Published online 17 September 2019

Nucleic Acids Research, 2019, Vol. 47, No. 21 ***e141***
doi: 10.1093/nar/gkz811

Low temperature isothermal amplification of microsatellites drastically reduces stutter artifact formation and improves microsatellite instability detection in cancer

Antoine Daunay[1], Alex Duval[2], Laura G. Baudrin[1,3], Olivier Buhard[2], Victor Renault[4], Jean-François Deleuze[1,5] and Alexandre How-Kit[1,*]

[1]Laboratory for Genomics, Foundation Jean Dausset – CEPH (Centre d'Etude du Polymorphisme Humain), Paris, France, [2]Sorbonne-Université, Université Pierre et Marie Curie – Paris 6, Paris, France, INSERM, UMRS 938-Centre de Recherche Saint-Antoine, Equipe 'Instabilité des Microsatellites et Cancers', Equipe labellisée par la Ligue Nationale contre le Cancer, and SIRIC CURAMUS, Paris, France Université Pierre et Marie Curie, Paris, France, [3]Laboratory of Excellence GenMed, Paris, France, [4]Laboratory for Bioinformatics, Foundation Jean Dausset – CEPH (Centre d'Etude du Polymorphisme Humain), Paris, France and [5]Centre National de Recherche en Génomique Humaine, CEA-Institut François Jacob, Evry, France

Received June 13, 2019; Revised September 03, 2019; Editorial Decision September 09, 2019; Accepted September 11, 2019

ABSTRACT

Microsatellites are polymorphic short tandem repeats of 1–6 nucleotides ubiquitously present in the genome that are extensively used in living organisms as genetic markers and in oncology to detect microsatellite instability (MSI). While the standard analysis method of microsatellites is based on PCR followed by capillary electrophoresis, it generates undesirable frameshift products known as 'stutter peaks' caused by the polymerase slippage that can greatly complicate the analysis and interpretation of the data. Here we present an easy multiplexable approach replacing PCR that is based on low temperature isothermal amplification using recombinase polymerase amplification (LT-RPA) that drastically reduces and sometimes completely abolishes the formation of stutter artifacts, thus greatly simplifying the calling of the alleles. Using HT17, a mononucleotide DNA repeat that was previously proposed as an optimal marker to detect MSI in tumor DNA, we showed that LT-RPA improves the limit of detection of MSI compared to PCR up to four times, notably for small deletions, and simplifies the identification of the mutant alleles. It was successfully applied to clinical colorectal cancer samples and enabled detection of MSI. This easy-to-handle, rapid and cost-effective approach may deeply improve the analysis of microsatellites in several biological and clinical applications.

INTRODUCTION

Microsatellites are DNA sequences also known as short tandem repeats (STRs) or simple sequence repeats (SSRs) located throughout the genome that are composed of a unit of 1–6 nucleotides repeated in tandem and whose polymorphism is based on the number of repetitions (1). These highly polymorphic sequences have extensively been used for ∼30 years in humans, animals and plants as genetic markers for applications as diverse as gene mapping (2), conservation and population genetic studies (3), marker-assisted selection (4), kinship analysis (5), forensic DNA fingerprinting (6) or cancer diagnosis (7). In the last mentioned application microsatellite genotyping is used to detect microsatellite instability (MSI), a widespread alteration present in cancer caused by DNA mismatch repair system deficiency (dMMR) and characterized by the accumulation of mutations (insertions or deletions of several nucleotides that are multiples of the length of the repeated unit) in microsatellites (7–9). MSI has a great clinical interest for cancer patients as it bears prognostic information and is also a predicator of the efficacy of immune checkpoint blockade therapy (8,10,11). In colorectal cancer (CRC), about 15% of the tumors present MSI with a sporadic origin or arising in the context of Lynch syndrome (7). Moreover, MSI testing

*To whom correspondence should be addressed. Tel: +33 1 53725183; Email: alexandre.how-kit@fjd-ceph.org

is recommended for every CRC patients as a first screen for Lynch syndrome (12).

The gold standard approach for the genotyping of microsatellites for all the above-mentioned applications relies on a PCR amplification of microsatellite loci with fluorescent primers followed by capillary electrophoresis fragment analysis (FA) (13,14). More recently, the analysis of microsatellites has also been performed using next generation sequencing (NGS) of PCR-amplified libraries (8,15–17), thereby allowing a higher throughput of analyzed microsatellite loci but requiring the development of different bioinformatics algorithms (18–25). While PCR allows specific and rapid amplification of microsatellites, it also induces the presence of undesirable frameshift products known as stutter or shadow bands/peaks (26–29). These artifacts, whose formation has been attributed to slipped-strand mispairing during PCR, are multiples of the repeated nucleotide sequence and are generally shorter than the specific amplification product (26–29). Moreover, the length of the repeat unit and the number of repetitions of the repeat unit have been shown to influence negatively and positively the formation of stutter respectively, so that long mononucleotide microsatellites presented generally a higher number of stutter peaks (1,24,29–31).

The presence of these stutter peaks has rendered difficult the manual and automatic genotype calling of the microsatellites, and even more difficult in some complex biological and clinical contexts such as the genetic fingerprinting of mixed or low template DNA samples in forensics (32) or the sensitive detection and identification of mutant microsatellite alleles in microsatellite unstable (MSI) cancer (14,33). Some attempts to decrease the formation of stutter peaks during the PCR amplification have been made, including the modification of pH, dNTP and Mg^{2+} concentrations or the use of different proofreading polymerases, but with limited success (29,34). Only the use of a lower elongation temperature at 37°C during the PCR has shown a decrease of stutter formation, but requires a new addition of a thermolabile DNA polymerase after each denaturation step, which is laborious, time-consuming and incompatible for most applications (34). Thus, in order to deal with stutter peaks and to overcome the difficulties of microsatellite allele calling due to them, several complex post-PCR analysis approaches have been developed for fragment analysis and next generation sequencing data, notably in the fields of forensics and cancer diagnosis (14,33,35–40), while the PCR remained the standard method for microsatellite amplification.

Several approaches based on isothermal amplification of nucleic acids have been developed as alternatives to PCR (41). Among them recombinase polymerase amplification (RPA) is a very simple, rapid and convenient isothermal amplification approach for the amplification of DNA performed between 37 and 42°C whose applications to date have focused principally on the detection of pathogens (virus, bacteria, fungi and parasites) or genetically modified crops (42,43). The principle of RPA relies on the repeated action of recombinase proteins that bind to primers and guide them to the complementary DNA sequence where they hybridize (42,43). The displaced DNA strand is then stabilized by single-strand DNA binding proteins (SSBP) while the primer extension is mediated by a DNA polymerase with strand-displacement activity resulting in isothermal amplification of DNA (42,44).

In the present study and for the first time, we replaced PCR by low temperature isothermal amplification based on RPA for the amplification of microsatellites. We first compared the formation of stutter peaks between PCR and RPA using the stutter ratio (SR) and HT17, a T_{17} mononucleotide repeat microsatellite of great clinical interest in oncology. It presents a higher sensitivity for MSI detection compared to the gold standard pentaplex panel (45) and also carries prognostic and predictive information of the response to chemotherapy (5-FU and oxaliplatin) (46,47). The effects of several RPA reaction parameters ($Mg(OAc)_2$ concentration, temperature, duration and DNA template quantity) were evaluated on SR to identify the optimal reaction conditions. We further tested whether low-temperature RPA (LT-RPA) was applicable to several other mono- to tetra-nucleotide repeat microsatellites, including tri-nucleotide repeat microsatellites known to form secondary structures, and measured the SR reduction compared to PCR. The ability of multiplexing of LT-RPA was also assessed as well as the improvement of allele calling in four blood samples genotyped by LT-RPA combined to capillary electrophoresis or amplicon sequencing. Finally we evaluated LT-RPA in the context of disease diagnosis, including the detection of expanded triplet repeats in Huntington disease and MSI in cancer. The limit of detection of MSI obtained with LT-PRA and PCR was estimated and compared using dilution series of cancer cell lines bearing different HT17 mutations. Moreover, HT17 LT-RPA assay was tested on several fresh frozen and FFPE tumor samples of colorectal cancer patients to evaluate its potential application to MSI detection in routine clinical testing.

MATERIALS AND METHODS

Cell lines and dilution series

The CEPH lymphoblastoid cell line 0144711 (NA12762) from a male donor of CEU/UTAH population bearing a T17/T17 homozygous genotype for HT17 was used for all optimization experiments. For dilution series, different CEPH lymphoblastoid cell lines from individuals of CEU/UTAH population bearing T16/T16 homozygous genotypes for HT17 were used including 0145412 (NA12812), 0145413 (NA12813), 0145414 (NA12814) and 0145613 (NA12829). All DNA of CEPH cell lines were provided by the CEPH Biobank.

Four CRC cell lines with MSI phenotype harboring different HT17 deletions were used in this study: SNU-1 (T15/T15), HCT-15 (T13/T14), LoVo (T12/T12) and Co-115 (T10/T11). Each cell line DNA was diluted serially into DNA from CEPH lymphoblastoid cell lines harboring T16/T16 genotypes to the following fractions of mutations: 50%, 25% 10%, 5%, 1%.

