Published online 3 August 2015

Nucleic Acids Research, 2015, Vol. 43, No. 22 ***e151***
doi: 10.1093/nar/gkv772

Oligonucleotide gap-fill ligation for mutation detection and sequencing *in situ*

Marco Mignardi[1,†], Anja Mezger[1,†], Xiaoyan Qian[1], Linnea La Fleur[2], Johan Botling[2], Chatarina Larsson[2] and Mats Nilsson[1,*]

[1]Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Stockholm, SE-17121 Sweden and [2]Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala, SE-75185, Sweden

Received November 13, 2014; Revised July 07, 2015; Accepted July 19, 2015

ABSTRACT

In clinical diagnostics a great need exists for targeted *in situ* multiplex nucleic acid analysis as the mutational status can offer guidance for effective treatment. One well-established method uses padlock probes for mutation detection and multiplex expression analysis directly in cells and tissues. Here, we use oligonucleotide gap-fill ligation to further increase specificity and to capture molecular substrates for *in situ* sequencing. Short oligonucleotides are joined at both ends of a padlock gap probe by two ligation events and are then locally amplified by target-primed rolling circle amplification (RCA) preserving spatial information. We demonstrate the specific detection of the A3243G mutation of mitochondrial DNA and we successfully characterize a single nucleotide variant in the *ACTB* mRNA in cells by *in situ* sequencing of RCA products generated by padlock gap-fill ligation. To demonstrate the clinical applicability of our assay, we show specific detection of a point mutation in the *EGFR* gene in fresh frozen and formalin-fixed, paraffin-embedded (FFPE) lung cancer samples and confirm the detected mutation by *in situ* sequencing. This approach presents several advantages over conventional padlock probes allowing simpler assay design for multiplexed mutation detection to screen for the presence of mutations in clinically relevant mutational hotspots directly *in situ*.

INTRODUCTION

Many *in vitro* molecular assays exploit enzymatic ligation of oligonucleotides as a means to analyze nucleic acids and proteins and applications range from mutation analysis, to cloning and sequencing of genes (1–4). DNA ligases are enzymes that mediate the junction of a DNA strand that is phosphorylated at the 5′-end to a DNA strand 3′-end upon hybridization to a complementary target DNA (5). The high specificity of this molecular process provides a powerful mechanism to analyze single nucleotide variants (SNV) in target DNA strands.

Based on DNA ligation as a means to provide specificity, padlock probes enable sensitive nucleic acid detection and genotyping of DNA molecules in solution and *in situ* (4,6). Padlock probes are linear oligonucleotides whose ends are joined together by template-dependent DNA ligation. The circular molecule resulting from this reaction is amplified via RCA and the RCA product (RCP) generated consists of a tandem-repeated sequence that forms submicron sized random coils. Upon labeling by fluorescent *in situ* hybridization the RCPs appear as intense dot-like signals suitable for digital quantification. The strong fluorescent signal allows molecular analyses to be performed directly in fixed cells or tissue sections, preserving sample morphology information. *In situ* assays based on the ligation and amplification of padlock probes have been used for, e.g. pathogen detection (7), detection of specific sequences in genomic DNA (8) and genotyping of mitochondrial DNA (mtDNA) (6) and mRNA molecules (9).

Padlock gap probes are an alternative version of padlock probes where the probe arms hybridize to the target with a gap of a defined number of nucleotides between the 5′-end and 3′-end of the probe. Assays based on padlock gap probes come in two variants. In the first version, enzymatic polymerization and ligation is used to fill the sequence and join the ends between the two target-complementary arms to circularize the padlock gap probe. Gaps ranging from one nucleotide to the entire length of an exon have been used for highly multiplexed genotyping (10) and sample enrichment (11) in solution-phase methods based on this approach. In the second version, a short oligonucleotide of the same length as the gap is hybridized between the probe arms and ligation is used to circularize the padlock gap probe to

*To whom correspondence should be addressed. Tel: +46 8 524 811 58; Email: mats.nilsson@scilifelab.se
†These authors contributed equally to the paper as first authors.

distinguish between the variants. This approach was used to discriminate alleles of genes in genomic DNA from salt-extracted halo preparations (12).

