and TE wrote the manuscript. HK and SM pathologically diagnosed tissue samples and interpreted the data. ME designed the study, and reviewed and edited the manuscript. MM reviewed the study design, interpreted data, and critically revised the manuscript for medical and scientific content. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

The present study was approved by the Ethics Committee of Nihon University School of Medicine (approval nos. 147 and 115). Informed consents were obtained from patients prior to the start of the study.

## Patient consent for publication

Written informed consents were obtained from patients for publication.

## Competing interests

The authors declare that they have no competing interests.

## References

- Wallace DC: Mitochondrial DNA sequence variation in human evolution and disease. *Proc Natl Acad Sci USA* 91: 8739-8746, 1994.
- Sanchez-Cespedes M, Parrella P, Nomoto S, Cohen D, Xiao Y, Esteller M, Jeronimo C, Jordan RC, Nicol T, Koch WM, *et al*: Identification of a mononucleotide repeat as a major target for mitochondrial DNA alterations in human tumors. *Cancer Res* 61: 7015-7019, 2001.
- Suzuki M, Toyooka S, Miyajima K, Iizasa T, Fujisawa T, Bekele NB and Gazdar AF: Alterations in the mitochondrial displacement loop in lung cancers. *Clin Cancer Res* 9: 5636-5641, 2003.
- Lee HC, Li SH, Lin JC, Wu CC, Yeh DC and Wei YH: Somatic mutations in the D-loop and decrease in the copy number of mitochondrial DNA in human hepatocellular carcinoma. *Mutat Res* 547: 71-78, 2004.
- Lievre A, Blons H, Houllier AM, Laccourreye O, Brasnu D, Beaune P and Laurent-Puig P: Clinicopathological significance of mitochondrial D-loop mutations in head and neck carcinoma. *Br J Cancer* 94: 692-697, 2006.
- Chen D and Zhan H: Study on the mutations in the D-loop region of mitochondrial DNA in cervical carcinoma. *J Cancer Res Clin Oncol* 135: 291-295, 2009.
- Chatterjee A, Dasgupta S and Sidransky D: Mitochondrial subversion in cancer. *Cancer Prev Res (Phila)* 4: 638-654, 2011.
- Guo W, Yang D, Xu H, Zhang Y, Huang J, Yang Z, Chen X and Huang Z: Mutations in the D-loop region and increased copy number of mitochondrial DNA in human laryngeal squamous cell carcinoma. *Mol Biol Rep* 40: 13-20, 2013.
- Geurts-Giele WR, Gathier GH, Atmodimedjo PN, Dubbink HJ and Dinjens WN: Mitochondrial D310 mutation as clonal marker for solid tumors. *Virchows Arch* 467: 595-602, 2015.
- Harfe BD and Jinks-Robertson S: DNA mismatch repair and genetic instability. *Annu Rev Genet* 34: 359-399, 2000.
- Naue J, Hörer S, Sängler T, Strobl C, Hatzler-Grubwieser P, Parson W and Lutz-Bonengel S: Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. *Mitochondrion* 20: 82-94, 2015.
- Stewart JB and Chinnery PF: The dynamics of mitochondrial DNA heteroplasmy: Implications for human health and disease. *Nat Rev Genet* 16: 530-542, 2015.
- Cline J, Braman JC and Hogrefe HH: PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res* 24: 3546-3551, 1996.
- Shinde D, Lai Y, Sun F and Arnheim N: Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)<sub>n</sub> and (A/T)<sub>n</sub> microsatellites. *Nucleic Acids Res* 31: 974-980, 2003.
- Einaga N, Yoshida A, Noda H, Suemitsu M, Nakayama Y, Sakurada A, Kawaji Y, Yamaguchi H, Sasaki Y, Tokino T and Esumi M: Assessment of the quality of DNA from various formalin-fixed paraffin-embedded (FFPE) tissues and the use of this DNA for next-generation sequencing (NGS) with no artifactual mutation. *PLoS One* 12: e0176280, 2017.
- Nakayama Y, Yamaguchi H, Einaga N and Esumi M: Pitfalls of DNA quantification using DNA-binding fluorescent dyes and suggested solutions. *PLoS One* 11: e0150528, 2016.
- Simone D, Calabrese FM, Lang M, Gasparre G and Attimonelli M: The reference human nuclear mitochondrial sequences compilation validated and implemented on the UCSC genome browser. *Bmc Genomics* 12: 517, 2011.
- Howell N and Smejkal CB: Persistent heteroplasmy of a mutation in the human mtDNA control region: Hypermutation as an apparent consequence of simple-repeat expansion/contraction. *Am J Hum Genet* 66: 1589-1598, 2000.
- Alvi AJ, Rader JS, Broggin M, Latif F and Maher ER: Microsatellite instability and mutational analysis of transforming growth factor beta receptor type II gene (TGFBR2) in sporadic ovarian cancer. *Mol Pathol* 54: 240-243, 2001.
- Maehara Y, Oda S and Sugimachi K: The instability within: Problems in current analyses of microsatellite instability. *Mutat Res* 461: 249-263, 2001.
- Enomoto A, Esumi M, Yamashita K, Takagi K, Takano S and Iwai S: Abnormal nucleotide repeat sequence in the TGF-betaRII gene in hepatocellular carcinoma and in uninvolved liver tissue. *J Pathol* 195: 349-354, 2001.
- Longley MJ, Nguyen D, Kunkel TA and Copeland WC: The fidelity of human DNA polymerase gamma with and without exonucleolytic proofreading and the p55 accessory subunit. *J Biol Chem* 276: 38555-38562, 2001.
- Mason PA, Matheson EC, Hall AG and Lightowlers RN: Mismatch repair activity in mammalian mitochondria. *Nucleic Acids Res* 31: 1052-1058, 2003.
- Kazak L, Reyes A and Holt IJ: Minimizing the damage: Repair pathways keep mitochondrial DNA intact. *Nat Rev Mol Cell Biol* 13: 659-671, 2012.
- Carelli V, Maresca A, Caporali L, Trifunov S, Zanna C and Rugolo M: Mitochondria: Biogenesis and mitophagy balance in segregation and clonal expansion of mitochondrial DNA mutations. *Int J Biochem Cell Biol* 63: 21-24, 2015.
- Larsson NG: Somatic mitochondrial DNA mutations in mammalian aging. *Annu Rev Biochem* 79: 683-706, 2010.
- Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, Taylor RW, Samuels DC, Santibanez-Koref M and Chinnery PF: Universal heteroplasmy of human mitochondrial DNA. *Hum Mol Genet* 22: 384-390, 2013.
- He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, Diaz LA Jr, Kinzler KW, Vogelstein B and Papadopoulos N: Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* 464: 610-614, 2010.
- Szulwach KE, Chen P, Wang X, Weaver LS, Gonzales ML, Sun G, Unger MA and Ramakrishnan R: Single-cell genetic analysis using automated microfluidics to resolve somatic mosaicism. *PLoS One* 10: e0135007, 2015.