Human blood samples

Peripheral blood samples were obtained from the French blood bank, EFS (Etablissement Français du Sang, Paris, France) and were derived from four healthy donors. Buffy

coats were obtained from blood after 10 min centrifugation at 1600g and frozen at −80°C before DNA extraction. DNA extraction was performed using the QIAmp DNA blood mini Kit (Qiagen) on a QIAcube robotic workstation (Qiagen) according to the manufacturer's instructions. Informed consent was obtained from all donors and all methods were performed in accordance to French laws and recommendations of the French National Committee of Ethics (Comité Consultatif National d'Ethique pour les Sciences de la Vie et de la Santé).

Patient colorectal cancer samples

The cohort consisted of 20 CRC patients (12 fresh frozen and 8 FFPE tumor samples) with well-differentiated colorectal cancer who underwent surgery at Saint-Antoine hospital, Paris, France. Tumor samples were snap-frozen in liquid nitrogen or formalin-fixed and paraffin embedded and stored in the Biobank of this Hospital. DNA isolation was performed using the QIAamp DNA Mini Kit (Qiagen) or QIAamp DNA FFPE kit (Qiagen) for fresh frozen or FFPE samples respectively, according to the manufacturer's instructions. The use of clinical and pathologic records was in agreement with French laws and ethical guidelines related to the protection of the patient.

DNA quantification and primer design

The DNA quantification was performed using nanodrop 2000c (Thermo scientific, Wilmington/USA) or the Qubit™ dsDNA HS assay Kit on a Qubit 3 Fluorometer (Thermo Fischer Scientific) according to the manufacturer's instructions. All the primers used in this study have been either designed with Beacon Designer 8 (PREMIER Biosoft) or taken from previous publications (Supplementary Table S1) and were purchased from EurofinsGenomics.

Polymerase chain reaction (PCR)

Standard real-time PCR was performed on a LightCycler 480 II thermocycler (Roche Applied Science). From 50 pg to 20 ng of genomic DNA were used as template in a 20 μl PCR mix including 1x HotStarTaq DNA polymerase Buffer, 1.6 mM of additional $MgCl_2$, 200 μM of dNTPs, 200 nM of each primer, 2 μM of SYTO9 as fluorescent dye (Invitrogen) and 0.8 U of HotStarTaq DNA polymerase (Qiagen). An initial denaturation step was performed for 10 min at 95°C, followed by 50 cycles of 30 s denaturation at 95°C, 30 s annealing (58–68°C) and 20 s elongation at 72°C. The final step included a melting curve (five acquisitions per degree) from 65 to 95°C. For multiplex PCR 200 nM of each primer (HT17, NR24 and BAT26) was added. For small-pool PCR, the reactions were performed in 384-well PCR plates in a final reaction volume of 10 μl using about 1 pg of DNA per reaction and the same PCR conditions described above.

Recombinase polymerase amplification (RPA)

Isothermal amplification of DNA by RPA was performed using TwistAmp™ Liquid basic or TwistAmp™ basic Kits (TwistDx™). From 500 pg to 10 ng of DNA were used as template in the RPA reaction in a final volume of 10 μl. The RPA mix using TwistAmp™ Liquid basic Kit included 1× Reaction buffer, 230 μM dNTP, 1x Basic E-mix, 480 nM of each primer added in that order followed by homogenization with pipetting, supplementation with 1x CORE Reaction Mix and template DNA. From 5 to 14 mM $Mg(OAc)_2$ were added in the reaction followed by 10–40 min incubation at 25–42°C in a preheated Mastercycler Pro S (Eppendorf) and 2 min at 95°C for enzyme inactivation. For TwistAmp™ basic Kit RPA reaction, the dried pellets, which are provided for a reaction volume of 50 μl, were resuspended in 29.5 μl of buffer. The mix was aliquoted in 8.5 μl and supplemented with 5 ng of DNA and 10.5 mM of $Mg(OAc)_2$ for a final reaction volume of 10 μl followed by 40 min incubation at 32°C in a preheated Mastercycler Pro S (Eppendorf) and 2 min at 95°C for enzyme inactivation. For multiplex RPA reaction, 200 nM of NR24 and BAT26 primers and 100 nM of HT17 primers were added in the RPA reaction.

Genotyping by capillary electrophoresis fragment analysis

Genotyping of microsatellites was performed on a 3500 Genetic Analyzer (Applied Biosystems) using POP-7 polymer™ and 50 cm capillary array. Genotyping reaction included 0.5 μl of 1:25–1:40 PCR or 1:10–1:80 RPA product, 0.5 μl of 1:4 GENESCAN-600 LIZ size standard and 9 μl of Hi-Di™ Formamide (Applied Biosystems) and was heated 5 min at 95°C then cooled on ice before injection. Fragment length analysis was performed using the Microsatellite Default analysis module of GeneMapper® software v5.0 (Applied Biosystems). The stutter ratio (SR) was calculated as the ratio of peak intensities of the $n-1$ stutter over the n genuine allele expressed in percentage.

DNA amplicon sequencing

PCR (20 ng of DNA) and LT-RPA (5 ng of DNA, 10.5 mM of $Mg(OAc)_2$ and 40 min incubation at 32°C) amplicons of HT17, NR24, CAT25, BAT26, D2S123, D18S61, D12ATA63, REN and HPRTII microsatellites (shorter primers were used for REN and HPRTII and are indicated in Supplementary Table S1) obtained for each blood sample were purified using Beckman Coulter™ Agencourt AMPure XP beads (Thermo Fischer Scientific), quantified and pooled in equimolar ratio. 900 ng of each pool were used for the ligation of dual-indexed adapters using QIAseq 1-Step Amplicon Library Kit (Qiagen) and no supplemental PCR amplification of the libraries was required. Libraries were assessed for quality and quantity with a Fragment Analyzer (Agilent) and QIAseq™ Library Quant Assay Kit (Qiagen) respectively. 50 pM of the pooled libraries were deposited on an iSeq 100 cartridge (Illumina) together with 20% of 50 pM PhiX control v3 Library (Illumina). Amplicon sequencing was performed on an iSeq 100 using 151 cycles of paired-end sequencing. FastQ files were generated for each sample using Local Run Manager Software (Illumina) and the read counts and sequences of each microsatellite allele were obtained from the aligned reads using an in-house developed Python code.

RESULTS

Effects of RPA conditions on stutter formation in HT17 microsatellite

To evaluate the impact of PCR (Supplementary Figure S1A) and RPA (Supplementary Figure S1B) on stutter formation, we compared the SR obtained by FA following both DNA amplification approaches in different reaction conditions with the same primers amplifying HT17, a mononucleotide microsatellite, and using CEPH 0144711, a lymphoblastoid cell line harboring a T17/T17 genotype.

Four DNA polymerases were evaluated for PCR amplification of HT17 microsatellite using two primer pairs (Supplementary Figure S2). Contrary to Phusion U Hot Start and Q5 High-Fidelity DNA polymerases, HotStarTaq and Platinium *Taq* DNA polymerases systematically introduced $n-1/n(S)-1$ stutter peaks of lesser intensity than $n/n(S)$ genuine allele peaks in samples of any genotype, thereby allowing easy genotype calling (Supplementary Figure S2). We chose the HotStarTaq DNA polymerase for all other PCR experiments.

For RPA, we first used 5 ng of DNA template in 10 μl reaction volume and varied the concentration of $Mg(OAc)_2$ under different temperatures ranging from 25 to 42°C (Figure 1A, Supplementary Figure S3). While the SR was above 88% in our standard PCR conditions (20 ng of DNA template in a final volume reaction of 20 μl), a strong decrease of the SR ranging from 21% to 67% was observed for every RPA condition tested SR (Figure 1A and B). The results showed that the SR was directly under the influence of temperature and of $Mg(OAc)_2$ concentration (Figure 1A). The SR gradually decreased with decreasing RPA temperature with the lowest SR value of 21% obtained at 27°C, while the decrease of $Mg(OAc)_2$ concentrations increased the SR for RPA temperature below 32°C or over 37°C (Figure 1A).

We also evaluated the effect of the DNA template quantity on RPA and on stutter formation compared to PCR using 10 ng to 50 pg of DNA input. We observed an increase of the SR for decreasing quantity of template DNA under PCR, i.e. 89% for 10 ng of DNA to 106% with 50 pg of DNA, associated with an increased variability (Figure 1B, Supplementary Figure S4A). Thus, the genuine n allele peak could be lower than the $n-1$ stutter peak for DNA input <500 pg, which could complicate the correct genotype calling for HT17 (Figure 1B, Supplementary Figure S4A). On the contrary, when RPA was used on decreasing quantity of DNA template, the SR remained constant (∼ 24%) and the genotyping of HT17 remained easy in spite of an increased variability of the SR for inputs 500 pg (Figure 1C, Supplementary Figure S4A). Finally, we evaluated the effect of different durations of the RPA reaction on the stutter formation and observed that the SR decreased with the increase of the duration of the reaction (Figure 1D, Supplementary Figure S4B).

Taken together these results indicate that low temperature RPA have reduced the formation of stutter peaks in HT17 and that a minimum duration of at least 30 min is required to achieve the lowest SR. Contrary to PCR, LT-RPA can also amplify HT17 with inputs as low as 50 pg without any shift of the majority allele. Thus, the optimal HT17 RPA reaction included 5 ng of DNA template with 10.5 mM of $Mg(OAc)_2$ at 32°C during 40 min. The microsatellite profiles obtained by LT-RPA are greatly simplified for every HT17 genotype allowing an easy allele calling (Figure 1E).