Recently, we used polymerization and ligation of padlock gap probes to generate a substrate which can be sequenced directly *in situ* (13). Hybridization and ligation of a cocktail of short probes to the padlock gap probe represents an alternative approach to generate substrates for *in situ* sequencing, as presented herein. The strict requirement for two highly specific ligation events to occur using this strategy theoretically offers an increased level of specificity compared to the single ligation event required in classical ligation based assays.

In this paper we demonstrate that padlock gap probes and oligonucleotide gap-fill ligation can be used for highly specific *in situ* mutational analysis. As a proof of concept, we use this approach to detect the mitochondrial A3243G mutation, causative of mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) syndrome (14), in human fibroblast cell lines. We also demonstrate the use of gap-fill ligation in combination with *in situ* sequencing for mRNA analysis by successfully targeting an expressed SNV in the beta actin (*ACTB*) gene. Furthermore, we show clinical applicability of this method by specific detection of a point mutation in the epidermal growth factor receptor (*EGFR*) mRNA in fresh frozen and FFPE lung cancer tissues.

MATERIALS AND METHODS

Cell and sample pretreatment

Human osteosarcoma 143B and human fibroblast M50 cell lines were cultured in Dulbecco's Modified Eagles medium (Gibco) without phenol red, with 10% fetal calf serum (Sigma), 50 μg/ml uridine (Sigma), 2 mM L-glutamine (Sigma) and 1 × Penicillin-Streptomycin (PEST, Sigma). Cells were cultured on Superfrost plus slides (Thermo Scientific) and fixed in 70% ethanol at room temperature after a wash in PBS. Permeabilization was made by using 0.01% pepsin from porcine gastric mucosa (Sigma, CAS# 9001–75–6) in 0.1 M HCl for 90 s at 37°C, followed by washing in PBS and dehydration by serial incubation in ethanol (70%, 85%, 100%).

Mouse fibroblast MEF and human fibroblast BJ-hTERT cell lines were cultured as described above except the medium was not supplemented with uridine. Human colorectal carcinoma HCT 116 cells were cultured in McCoy's 5a medium (Gibco) supplemented with 10% fetal bovine serum (Sigma), 2 mM L-glutamine (Sigma) and 1 × Penicillin-Streptomycin (PEST, Sigma). Cells were incubated overnight and washed in PBS before fixation with 3.7% paraformaldehyde in PBS for 15 min at room temperature, followed by two washes in PBS and dehydration in ethanol (70%, 85%, 100%; 3 min each). For RNA analysis, all solutions used after fixation were DEPC treated or prepared in DEPC treated water. Before reverse transcription, cells were washed in PBS, incubated in 0.1 M HCl for 3 min at room temperature, followed by two washes in PBS.

Subsequent incubations and reactions were performed in a volume of 50 μl in a 9 mm diameter Secure-Seal hybridization chamber (Invitrogen) for experiments on cell lines and in 100 or 200 μl volume (respectively 13 and 20 mm diameter, respectively) Secure-Seal hybridization chambers for experiments on tissue sections. The chamber was covered with a piece of common plate sealing tape for incubations lasting longer than 30 min to avoid evaporation.

Restriction digestion and exonucleolysis for mitochondrial DNA detection

We adapted a previously described protocol for *in situ* detection of mitochondrial DNA to work with the padlock gap probes (6). Genomic and mitochondrial DNA was restriction digested using 0.5 U/μl of *Msc*I (New England Biolabs) in NEB buffer 4 supplemented with 0.2 μg/μl BSA (New England Biolabs) at 37°C for 30 min. Exonucleolysis was carried out in the same step by adding 0.4 U/μl of the 5′-3′ T7 exonuclease (New England Biolabs) or the 5′-3′ λ exonuclease (New England Biolabs) to the reaction mix. After the incubation, slides were rinsed in wash buffer (0.1 M Tri-HCl pH 7.5, 0.15 M NaCl and 0.05% Tween-20).

Tissue sections

Tissue samples, fresh-frozen and in FFPE format, were obtained from a patient cohort of non-small cell lung cancer (15). Samples and data were handled in accordance with the Swedish Biobank Legislation and Ethical Review Act (reference 2006/325, Ethical review board in Uppsala). Fresh frozen sections were fixed in 3.7% paraformaldehyde in PBS for 45 min at room temperature, washed twice in PBS and permeabilized for 5 min in 0.01% pepsin (Sigma, CAS# 9001–75–6) in 0.1 M HCl at 37°C. After permeabilization, slides were washed in milli-Q water for 5 min, followed by a 2 min PBS wash and dehydration in ethanol (70%, 85%, 100%; 1 min each). FFPE sections were baked at 60°C for 30 min, immersed in xylene for 15 + 10 min and rehydrated in ethanol (2x 100%, 2x 95%, 2x 70%; 2 min each) and subsequently washed in milli-Q water for 5 min and in PBS for 2 min. After deparaffinization, the FFPE sections were fixed in 3.7% paraformaldehyde in PBS for 10 min at room temperature and washed in PBS for 2 min. Permeabilization was done by incubation in 0.01% pepsin in 0.1 M HCl for 30 min at 37°C, followed by a wash in milli-Q water for 5 min and one wash in PBS for 2 min. Samples were dehydrated by a serial ethanol incubation (70%, 85%, 100%; 1 min each).

cDNA synthesis for mRNA detection

cDNA synthesis was performed in the supplied M-MuLV RT reaction buffer (50mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT), 0.5 mM dNTPs, 0.2 μg/μl BSA (New England Biolabs), 1 μM of each primer and 10 U/μl of TRANSCRIPTME M-MuLV reverse transcriptase (DNA Gdansk, Poland) for 2 h at 45°C when testing cell lines and 3 h when testing tissue sections. Slides were washed twice in PBS and then fixed in 3% paraformaldehyde for 5 min and washed twice in PBS before probe hybridization and ligation.

Padlock gap probe hybridization and gap probe ligation

The sequences of the oligonucleotides used are listed in Table 1. Because an available phosphate group on the 5′-end

Table 1. Oligonucleotide sequences. Probes were designed based on accession numbers NC_012920.1 (mtDNA), NM_001101.3 (human *ACTB*), NM_007393.3 (mouse *ACTB*), KJ685774.1 (*EGFR*) and NM_004985.4 (*KRAS*)

	Probe name	Sequence
Padlock Probes	PdMelas_WT	5′-CTGCCATCTTAACAAACCCTCAATGCTGCTGCTGTACTACTTTTATGCGA TTACCGGGCT-3′
Padlock Gap Probes	PdGp6	5′-CCATCTTAACAAACCCTGTTCTTGGTTCCTCTATGATTACTGACCT ACCTCAATGCTGCTGCTGTACTACTCTTCTTATGCGATTACCGG-3′
	PdGp5	5′-GCCATCTTAACAAACCCTGTTCTTGTTCCTCTATGATTACTGACCT ACCTCAATGCTGCTGCTGTACTACTCTTCTTATGCGATTACCGG-3′
	ActbGp6	5′-GCTTCGCGGGCGACGTTCCTTTTACG ACCTCAATGCTGCTGCTGTACTACTCTTGGCTCCGGCATGTGCA-3′
	KrasGp6	5′-GTAGGCAAGAGTGCCTCCTCTATGATTACTGACCAAGA TCCCTCAATGCTGCTGCTGTACTACGGTTCAAGTGTGGTAGTTGGAGCT-3′
	EGFRGp6	5′-AAACTGCTGGGTGCGTCCTAGTAA TCAGTAGCCGTGACTATCGACTGGTTCAAAGATCACAGATTTTGGG-3′
Gap Probes	Gp5_WT	5′- GCTCT -3′
	Gp_WT	5′- GCTCTG -3′
	Gp_6T	5′- GCTCT**T** -3′
	Gp_6A	5′- GCTCT**A** -3′
	Gp_5C	5′- GCTC**C**G -3′
	Gp_5G	5′- GCTC**G**G -3′
	Gp_4A	5′- GCT**A**TG -3′
	Gp_3A	5′- GC**A**CTG -3′
	Gp_3G	5′- GC**G**CTG -3′
	Gp_MELAS	5′- GC**C**CTG -3′
	Gp_2A	5′- G**A**TCTG -3′
	Gp_2T	5′- G**T**TCTG -3′
	Gp_1A	5′- **A**CTCTG -3′
	Gp_Act_mus	5′-AAGCCG-3′
	Gp_Act_hum	5′-AGGCCG-3′
	WT_KRAS	5′-GGTGGC-3′
	G13D	5′-GGTG**A**C-3′
	G12D	5′-G**A**TGGC-3′
	WT_EGFR	5′- CTGGCC-3′
	2572 C>A	5′- **A**TGGCC-3′
	2572 C>T	5′- **T**TGGCC-3′
	2573 T>A	5′- C**A**GGCC-3′
	2573 T>G	5′- C**G**GGCC-3′
	2574 G>A	5′- CT**A**GCC-3′
	2575 G>A	5′- CTG**A**CC-3′
	2576 C>A	5′- CTGG**A**C-3′
Detection oligo	DO_Cy3	5′-Cy3–CCTCAATGCTGCTGCTGTACTAC-3′
	DO_Cy5	5′-Cy5–CCTCAATGCTGCTGCTGTACTAC-3′
	DO_EGFR	5′-Cy3-AGTAGCCGTGACTATCGACT-3′
	DO_EGFR_U	5′-Cy3-AGUAGCCGUGACUAUCGACU-3′
Sequencing library	Cy51NA	5′-PO_4-NANNNNNNN-Cy5–3′
	Cy31NG	5′-PO_4-NGNNNNNNN-Cy3–3′
	FITC1NT	5′-PO_4-NTNNNNNNN-FITC-3′
	TR1NC	5′-PO_4-NCNNNNNNN-TR-3′
	TR0NC	5′-PO_4-CNNNNNNNN-TR-3′
cDNA Primers	P_Actb_hum	5′-C+TG+AC+CC+AT+GC+CC+ACCATCACGCCC-3′
	P_Actb_mou	5′-C+TG+AC+CC+AT+TC+CC+ACCATCACACCC-3′
	P_KRAS	5′- C+CT+CT+AT+TG+TT+GG+ATCATATTCGTC
	P_EGFR	5′- G+TA+TT+CT+TT+CT+CT+TCCGCAC-3′
Anchor Primer	Act_AnchP	5′-GGCUCCGGCAUGUGCA-3′
	EGFR_AnchP	5′- AAAGAUCACAGAUUUUGGG-3′