Evaluation of LT-RPA on stutter formation using different mono- to tetra-nucleotide repeat microsatellites

In order to evaluate the effect of RPA on the formation of stutter peaks compared to PCR in microsatellites other than HT17, we performed RPA on mono- (NR24, CAT25 and BAT26), di- (D18S61 and D2S123), tri- (D12ATA63) and tetra- (REN and HPRTII) nucleotide repeat microsatellites using the same CEPH cell line. When PCR was used, SRs of variable magnitudes were observed ranging from 5% to 120% that varied according to the type of microsatellite (mono- to tetra- nucleotide repeat) and to the number of repetitions (Figure 2A and B). Thus, the mono- and di-nucleotide microsatellites presented higher SR (≥ 70%) compared to tri- to tetra-nucleotide microsatellites (≤20%) while smaller S alleles presented lower SR compared to larger L alleles (Figure 2A).

For the LT-RPA of the microsatellites 5 ng of DNA were used with 10.5 mM $Mg(OAc)_2$ at 32°C during 40 min. Lower SRs ($0\% \leq SR \leq 49\%$) were observed for all tested microsatellites when the amplification was performed with LT-RPA compared to PCR (Figure 2A). Moreover the formation of stutter peaks was sometimes completely abolished for the small S allele of D18S61 and for the tri- and tetra-nucleotide repeat microsatellites (D12ATA63, REN and HPRTII), while it remained quite high for CAT25 (49%) and BAT26 (37%), that are two long mononucleotide microsatellites (Figure 2A). While the differences of SR between PCR and LT-RPA were higher for mono- and di-nucleotide repeat microsatellites (Supplementary Figure S5A), the SR reductions were the highest for the tri- and tetra-nucleotide repeat microsatellites (Supplementary Figure S5B). Thus, LT-RPA reduced the stutter formation and simplified the profile of every tested microsatellite compared to PCR that allowed an easier identification of the genuine alleles, notably for the mononucleotide repeat microsatellites for which the $n-1$ stutter peak were higher than the genuine n allele peak (Figure 2C).

The capacity of multiplexing of LT-RPA was also evaluated by combining the NR24, HT17, and BAT26 simplex assays in a triplex LT-RPA assay. These microsatellites were selected because they present different amplicon lengths allowing their separation by capillary electrophoresis. The multiplex LT-RPA conditions were the same as the simplex LT-RPA conditions except for the concentration of primers that were adjusted so that the final concentration of all primers (sum of the three pairs of primers) equaled the primer concentration in a simplex assay. Thus, both multiplex LT-RPA reactions allowed the amplification of the three microsatellite loci and their separation by capillary electrophoresis while maintaining the strong reduction of stutter peaks, indicating that LT-RPA is compatible with multiplexing (Figure 2C).

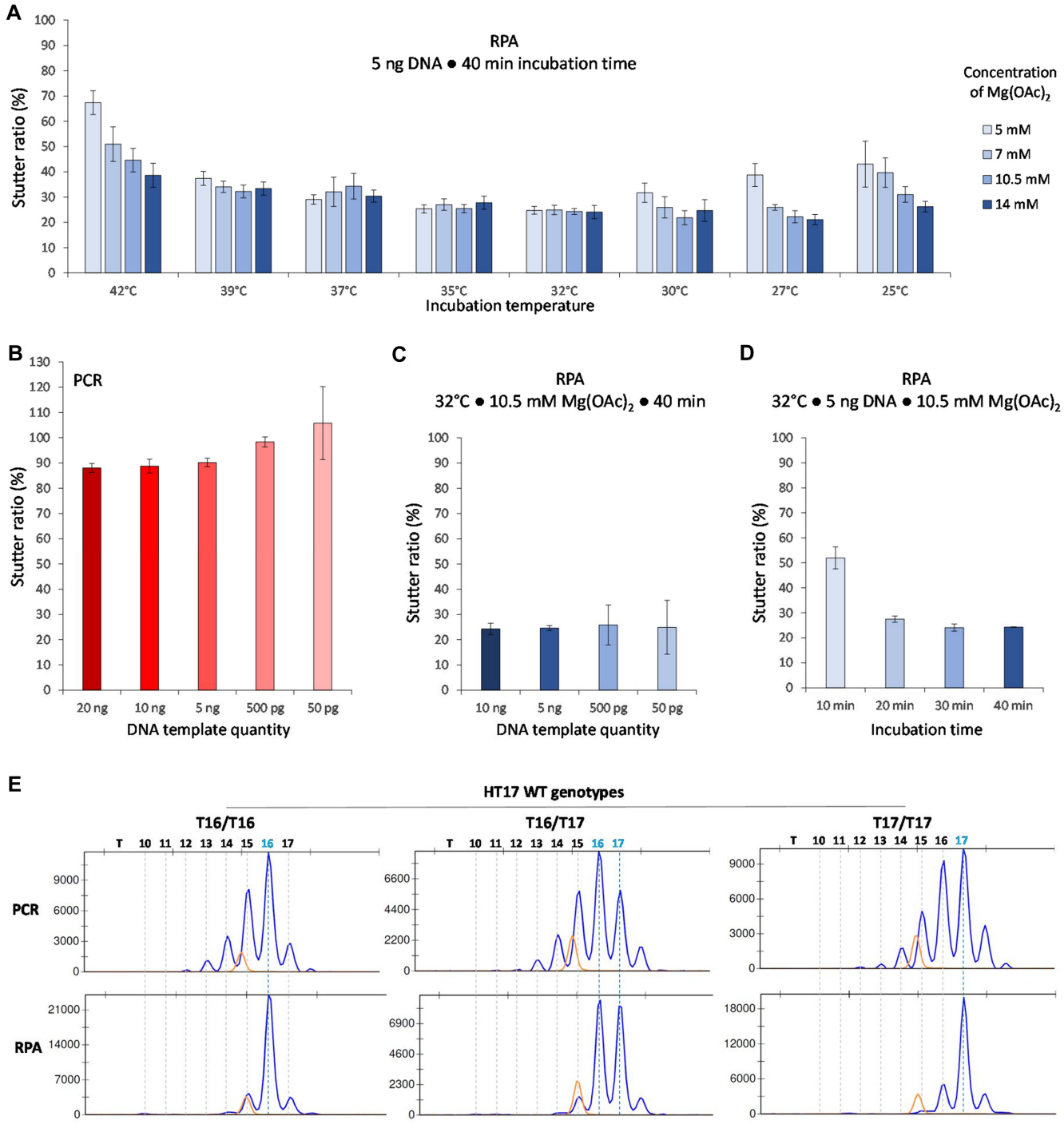

Figure 1. Effect of the variation of RPA conditions on stutter ratio (SR) compared to PCR using HT17 microsatellite. (**A**) RPA of HT17 using 5 ng of DNA of CEPH 0144711 T17/T17 cell line and 40 min incubation and variable concentrations of $Mg(OAc)_2$ (5 mM to 14 mM) and of isothermal amplification temperatures (25–42°C). (**B**) PCR amplification of HT17 using 50 pg to 20 ng of DNA of CEPH 0144711 T17/T17 cell line. (**C**) RPA of HT17 using 50 pg to 10 ng of DNA of CEPH 0144711 T17/T17 cell line, 10.5 mM of $Mg(OAc)_2$ at 32°C during 40 min. (**D**) RPA of HT17 using 5 ng of DNA of CEPH 0144711 T17/T17 cell line, 10.5 mM of $Mg(OAc)_2$ at 32°C during 10 to 40 min. (**E**) HT17 microsatellite profiles obtained after PCR and RPA (32°C incubation during 40 min with 10.5 mM of $Mg(OAc)_2$) using 20 ng and 5 ng DNA respectively of CEPH cell line harboring the three WT genotypes. Blue dotted lines indicate the HT17 alleles present in the WT cell lines used. Blue and orange peaks correspond to HT17 microsatellite and 140 nucleotide size standard respectively. All experiments were performed in triplicate (B–D), quadruplicate (A) or sextuplicate (B-C, for 50 pg).