Underlined are the sequences complementary to the detection oligo; mismatches in the gap probes are in bold; a + before the nucleotide indicates a LNA modified base.

of the ligation substrate is an absolute requirement for enzymatic ligation reactions, the padlock gap probe and the gap probes were phosphorylated with 0.2 U/μl of T4 PNK enzyme (Thermo Scientific) in supplied buffer A and 1 mM ATP at 37°C for 30 min, followed by inactivation of the enzyme at 65°C for 10 min. Hybridization and ligation were performed in a single reaction containing 0.1 μM of the padlock gap probe and 2 μM of each gap probe (1 μM when 7 gap probes were used to test FFPE samples) using 0.5 U/μl of Ampligase (Epicentre) in 1x Ampligase buffer (20 mM Tris-HCl pH 8.3, 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM NAD and 0.01% Triton X-100), 0.2 μg/μl BSA and 0,05 mM KCl (125 mM KCl was used for mtDNA detection) at 45°C for 60 min. When T4 DNA ligase was compared to Ampligase, 0.1 U/μl of T4 DNA ligase (Thermo Scientific) in 1X T4 DNA ligation buffer (40 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP) supplemented with 0.2 μg/μl BSA, 200 nM ATP and 200 mM NaCl was used to

ligate 0.1 μM of the padlock gap probe and 2 μM of each gap probe for 30 min at 37°C. After ligation the slides for mitochondrial DNA detection were incubated 5 min at 37°C in 2x SSC, 0.05% Tween-20, and rinsed in wash buffer to remove non-hybridized probes. All other slides were washed twice in PBS-Tween for 5 min each.

Rolling circle amplification and fluorescence staining

The RCA reaction was performed using 1 U/μl of Φ29 DNA Polymerase (Thermo Scientific) in the supplied Φ29 buffer, 0.25 mM dNTPs, 0.2 μg/μl BSA and 5% of glycerol at 37°C for about 2 and 5 h when testing cell lines and tissue sections, respectively. Slides were then rinsed in PBS-Tween (wash buffer for mitochondrial DNA detection) and RCA products were detected by hybridization of 100 nM of fluorescence-labeled detection oligonucleotide probe in 2x SSC with 20% formamide for 30 min at 37°C. After brief washes in PBS-Tween (wash buffer for mitochondrial DNA detection), Secure-Seals were removed and slides were dehydrated in ethanol (70%, 85% and 100% for 3 min each) before being mounted in Vectashield mounting medium (Vector Laboratories) containing 10 ng/ml DAPI. Slides containing MEF or BJ-hTERT cells were stained with Hoechst (1 μg/ml) for 20 min at room temperature and tissue slides were stained with DAPI (2 μg/ml) for 10 min at room temperature before removing Secure-seals. Slides were washed in PBS and dehydrated in ethanol (70%, 85%, 100%; 3 min each) and mounted using SlowFade Gold antifade reagent (Life Technologies).