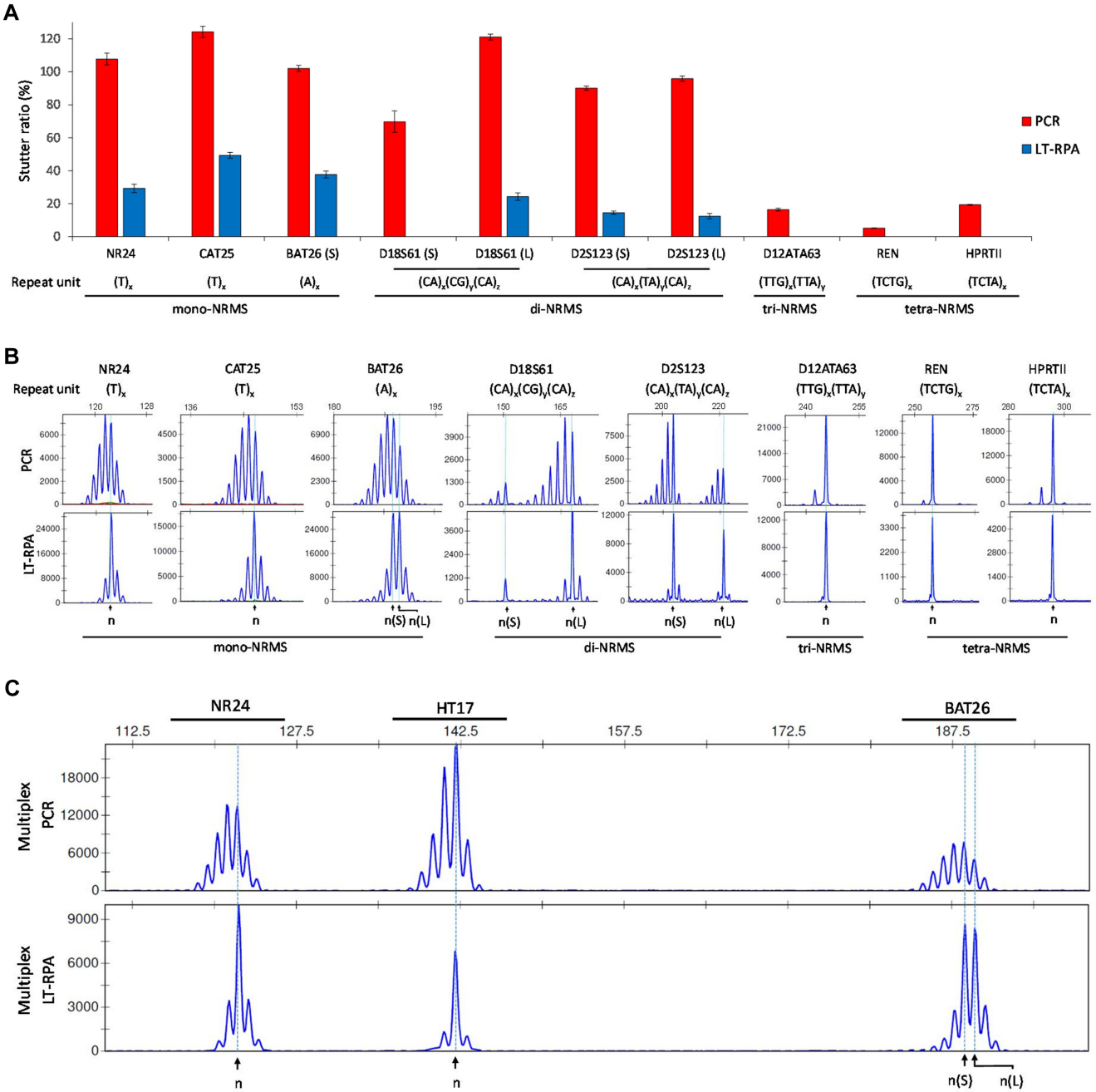

Figure 2. Effect of LT-RPA on stutter peak formation using mono- to tetra-nucleotide repeat microsatellites and CEPH 0144711 cell line. (**A**) Stutter ratio obtained with PCR (20 ng of DNA) and RPA (5 ng DNA, 10.5 mM $Mg(OAc)_2$ and 40 min incubation at 32°C). (**B**) Microsatellite profiles of the tested mono- to tetra-nucleotide repeat microsatellites using CEPH 0144711 cell line. (**C**) Multiplexing of PCR and low temperature RPA using NR24, HT17 and BAT26 assays and CEPH 0144711 cell line. *n*(S) and *n*(L) indicate the small and large allele of a same microsatellite respectively. NRMS, nucleotide repeat microsatellites. All experiments were performed in quadruplicates.

Evaluation of the mono- to tetra-nucleotide repeat microsatellite LT-RPA assays on blood samples of healthy individuals

We evaluated the different mono- to tetra-nucleotide repeat microsatellite assays on four blood samples from healthy donors using either PCR (20 ng of DNA in 20 μl) or LT-RPA (5 ng of DNA and 10.5 mM of $Mg(OAc)_2$ in 10 μl at 32°C). Compared to PCR, a simplification of the profiles of all tested microsatellites was visible with LT-RPA due to the decrease (mono- and di-nucleotide repeat microsatellites) or disappearance (tri- and tetra-nucleotide repeat microsatellites) of stutter peaks (Figure 3 and Supplementary Figure S6). In all cases, the alleles were clearly identifiable and the genotype calling was much easier with LT-RPA than with PCR (Figure 3A and Supplementary Figure S6). Thus, the identification of both alleles of BAT26 in B1 and of D18S61 in B3 were quite complicated after PCR but were very easy

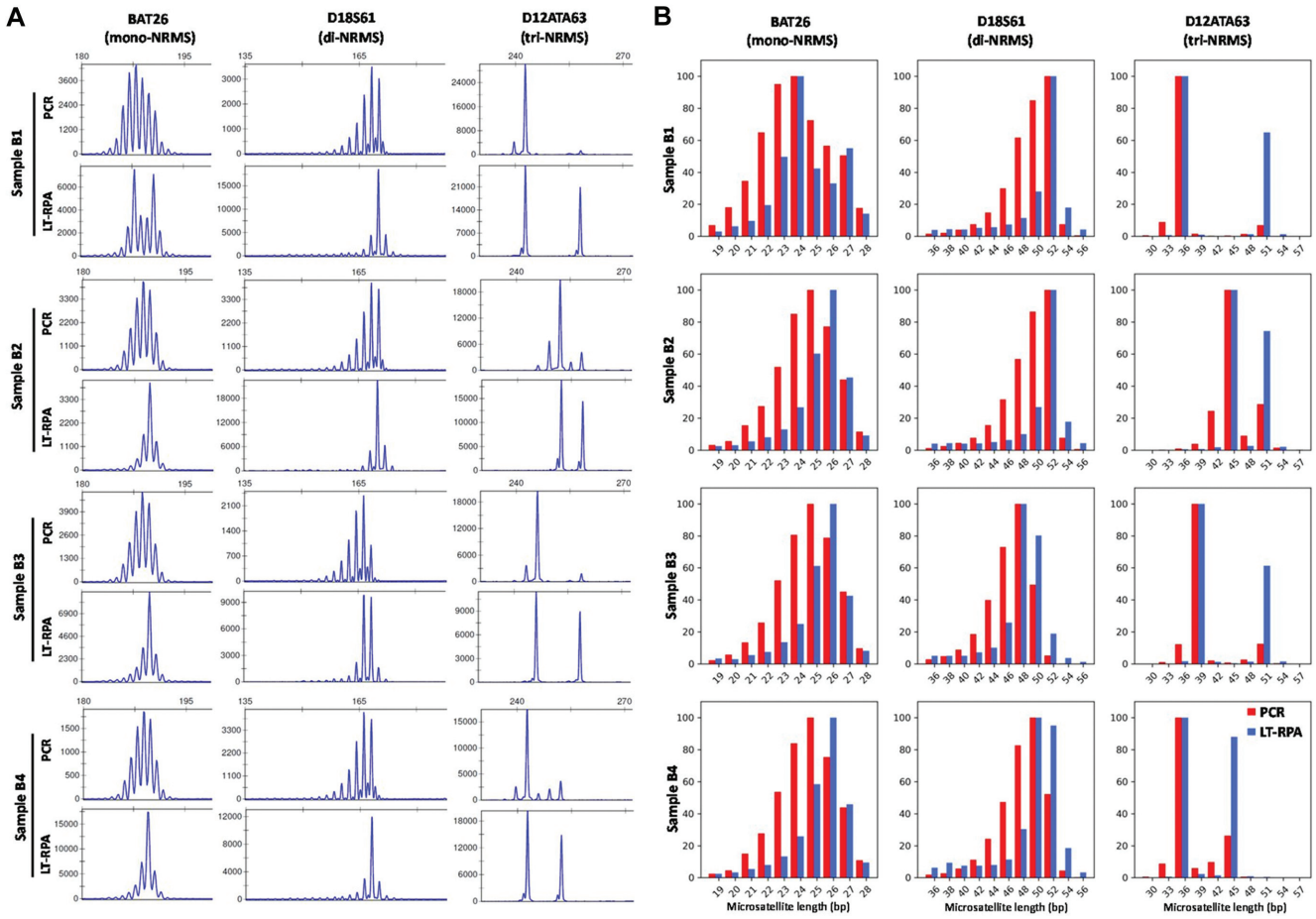

Figure 3. Microsatellite profiles of BAT26, D18S61 and D12ATA63 obtained with four blood samples (named B1, B2, B3 and B4) from healthy individuals using PCR (20 ng of DNA in 20 μl) and LT-RPA (5 ng of DNA and 10.5 mM of $Mg(OAc)_2$ in 10 μl at 32°C). (**A**) Profiles obtained by capillary electrophoresis fragment analysis. (**B**) Profiles obtained by amplicon sequencing. Capillary electrophoresis experiments were performed in duplicates and a representative profile is shown for each duplicated experiment. For amplicon sequencing experiments, the y-axis indicates the read counts for each alleles on a scale of 0–100. NRMS, nucleotide repeat microsatellite.

with LT-RPA (Figure 3A). Moreover, for the D12ATA63 microsatellite assay, a strong PCR amplification bias was observed between the small and the large alleles. The large allele tended to disappear as the gap between the two alleles increased and this could result in false genotype calling (Figure 3A). On the other hand, both alleles of D12ATA63 microsatellite were always clearly visible when LT-RPA was used which should allow to avoid potential genotyping errors (Figure 3A).

To further confirm these results and to identify precisely the repeated nucleotides in each allele, we performed amplicon sequencing of the nine microsatellites. PCR and LT-RPA amplicons were amplified, pooled by sample and amplification method and ligated to dual-indexed adapters before being sequenced on an iSeq 100 with an average depth of 27 000× (Supplementary Table S2). The microsatellite profiles obtained by amplicon sequencing were almost identical to fragment analysis profiles for all samples except for D18S61 in sample B4 that appeared homozygous by capillary electrophoresis and heterozygous by amplicon sequencing (Figure 3B and Supplementary Figure S7). This was explained by a 2 bp deletion corresponding to a rare polymorphism (rs139180351) located 37 nucleotides downstream the microsatellite in the larger allele resulting in two different amplicons of the same length in capillary electrophoresis experiments. Amplicon sequencing also allowed the accurate sequence identification of each microsatellite allele that was not feasible for compound microsatellites (D18S61, D2S123 and D12ATA63) by fragment analysis (Supplementary Table S3).

Amplification of tri-nucleotide repeat microsatellites implicated in triplet repeat diseases by LT-RPA

We assessed the effect of LT-RPA on the amplification of trinucleotide repeat microsatellites presenting repeat motifs (CGG, CTG, CAG and GAA) that are known to form secondary structures and whose allele expansion is responsible for several genetic diseases (48,49). We selected four microsatellites implicated in fragile X syndrome, Huntington disease like 2, Huntington disease and Friedrich ataxia that are located in *FMR1* (CGG repeat), *JPH3* (CTG repeat), *HTT* (CAG repeat) and *FXN* (GAA repeat) genes respectively. Concerning *FMR1*, we were not able to amplify the

microsatellite by LT-RPA, despite the addition of DMSO (3% to 9%), the use of five different primer pairs and the increase of reaction duration (40 min to 3 h) and temperature (32–42°C), indicating that CG-rich templates could prevent DNA amplification by RPA (data not shown). On the contrary, CTG and GAA triplet repeats were easily amplified in the standard LT-RPA conditions and presented a strong reduction of the stutter peaks compared to PCR in blood and CEPH cell line DNA samples (Supplementary Figure S8). Of note, the GAA triplet repeat located in *FXN* is contiguous to an A_{16} microsatellite explaining the compound microsatellite profile (Supplementary Figure S8). The amplification of *HTT* CAG trinucleotide microsatellite was also feasible with LT-RPA and also induced a reduction of stutter artifacts compared to PCR (Supplementary Figure S9). However, it required an increase of reaction duration (90 min) and temperature (37°C) in order to enhance the amplicon yield and reduce slight background noise present on the baseline at 32°C respectively. In the three cell lines derived from Huntington disease patients, LT-RPA allowed the detection of *HTT* CAG-expanded alleles with less amplification bias in favor of the short allele compared to PCR (Supplementary Figure S9).

HT17 LT-RPA on serial dilutions of cancer cell lines with MSI phenotype

In order to evaluate the potential benefit of low temperature RPA on the detection of MSI in cancer, we performed PCR and LT-RPA experiments using HT17 on serial dilutions of MSI cancer cell lines bearing different HT17 mutations (T10 to T15) mixed with wild-type T16/T16 cell lines to the following fractions: 50%, 25%, 10%, 5%, 1%. After PCR, the limit of detection (LOD) of MSI was dependent of the length of HT17 deletion so that large deletions presented lower LOD, due to the presence of stutter peaks (n –1 and lower) of the T16 WT alleles that masked the presence of mutant alleles (Figure 4 and Supplementary Figure S10A). Thus, a LOD of 5%, 10%, 25% and 50% was observed for Co-115, LoVo, HCT-15 and SNU-1 respectively (Figure 4). When LT-RPA was used on the cell line dilutions, an improvement of the LOD was observed from T12-T15 mutant alleles while larger deletions (T10–T11) presented the same LOD as for PCR (Figure 4 and Supplementary Figure S10B). Due to the reduction (n – 1) or disappearance (n – 2 and lower) of stutter peaks, the LOD of MSI decreased to 5% for LoVo and HCT-15 and to 25% for SNU-1, while it remained at 5% for Co-115 (Figure 4). Moreover, the simplification of the microsatellite profile with LT-RPA allowed an easy identification of the mutant alleles, notably for HCT-15 where the presence of stutter peaks complicated the mutant allele identification (Figure 4 and Supplementary Figure S10). Thus, replacing PCR by LT-RPA simplified and improved the detection of MSI.

HT17 LT-RPA testing in fresh frozen and FFPE tumor samples of dMMR CRC patients

A collection of 20 genomic DNAs extracted from fresh frozen and FFPE CRC samples previously assessed as MMR deficient by IHC was evaluated for the detection

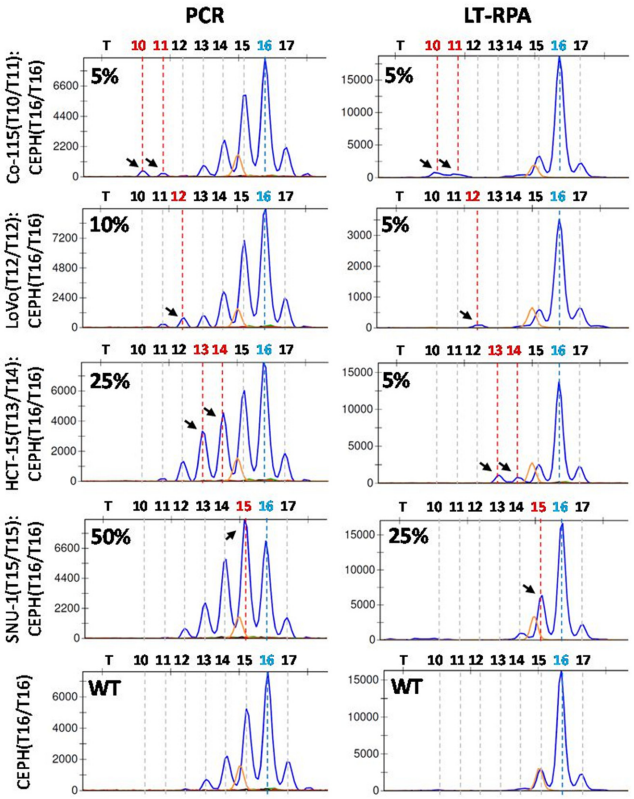

Figure 4. HT17 LT-RPA improves the limit of detection of MSI compared to PCR. PCR (20 ng of DNA) and LT-RPA (5 ng of DNA and 10.5 mM of $Mg(OAc)_2$ at 32°C during 40 min) experiments were conducted on dilution series of cancer cell lines (50%, 25%, 10%, 5% and 1%) harboring T10 to T15 HT17 mutant alleles mixed with T16/T16 WT CEPH cell lines. Only HT17 profiles of the fractions corresponding to the limit of detection of the mutant alleles are shown. Arrows indicate for each cancer cell line used the lowest amount of mutant alleles visible in mutant profiles compared to WT, defining the limit of detection of MSI for each type of mutation. Red and blue dotted lines indicate HT17 genotype of cancer and WT cell lines used respectively. Orange peaks correspond to 140 nucleotide size standard. All experiments were performed in triplicates.

of MSI using HT17 PCR and LT-RPA assays (Figure 5 and Supplementary Figure S11). All LT-RPA reactions performed well whether with fresh frozen or FFPE samples indicating that the use of degraded DNA extracted from FFPE samples did not affect RPA reactions. After standard PCR, MSI was detectable in most CRC samples (17 out of 20) except for three samples (S10, S11 and S12) for which the mutant alleles were difficult to distinguish from the stutter peaks as they were located in the stutter peak tail (Figure 5 and Supplementary Figure S11). Moreover, although the detection of MSI was easily performed with PCR, the precise identification of the mutant HT17 alleles was sometimes complicated due to the presence of stutter peaks (Figure 5 and Supplementary Figure S11). On the contrary, MSI was detected in all the CRC samples after LT-RPA, and the identification of the genuine mutant alleles was much easier compared to PCR, due to the absence of most stutter artefacts (Figure 5 and Supplementary Figure S11). Of note, while the microsatellite profiles obtained by PCR for samples S11 and S12 are nearly indistinguishable,