In-situ sequencing by ligation

The sequencing of the gap probes incorporated into the RCPs was performed as described in Ke *et. al.* (13). For fibroblast containing slides, only the SNV on the *ACTB* transcript was sequenced using the Cy51NA and Cy31NG sequencing libraries in the ligation reaction to the Act_AnchP anchor primer on the sequencing substrate. After the sequencing cycle, nuclei were counterstained again with Hoechst (1 μg/ml) for 20 min. Cells were washed in PBS and dehydrated by a serial incubation in ethanol before mounting in SlowFade Gold antifade reagent (Life Technologies). In tissue samples, the first three bases of *EGFR*-specific gap probes were sequenced using the full sequencing library and the *EGFR*_AnchP anchor primer (Table 1).

Image acquisition and analysis

For mtDNA analysis, slides were imaged using an Axioplan II epifluorescence microscope (Zeiss) equipped with a 100W mercury lamp and excitation and emission filters for DAPI, Cy3 and Cy5, and with a CCD camera (C4742–95, Hamamatsu) for image acquisition. A 20x or 40x objective (Plan-neofluar, Zeiss) was used and images were collected using the AxioVision software (version 4.4, Zeiss). For RNA analysis, automatic scanning of the slides was performed at 20x magnification with an Imager.Z2 epifluorescence microscope (Zeiss) equipped with 20x objective (Plan-neofluar, Zeiss) and a CMOS Orca-Flash 4.0 camera (C11440, Hamamatsu). For experiments on cell lines, RCPs were quantified using CellProfiler software (version r10997) (16) to extract the number of cells in each single image and the total number of RCPs. The mean number of RCPs per cell was then calculated for each tested condition. All statistical analyses were performed using a two-tailed Mann–Whitney U test (GraphPad Prism, CA, USA). CellProfiler and Matlab scripts were used to analyze images and to do base calling for *in situ* sequencing as described previously (13). Base-calling in cell lines experiments was made using images acquired with Cy3 and Cy5 filters and two black images were used for FITC and TexasRed filters because these fluorophores were not used in the sequencing reaction. The quality threshold (Qt) for base calling was set to 0.5. The scripts are available for downloading at http://www.cellprofiler.org/examples.shtml.

Images from tissue sections were analyzed using CellProfiler software (version 2.1.1 rev. 6c2d896). Tissue samples can show considerable level of auto-fluorescence background, particularly in the green part of the fluorescent spectra, which can give rise to false signals. These false signals typically can be observed at different wavelengths, in contrast to true signals that are present only at the wavelength of the used fluorophore. Therefore, an extra 'empty' channel can be imaged and used to identify false signals (17). Because RCPs were stained in Cy3, we imaged tissue sections using DAPI for nuclear staining, Cy3 for RCPs and FITC to identify fluorescent background. Before quantifying the RCPs, for each image the three channels were exported as RGB images and converted to grayscale. The intensity from the green image (FITC) was subtracted from the red image (Cy3) and the resulting image was used to quantify the RCPs. Sequencing analysis of RCPs in tissue samples was done in the same way as described in Ke. *et al.* (13). The nuclear staining was used to quantify the area of the tissue imaged from each sample. Briefly, a mask was created in Cellprofiler, based on the intensity level, assigning each pixel to foreground or background and the number of pixels belonging to the foreground was calculated. Based on the pixel size at the imaging condition the imaged tissue area can be calculated.

Images featured in the publication were thresholded for visualization in Adobe Photoshop CS3 (Adobe Systems Inc).

RESULTS

Design of the padlock gap-fill ligation assay

Padlock gap probes were designed to have arms that hybridize six nucleotides apart from each other on the targeted mtDNA or cDNA (Figure 1). In this assay, short sequence-specific gap probes, phosphorylated at the 5′-end, are added to circularize the padlock gap probe in a double ligation reaction. In absence of mutations, the wild type gap probe perfectly matches the entire target region and will be able to undergo the ligation reaction. After *in situ* RCA and hybridization of fluorescence-labeled probes to the RCP, the presence of signals confirms the wild type sequence. If one or more mutations are present within the target region, ligation of the wild type probe is prevented, and thus no signal is detected for those bound probes. In the multiplex format of

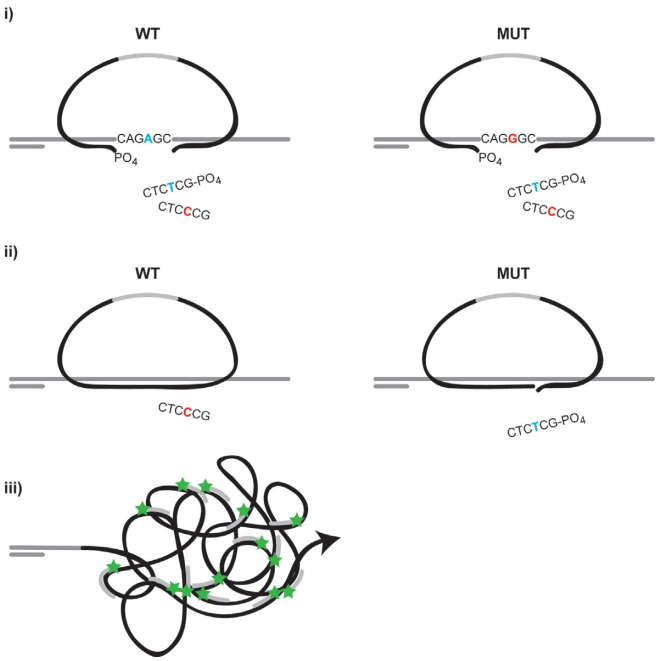

Figure 1. Schematic representation of a two-plex oligonucleotide gap-fill ligation assay for detection of an SNV. The padlock gap probe (black) hybridizes to a single stranded target sequence (gray) leaving a six nucleotide gap between the 5′- and 3′-ends of the probe (i). In a competitive two-plex reaction, wild-type 5′-phosphorylated and mutant non-phosphorylated gap probes are present in the reaction mix. In presence of a wild-type target, the wild-type gap probe will hybridize, circularize the padlock gap probe and will be amplified via RCA, generating an RCP detectable with fluorescently-labelled oligonucleotides (i–iii left). If a mutation is present in the target region, a mutant gap probe hybridizes more stably to the target, but as it is not phosphorylated ligation at the 3′-end of the padlock gap probe will not occur. This prevents RCA of the padlock gap probe and no signal is generated (i–ii right). Adding non-phosphorylated wild-type gap probe and 5′-phosphorylated mutant gap probe would in the same scenario result in the detection of mutant sequence whereas no signal would be generated at wild-type loci.

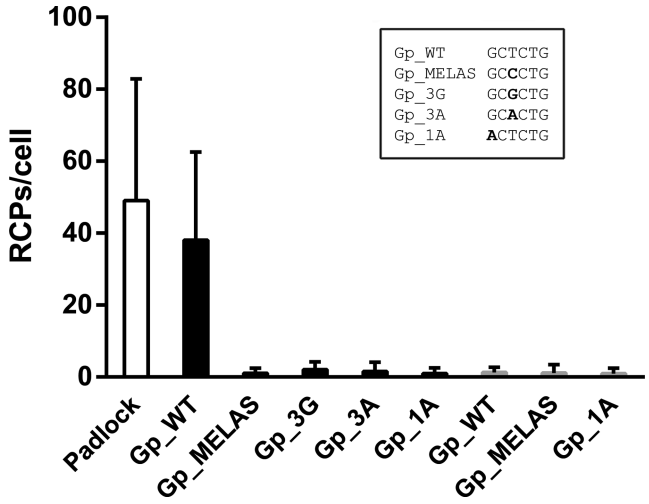

Figure 2. Efficiency and specificity of *in situ* gap-fill ligation for mtDNA detection in the 143B cell line. The efficiency of the gap probe ligation in singleplex reactions is compared to a regular padlock probe. The white bar indicates the padlock probe; black bars are for the different hexamers included in each gap-fill ligation reaction. The sequence of each specific gap probe is reported in the box with specific mutations indicated in bold. Error bars = SD.