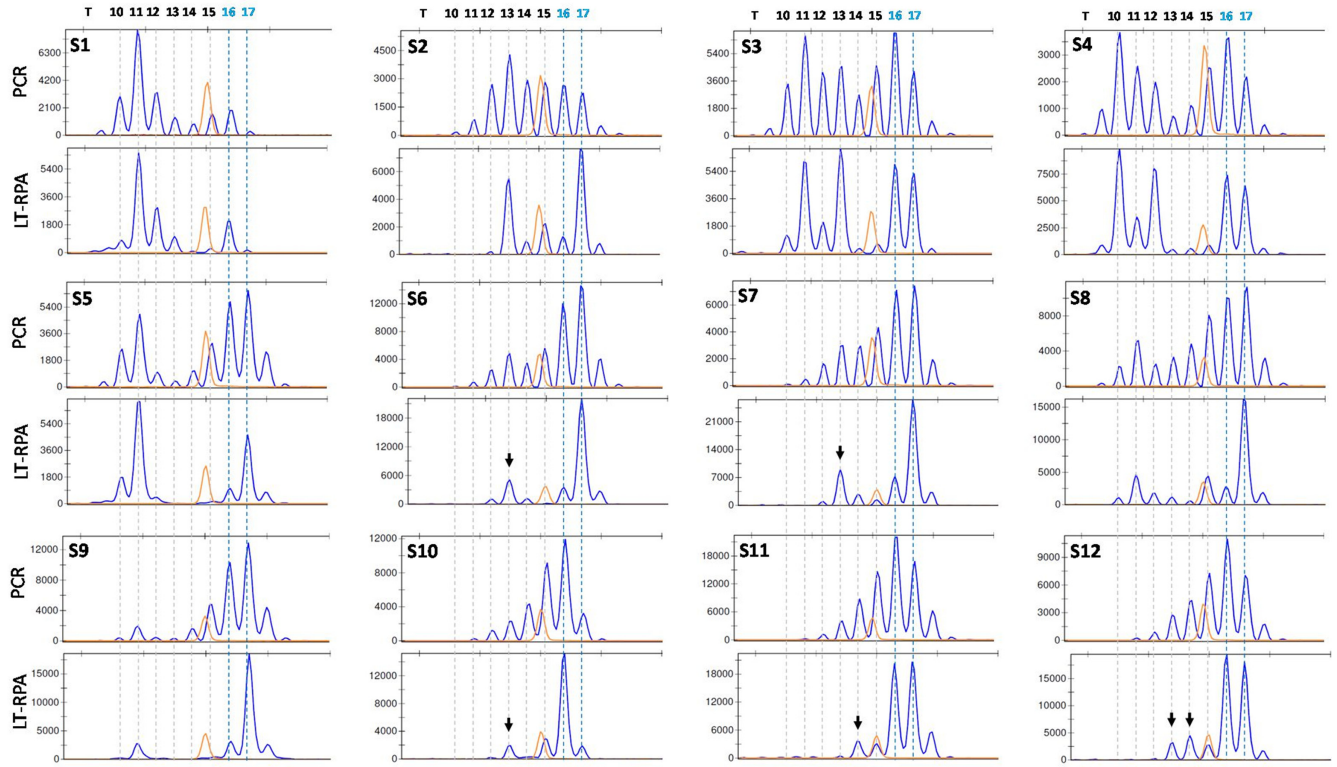

Figure 5. Detection of MSI in 12 fresh frozen MMR-deficient CRC samples (S1-S12) using HT17 microsatellite and LT-RPA and PCR. Blue dotted lines indicate the two possible WT alleles for HT17. Arrowheads indicate mutant alleles clearly visible and/or identifiable in the HT17 profile obtained with low temperature RPA but not with PCR. All experiments were performed in duplicates.

LT-RPA allowed a clear identification of T14 and T13/T14 mutant alleles in sample S11 and S12 respectively (Figure 5). Thus, these results demonstrated that LT-RPA improved MSI detection and identification in CRC samples compared to PCR.

MSI detection was further assessed in CRC samples S10, S11 and S12 and three control cell line samples by small-pool PCR (SP-PCR), a method based on multiple PCR experiments using 0–3 genome equivalents allowing to lower the LOD of MSI (14,50,51). 384 PCR reactions were performed per sample and resulted in 16–42% positive PCR, most of which (76–91%) being amplified from a single molecule (Supplementary Table S4). The expected genuine WT and/or mutant alleles were found in the three control samples but also a smaller proportion of −1 and/or +1/+2 alleles (Supplementary Table S4), which was probably due to a stochastic effect of very low-template PCR known as allele drop-in (32,52). In the three CRC samples MSI was easily detected by SP-PCR, thereby confirming LT-RPA results (Figure 5 and Supplementary Table S4). While SP-PCR identified additional mutated alleles compared to LT-RPA (T12 in S10 and T13 in S11), they could also result from allele drop-in (Figure 5 and Supplementary Table S4).

DISCUSSION

Since their discovery, the analysis of microsatellites has always been challenging in spite of the constant evolution of the different genotyping technologies, from capillary electrophoresis fragment analysis (FA) to next-generation sequencing (NGS). One of the main source of difficulty and inaccuracy of microsatellite allele calling came from the presence of stutter peaks introduced during the PCR and attributed to slipped-strand mispairing that greatly complicated the microsatellite profiles obtained by FA or NGS (14,32–34,36,37). In 1996, Hite *et al.*, showed that a decrease of the elongation temperature to 37°C during the PCR reduced the stutter artifacts but required the addition of a new thermolabile DNA polymerase at each cycle, thereby rendering this method incompatible to most applications (34). Since then, no improvement in the amplification of microsatellites has been proposed to reduce the formation of stutter peaks, while the efforts to deal with these artifacts were mainly focused on the development of complex post-PCR analysis approaches (14,33,36–40).

In the present study, we presented an alternative to PCR for the amplification of microsatellites, that is easier-to-handle and with a faster time-to-result. It relied on the low temperature isothermal amplification of microsatellites using RPA that drastically reduced and sometimes completely abolished the formation of stutter artifacts depending on the type (mono- to tetra-nucleotide repeat microsatellites) or the number of repetitions (Figures 1 and 2). We chose RPA among the isothermal methods, as it allows an amplification of DNA at temperatures lower than 42°C while most of the others required a higher reaction temperature that would have no or little effects on SR (41). For example, loop-mediated isothermal amplification (LAMP) has al-

ready been evaluated on microsatellites with a reaction temperature of 55°C without reducing the formation of stutter peaks compared to PCR (53). Similarly, the reduction of the annealing and elongation of PCR at 56°C had only produced a very slight reduction of the stutter peaks compared to the standard PCR (54).

Using LT-RPA, we obtained a strong reduction of the stutter peak formation and of SR that tended to completely disappear for tri- and tetra-nucleotide repeat microsatellites, thereby greatly simplifying the profile of microsatellites and the identification of the genuine alleles compared to PCR (Figures 1 and 2). This was particularly true when the $n-1$ stutter peak became higher than the n genuine allele peak with PCR but never with low temperature RPA, as showed for PCR with low quantity of DNA template (Figure 1), a well-known phenomenon already described in the literature (32) or for PCR on long mononucleotide repeat microsatellites such as NR24, CAT25 and BAT26 (Figure 2). In both cases, the n allele peak remained the highest with LT-RPA. However, it should be noted that our PCR conditions may not be optimized to produce the least stutter artifacts and that SR differences between PCR and LT-RPA may be lower under other PCR conditions. Moreover, since tri- and tetra-nucleotide repeat microsatellites present only few stutter peaks when amplified by PCR, the major advantage of LT-RPA for the reduction of stutter artifact formation concerns mono- and di-nucleotide microsatellites. We also evaluated LT-RPA on blood samples from healthy individuals that allowed easier and unbiased calling of alleles and genotypes for every microsatellite compared to PCR (Figure 3), which could potentially greatly improve the automatic genotyping from capillary electrophoresis fragment analysis data. Moreover, we showed that multiplexing the microsatellite assays was easily feasible with LT-RPA, which could be particularly useful to interrogate multiple loci when limited amount of DNA is available. Contrary to multiplex PCR, primers with similar annealing temperature were not required with multiplex LT-RPA. We also showed that LT-RPA allowed the amplification of CAG, CTG and GAA trinucleotide repeat microsatellites (Supplementary Figures S8 and S9) but not CGG triplet repeat, the latter being a CG-rich sequence that is particularly difficult to amplify by PCR (55). The expansion of these triplet repeats is implicated in several genetic disorders (48,49) and we showed that the detection of *HTT* CAG expanded repeats was possible by LT-RPA (Supplementary Figure S9), suggesting that this isothermal amplification approach could be used for the diagnosis of trinucleotide repeat disorders.

Our general recommendations for the design of a RPA microsatellite assay are first to design the primers the same way as for a PCR assay. Thus, the PCR assays used were fully compatible with RPA without any sequence modification. The use of longer primers (30–35 nucleotides) can also be performed as recommended in the original RPA study and by the manufacturer (42,44) but is not mandatory. To optimize the reaction conditions, we recommend a duration of RPA reaction of at least 30 min and an evaluation of different temperatures between 30°C and 37°C to obtain the lowest SR. We did not notice any difference in terms of performance between the TwistAmp™ Liquid basic and the TwistAmp™ basic Kits, but the TwistAmp™ basic Kit was easier to handle under optimized LT-RPA conditions. The only critical step in RPA was the addition of the $Mg(OAc)_2$ that needed to be performed just before the incubation on a preheated thermocycler because its addition starts the reaction.

In the context of cancer diagnostics, we showed that replacing PCR by LT-RPA for the amplification of microsatellites could easily improve and simplify the detection of MSI. Using HT17, a quasi-monomorphic marker with improved sensitivity (98.4% against 95.1%) for MSI detection in CRC compared to the gold standard pentaplex panel (45), the LOD of MSI was improved, notably for small deletions that were masked by the stutter peaks (Figure 4). Due to the reduction of stutter peaks, the identification of the WT and mutant alleles were easier and could also benefit to CRC patients, as the HT17 mutant alleles bear predictive information of the response to chemotherapy (5-FU and oxaliplatin) depending on the length of HT17 deletions (46,47). LT-RPA was successfully evaluated on three different types of clinical samples that included blood from healthy donors (Figure 3), fresh frozen and FFPE tumor samples of CRC patients (Figure 5 and Supplementary Figure S11), suggesting that the method could be easily used and implemented in routine clinical testing. Thus, the use of LT-RPA for MSI detection might be valuable in oncology, especially since the regain of interest for MSI after the recent discovery that MSI is a predicator of the efficacy of immune checkpoint blockade therapy in solid tumors (8,10,11). Further validation studies on larger cohorts of patients will need to be performed in order to evaluate the benefits of LT-RPA over PCR for the detection of MSI in cancer. Compared to small-pool PCR, LT-RPA is easier-to-handle with a shorter time-to-result but presents a limit of detection of MSI that cannot be decreased $<5\%$ of mutant alleles in the standard conditions (Figures 4 and 5 and Supplementary Table S4). However, we also expect the rapid adaptation to LT-RPA of several PCR-based approaches improving the limit of detection of MSI in cancer such as small-pool PCR (51), HRM (56), E-*ice*-COLD-PCR (33), NaME-PrO (57) or NaMSIE (58), that could be required in some clinical applications (e.g. to detect MSI in blood, plasma, precancerous lesions, tumors with high level of stromal cell contamination or heterogeneous tumors where only a small subset of the tumor cells presents MSI).

More recently, NGS has allowed to increase the number of analyzed microsatellites compared to capillary electrophoresis from a few dozens to almost all microsatellites in several different applications including MSI detection in cancer (8,14–17). However in spite of this improvement, the analysis of microsatellites from NGS data is challenging due to multiple errors inherent in the method of library preparation and sequencing-by-synthesis of short-reads (24,59). Moreover the sequencing libraries for whole genome, whole exome and targeted gene sequencing are mostly based on PCR amplification, thereby introducing

stutter artifacts leading to inaccurate microsatellite allele calling and requiring the development of different computational approaches and bioinformatics correction algorithms (18–25). To overcome these difficulties, the use of a PCR-free protocol for the preparation of NGS libraries has shown a reduction of the formation of stutter artifacts up to nine fold compared to PCR-containing protocols and has improved the accuracy of microsatellite genotyping from NGS data (24). However, the use of this protocol is currently limited to whole genome sequencing. The use of low temperature RPA could also be an alternative to PCR for library preparation of NGS experiments that could greatly reduce the errors and complexity of microsatellite NGS data and simplify downstream analyses. This was exemplified in our study by the results of amplicon sequencing showing simplified microsatellite profiles and an improvement of allele calling with LT-RPA compared to PCR (Figure 3B and Supplementary Table S3).

CONCLUSION

In summary, we have developed an easy, rapid and cost-effective low isothermal amplification approach based on RPA of microsatellites that drastically reduces the formation of stutter artifacts and greatly simplifies the microsatellite profiles obtained by capillary electrophoresis, thereby improving the calling of the microsatellite alleles. Applied to the detection of MSI in CRC, it improves the limit of detection of small deletions and allows an easy identification of the mutant alleles that enhances the assessment of MSI in clinical samples. Due to its simplicity, low temperature RPA of microsatellites has a strong potential for routine clinical diagnosis and also for numerous other basic or translational applications where microsatellite genotyping is required. Thus, low temperature RPA may be considered as a better alternative to PCR for the amplification of microsatellites.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We want to acknowledge the CEPH Biobank for providing DNA of CEPH lymphoblastoid cell lines and Men Woon and Candide How Kit for their careful editing of the manuscript. LGB received support from the GENMED Laboratory of Excellence on Medical Genomics [ANR-10-LABX-0013 to L.B.].

Author contributions: All authors contributed significantly to this work. A.H.-K. conceived and supervised the study. A.D. and O.B. provided clinical samples. A.D., L.G.B., O.B. and A.H.-K. performed the experiments. A.D., V.R. and A.H.-K. analyzed the data. A.H.-K. drafted the manuscript. A.D., A.D., L.G.B., O.B., V.R., J.-F.D. and A.H.-K. read and approved the final version of the submitted manuscript.

FUNDING

Foundation Jean Dausset – Centre d'Etude du Polymorphisme Humain. Funding for open access charge: Foundation Jean Dausset – Centre d'Etude du Polymorphisme Humain.

Conflict of interest statement. None declared.

REFERENCES

1. Ellegren,H. (2004) Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.*, **5**, 435–445.
2. Gulcher,J. (2012) Microsatellite markers for linkage and association studies. *Cold Spring Harbor Protoc.*, **2012**, 425–432.
3. Putman,A.I. and Carbone,I. (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.*, **4**, 4399–4428.
4. Miah,G., Rafii,M.Y., Ismail,M.R., Puteh,A.B., Rahim,H.A., Islam Kh,N. and Latif,M.A. (2013) A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. *Int. J. Mol. Sci.*, **14**, 22499–22528.
5. Stadele,V. and Vigilant,L. (2016) Strategies for determining kinship in wild populations using genetic data. *Ecol. Evol.*, **6**, 6107–6120.
6. Gettings,K.B., Aponte,R.A., Vallone,P.M. and Butler,J.M. (2015) STR allele sequence variation: Current knowledge and future issues. *Forensic Sci. Int. Genet.*, **18**, 118–130.
7. Boland,C.R. and Goel,A. (2010) Microsatellite instability in colorectal cancer. *Gastroenterology*, **138**, 2073–2087.
8. Hause,R.J., Pritchard,C.C., Shendure,J. and Salipante,S.J. (2016) Classification and characterization of microsatellite instability across 18 cancer types. *Nat. Med.*, **22**, 1342–1350.
9. Cortes-Ciriano,I., Lee,S., Park,W.Y., Kim,T.M. and Park,P.J. (2017) A molecular portrait of microsatellite instability across multiple cancers. *Nat. Commun.*, **8**, 15180.
10. Le,D.T., Durham,J.N., Smith,K.N., Wang,H., Bartlett,B.R., Aulakh,L.K., Lu,S., Kemberling,H., Wilt,C., Luber,B.S. *et al.* (2017) Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science*, **357**, 409–413.
11. Le,D.T., Uram,J.N., Wang,H., Bartlett,B.R., Kemberling,H., Eyring,A.D., Skora,A.D., Luber,B.S., Azad,N.S., Laheru,D. *et al.* (2015) PD-1 Blockade in tumors with Mismatch-Repair deficiency. *N. Engl. J. Med.*, **372**, 2509–2520.
12. Stoffel,E.M., Mangu,P.B., Gruber,S.B., Hamilton,S.R., Kalady,M.F., Lau,M.W., Lu,K.H., Roach,N. and Limburg,P.J. (2015) Hereditary colorectal cancer syndromes: American society of clinical oncology clinical practice guideline endorsement of the familial risk-colorectal cancer: European society for medical oncology clinical practice guidelines. *J. Clin. Oncol.*, **33**, 209–217.
13. Romsos,E.L. and Vallone,P.M. (2015) Rapid PCR of STR markers: Applications to human identification. *Forensic Sci. Int. Genet.*, **18**, 90–99.
14. Baudrin,L.G., Deleuze,J.F. and How-Kit,A. (2018) Molecular and computational methods for the detection of microsatellite instability in cancer. *Front. Oncol.*, **8**, 621.
15. Zhang,S., Niu,Y., Bian,Y., Dong,R., Liu,X., Bao,Y., Jin,C., Zheng,H. and Li,C. (2018) Sequence investigation of 34 forensic autosomal STRs with massively parallel sequencing. *Sci. Rep.*, **8**, 6810.
16. Gan,C., Love,C., Beshay,V., Macrae,F., Fox,S., Waring,P. and Taylor,G. (2015) Applicability of next generation sequencing technology in microsatellite instability testing. *Genes*, **6**, 46–59.
17. Waalkes,A., Smith,N., Penewit,K., Hempelmann,J., Konnick,E.Q., Hause,R.J., Pritchard,C.C. and Salipante,S.J. (2018) Accurate Pan-Cancer molecular diagnosis of microsatellite instability by Single-Molecule molecular inversion probe capture and High-Throughput sequencing. *Clin. Chem.*, **64**, 950–958.
18. Maruvka,Y.E., Mouw,K.W., Karlic,R., Parasuraman,P., Kamburov,A., Polak,P., Haradhvala,N.J., Hess,J.M., Rheinbay,E., Brody,Y. *et al.* (2017) Analysis of somatic microsatellite indels identifies driver events in human tumors. *Nat. Biotechnol.*, **35**, 951–959.
19. Dashnow,H., Lek,M., Phipson,B., Halman,A., Sadedin,S., Lonsdale,A., Davis,M., Lamont,P., Clayton,J.S., Laing,N.G. *et al.* (2018) STRetch: detecting and discovering pathogenic short tandem repeat expansions. *Genome Biol.*, **19**, 121.
20. Leclercq,S., Rivals,E. and Jarne,P. (2007) Detecting microsatellites within genomes: significant variation among algorithms. *BMC Bioinformatics*, **8**, 125.