the assay, one or more mutant gap probes are used in combination with the wild-type gap probe. All 5′-phosphorylated gap probes present can undergo ligation and generate a signal, but the identity of the ligated gap probe is indistinguishable when using fluorescence *in situ* hybridization for detection. Therefore, to reveal presence of a specific sequence in the target region, only one gap probe of specified sequence is 5′-phosphorylated at a time, and the presence of signal therefore indicates that sequence being present in the sample. Alternatively, the presence of mutations can be revealed by direct *in situ* sequencing of the inserted gap probe. In the case of *in situ* sequencing, all gap probes participating in the reaction can therefore be phosphorylated and all assayed sequence variants at the locus can be recorded simultaneously at read out.

Mutation detection on mtDNA

To evaluate the performance of oligonucleotide gap-fill ligation we first examined the sensitivity and specificity of the method by detecting mtDNA sequences *in situ*. The new approach requires two ligations for circularization of the padlock gap probe compared to a single one required for conventional padlock probes. To evaluate if this requirement has any effect on the detection efficiency in our assay we compared the performance of the perfect match gap probe hexamer in padlock gap-fill ligation to a regular padlock probe targeting the MELAS locus on the mtDNA of a wild-type cell line. The padlock gap-fill ligation resulted in 20% reduction of the number of obtained RCPs compared to the regular padlock probe (Figure 2). Specificity is an important aspect when it comes to mutation detection. It has been reported that different nucleotide mismatches between the ligating oligonucleotides and their template, as well as the position of the mismatch relative to the ligation site, affect ligation specificity to different extents (18). To evaluate the specificity in our assay, we tested five different gap probes in singleplex reactions for their ability to give rise to signal when detecting the wild-type sequence (3243A) at the MELAS locus. Mismatched sequences resulted in maximum 2 RCPs counted per cell, demonstrating high specificity of our assay (Figure 2). Control reactions in which the gap probes had not been 5′-phosphorylated yielded signals similar to the negative control where no ligase was added to the reaction. To evaluate if we could increase the number of signals by using an alternative enzyme in the ligation reaction we tested the specificity of ligation using the thermostable Ampligase and the mesophilic T4 DNA ligase. We compared the enzymes in two seven-plex reactions where the wild-type gap probe or the MELAS specific gap probes were phosphorylated. Ligation by T4 DNA ligase showed to be twice more efficient (136 and 69 signals per cell). The use of Ampligase however resulted in a higher signal-to-noise ratio compared to the T4 DNA ligase, calculated as wild-type signal:MELAS signal (115 and 3, respectively).

In order to demonstrate the applicability of the padlock gap-fill ligation approach to detect a clinically relevant mutation, we pooled seven gap probes, the perfect match gap

probe and six mutant hexamers, each mismatched to a specific base at a different position of the target gap. We then performed seven reactions in parallel, each reaction having only one probe in the pool of gap probes 5′-phosphorylated at a time whereas the remaining six probes were all non-phosphorylated. In each reaction, the 5′-phosphorylated gap probe was thus the only probe capable of undergoing a ligation reaction following a perfectly matched hybridization at the padlock gap site. We found that whereas the reaction that included the perfectly matched 5′-phosphorylated wild-type gap probe yielded abundant signals, none of the other reactions resulted in more than one RCP per cell (Figure 3A–C). The same pool of seven gap probes was then used to test a cell line with the MELAS mutation present on the mitochondrial DNA at a high level of heteroplasmy (>98%) (19). In three parallel reactions, one with 5′-phosphorylated wild-type gap probe, one with 5′-phosphorylated MELAS mutant specific gap probe, and a negative control where none of the gap probes were 5′-phosphorylated, we detected signals over the background level only for the MELAS mutation-specific gap probe, indicating the presence of the MELAS mutation in this cell line (Figure 3D–F).