21. Cao,M.D., Balasubramanian,S. and Boden,M. (2015) Sequencing technologies and tools for short tandem repeat variation detection. *Brief. Bioinform.*, **16**, 193–204.
22. Ganschow,S., Silvery,J., Kalinowski,J. and Tiemann,C. (2018) toaSTR: A web application for forensic STR genotyping by massively parallel sequencing. *Forensic Sci. Int. Genet.*, **37**, 21–28.
23. Tae,H., Kim,D.Y., McCormick,J., Settlage,R.E. and Garner,H.R. (2014) Discretized Gaussian mixture for genotyping of microsatellite loci containing homopolymer runs. *Bioinformatics*, **30**, 652–659.
24. Fungtammasan,A., Ananda,G., Hile,S.E., Su,M.S., Sun,C., Harris,R., Medvedev,P., Eckert,K. and Makova,K.D. (2015) Accurate typing of short tandem repeats from genome-wide sequencing data and its applications. *Genome Res.*, **25**, 736–749.
25. Fang,H., Wu,Y., Narzisi,G., O'Rawe,J.A., Barron,L.T., Rosenbaum,J., Ronemus,M., Iossifov,I., Schatz,M.C. and Lyon,G.J. (2014) Reducing INDEL calling errors in whole genome and exome sequencing data. *Genome medicine*, **6**, 89.
26. Schlotterer,C. and Tautz,D. (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Res.*, **20**, 211–215.
27. Hauge,X.Y. and Litt,M. (1993) A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum. Mol. Genet.*, **2**, 411–415.
28. Murray,V., Monchawin,C. and England,P.R. (1993) The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. *Nucleic Acids Res.*, **21**, 2395–2398.
29. Walsh,P.S., Fildes,N.J. and Reynolds,R. (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res.*, **24**, 2807–2812.
30. Shinde,D., Lai,Y., Sun,F. and Arnheim,N. (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Res.*, **31**, 974–980.
31. Brookes,C., Bright,J.A., Harbison,S. and Buckleton,J. (2012) Characterising stutter in forensic STR multiplexes. *Forensic Sci. Int. Genet.*, **6**, 58–63.
32. Gill,P., Haned,H., Bleka,O., Hansson,O., Dorum,G. and Egeland,T. (2015) Genotyping and interpretation of STR-DNA: Low-template, mixtures and database matches-Twenty years of research and development. *Forensic Sci. Int. Genet.*, **18**, 100–117.
33. How-Kit,A., Daunay,A., Buhard,O., Meiller,C., Sahbatou,M., Collura,A., Duval,A. and Deleuze,J.F. (2018) Major improvement in the detection of microsatellite instability in colorectal cancer using HSP110 T17 E-ice-COLD-PCR. *Hum. Mutat.*, **39**, 441–453.
34. Hite,J.M., Eckert,K.A. and Cheng,K.C. (1996) Factors affecting fidelity of DNA synthesis during PCR amplification of d(C-A)n.d(G-T)n microsatellite repeats. *Nucleic Acids Res.*, **24**, 2429–2434.
35. Ingham,D., Diggle,C.P., Berry,I., Bristow,C.A., Hayward,B.E., Rahman,N., Markham,A.F., Sheridan,E.G., Bonthron,D.T. and Carr,I.M. (2013) Simple detection of germline microsatellite instability for diagnosis of constitutional mismatch repair cancer syndrome. *Hum. Mutat.*, **34**, 847–852.
36. Salipante,S.J., Scroggins,S.M., Hampel,H.L., Turner,E.H. and Pritchard,C.C. (2014) Microsatellite instability detection by next generation sequencing. *Clin. Chem.*, **60**, 1192–1199.
37. Raz,O., Biezuner,T., Spiro,A., Amir,S., Milo,L., Titelman,A., Onn,A., Chapal-Ilani,N., Tao,L., Marx,T. *et al.* (2019) Short tandem repeat stutter model inferred from direct measurement of in vitro stutter noise. *Nucleic Acids Res.*, **47**, 2436–2445.
38. Bright,J.A., Taylor,D., Curran,J.M. and Buckleton,J.S. (2013) Developing allelic and stutter peak height models for a continuous method of DNA interpretation. *Forensic Sci. Int. Genet.*, **7**, 296–304.
39. Kedzierska,K.Z., Gerber,L., Cagnazzi,D., Krutzen,M., Ratan,A. and Kistler,L. (2018) SONiCS: PCR stutter noise correction in genome-scale microsatellites. *Bioinformatics*, **34**, 4115–4117.
40. Coble,M.D. and Bright,J.A. (2019) Probabilistic genotyping software: an overview. *Forensic Sci. Int. Genet.*, **38**, 219–224.
41. Zhao,Y., Chen,F., Li,Q., Wang,L. and Fan,C. (2015) Isothermal amplification of nucleic acids. *Chem. Rev.*, **115**, 12491–12545.
42. Daher,R.K., Stewart,G., Boissinot,M. and Bergeron,M.G. (2016) Recombinase polymerase amplification for diagnostic applications. *Clin. Chem.*, **62**, 947–958.
43. James,A. and Macdonald,J. (2015) Recombinase polymerase amplification: Emergence as a critical molecular technology for rapid, low-resource diagnostics. *Expert Rev. Mol. Diagn.*, **15**, 1475–1489.
44. Piepenburg,O., Williams,C.H., Stemple,D.L. and Armes,N.A. (2006) DNA detection using recombination proteins. *PLoS Biol.*, **4**, e204.
45. Buhard,O., Lagrange,A., Guilloux,A., Colas,C., Chouchene,M., Wanherdrick,K., Coulet,F., Guillerm,E., Dorard,C., Marisa,L. *et al.* (2016) HSP110 T17 simplifies and improves the microsatellite instability testing in patients with colorectal cancer. *J. Med. Genet.*, **53**, 377–384.
46. Collura,A., Lagrange,A., Svrcek,M., Marisa,L., Buhard,O., Guilloux,A., Wanherdrick,K., Dorard,C., Taieb,A., Saget,A. *et al.* (2014) Patients with colorectal tumors with microsatellite instability and large deletions in HSP110 T17 have improved response to 5-fluorouracil-based chemotherapy. *Gastroenterology*, **146**, 401–411.
47. Dorard,C., de Thonel,A., Collura,A., Marisa,L., Svrcek,M., Lagrange,A., Jego,G., Wanherdrick,K., Joly,A.L., Buhard,O. *et al.* (2011) Expression of a mutant HSP110 sensitizes colorectal cancer cells to chemotherapy and improves disease prognosis. *Nat. Med.*, **17**, 1283–1289.
48. Usdin,K., House,N.C. and Freudenreich,C.H. (2015) Repeat instability during DNA repair: Insights from model systems. *Crit. Rev. Biochem. Mol. Biol.*, **50**, 142–167.
49. Mirkin,S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932–940.
50. Coolbaugh-Murphy,M., Maleki,A., Ramagli,L., Frazier,M., Lichtiger,B., Monckton,D.G., Siciliano,M.J. and Brown,B.W. (2004) Estimating mutant microsatellite allele frequencies in somatic cells by small-pool PCR. *Genomics*, **84**, 419–430.
51. Coolbaugh-Murphy,M.I., Xu,J.P., Ramagli,L.S., Ramagli,B.C., Brown,B.W., Lynch,P.M., Hamilton,S.R., Frazier,M.L. and Siciliano,M.J. (2010) Microsatellite instability in the peripheral blood leukocytes of HNPCC patients. *Hum. Mutat.*, **31**, 317–324.
52. Taberlet,P., Griffin,S., Goossens,B., Questiau,S., Manceau,V., Escaravage,N., Waits,L.P. and Bouvet,J. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res.*, **24**, 3189–3194.
53. Lee,D., Ialicicco,M., Akkinepalli,H., Morreale,G., Liu,Y., Leung,H., Scippa,G.S., Greenland,A. and Mackay,I. (2012) Genotyping SSR length variants by isothermal DNA amplification. *Genome*, **55**, 691–695.
54. Seo,S.B., Ge,J., King,J.L. and Budowle,B. (2014) Reduction of stutter ratios in short tandem repeat loci typing of low copy number DNA samples. *Forensic Sci Int Genet.*, **8**, 213–218.
55. Rajan-Babu,I.S., Lian,M., Cheah,F.S.H., Chen,M., Tan,A.S.C., Prasath,E.B., Loh,S.F. and Chong,S.S. (2017) FMR1 CGG repeat expansion mutation detection and linked haplotype analysis for reliable and accurate preimplantation genetic diagnosis of fragile X syndrome. *Expert Rev. Mol. Med.*, **19**, e10.
56. Janavicius,R., Matiukaite,D., Jakubauskas,A. and Griskevicius,L. (2010) Microsatellite instability detection by high-resolution melting analysis. *Clin. Chem.*, **56**, 1750–1757.
57. Ladas,I., Yu,F., Leong,K.W., Fitarelli-Kiehl,M., Song,C., Ashtaputre,R., Kulke,M., Mamon,H. and Makrigiorgos,G.M. (2018) Enhanced detection of microsatellite instability using pre-PCR elimination of wild-type DNA homo-polymers in tissue and liquid biopsies. *Nucleic Acids Res.*, **46**, e74.
58. Baudrin,L.G., Duval,A., Daunay,A., Buhard,O., Bui,H., Deleuze,J.F. and How-Kit,A. (2018) Improved microsatellite instability detection and identification by nuclease-assisted microsatellite instability enrichment using HSP110 T17. *Clin. Chem.*, **64**, 1252–1253.
59. Treangen,T.J. and Salzberg,S.L. (2011) Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.*, **13**, 36–46.