SNV detection on mRNA

In addition to mutation profiling on mtDNA, we wanted to investigate the applicability of the oligonucleotide gap-fill ligation mechanism for capture and sequencing of mRNA molecules *in situ*. As a proof of concept we chose to detect an SNV at position 133 of the *ACTB* mRNA transcript, differentiating mouse from human *ACTB* in a conserved region of the transcript. We tested *ACTB* detection in the human BJ-hTERT and a mouse embryonic fibroblast cell line in a set of parallel reactions with different reaction conditions. In the first reaction only the human gap probe was 5′-phosphorylated while the mouse gap probe was non-phosphorylated. In the second reaction, the mouse gap probe was 5′-phosphorylated and the human gap probe was non-phosphorylated. To demonstrate that signals are dependent on the presence of a 5′-phosporylated gap probe we included one reaction with both of the gap probe hexamers being non-phosphorylated, and one reaction with no gap probe present. As expected, only the presence of the phosphorylated, fully complementary hexamers yielded a signal significantly above the level of the control reactions (30 and 31 RCPs per cell for BJ-hTERT and MEF, respectively) (Figure 4). The control conditions had fewer than 2 RCPs per cell, which was comparable to when no cDNA primers were included for cDNA synthesis, demonstrating that our method is specific and can be applied to mRNA genotyping.

To show the potential of this assay for multiplex mutation detection, we performed *in situ* sequencing-by-ligation (13) of the SNV which is incorporated into the RCPs. All experiments were performed on human BJ-hTERT and mouse embryonic fibroblast cells in separate cytological preparations. First we tested the sequencing error rate of our gap-fill approach combined with *in situ* sequencing. Either the 5′-phosphorylated human gap probe or the 5′-phosphorylated mouse gap probe was added to each reaction with the other gap probe (mouse and human, respectively) being present but non-phosphorylated. We sequenced more than 99% of all detected RCPs, scoring correctly RCPs in both human and mouse cell lines with a sequencing error rate lower than 1% (Figure 5 and Table 2). Nonspecific ligation detected in the reciprocal conditions was lower than one RCP per cell. Furthermore, we tested the approach for specific multiplex mutation detection, where in order to use the method, all gap probes present in the reaction need to be 5′-phosphorylated. Using these conditions to genotype the SNV in *ACTB* in human and mouse fibroblast cells again resulted in detection of human- and mouse-specific RCPs with low error rate, but with a small drop in the total number of RCPs per cell (Figure 5 and Table 2). Omitting the primer for cDNA synthesis yielded less than one sequenced RCP per cell confirming that sequenced RCPs are the product of specific cDNA detection. To demonstrate mutation detection of a lower expressed transcript we targeted the mutation G13D in the *KRAS* gene and analyzed the cell line HCT 116, heterozygous for this specific mutation (cancer.sanger.ac.uk) (20). We used specific gap probes to target the two alleles and tested them with only one of the hexamers being 5′-phosphorylated at a time. When normalized by cell density, the expression of the two alleles resulted in similar level, yielding 1.5 more wild-type RCPs than mutant RCPs. Importantly, we designed an additional hexamer targeting a different *KRAS* mutation in the same locus (G12D) and tested it in HCT 116 cells resulting in a about 30 times lower number of RCPs than using the gap probe for the G13D mutation (Supplementary Figure S1 and Table S1). These results provide support that oligonucleotide gap-fill ligation and *in situ* sequencing as described herein can be used for *in situ* multiplex mutation detection with a low error rate allowing for potential use in clinical diagnostics.

Mutation detection in tumor tissue samples

To demonstrate the applicability of the method on clinical patient samples, we tested fresh frozen and FFPE lung tissues for presence of clinically relevant mutations in *EGFR* which are predictive biomarkers for targeted treatment with *EGFR* tyrosine kinase inhibitors. In a first set of experiments on fresh frozen samples we used two gap probes, one wild-type probe and one probe specific for the mutation L858R (c.2573T>G). We successfully detected the c.2573T>G mutation in the cancer cells of an L858R mutation positive lung cancer tissue (Figure 6). The density of mutant signal in a normal tissue section was at the level of background measured in negative control experiment while the level of wild-type signals was similar in both tumor and normal tissue (Table 3 and Supplementary Figure S2). To demonstrate the capacity for simultaneous analysis of multiple mutations in FFPE tissue samples, we included five additional gap probes targeting the most frequent mutations in codon 858 and 859 in exon 21 of *EGFR* (Table 1). Using the complete set of phosphorylated mutant gap probes, the tissue harboring the L858R mutation resulted in a 30 times higher mutant signal density compared to an *EGFR* WT tissue. Importantly, the level of non-specific signal detected was lower than the level measured in a negative control sample, where no primer for cDNA synthesis was added

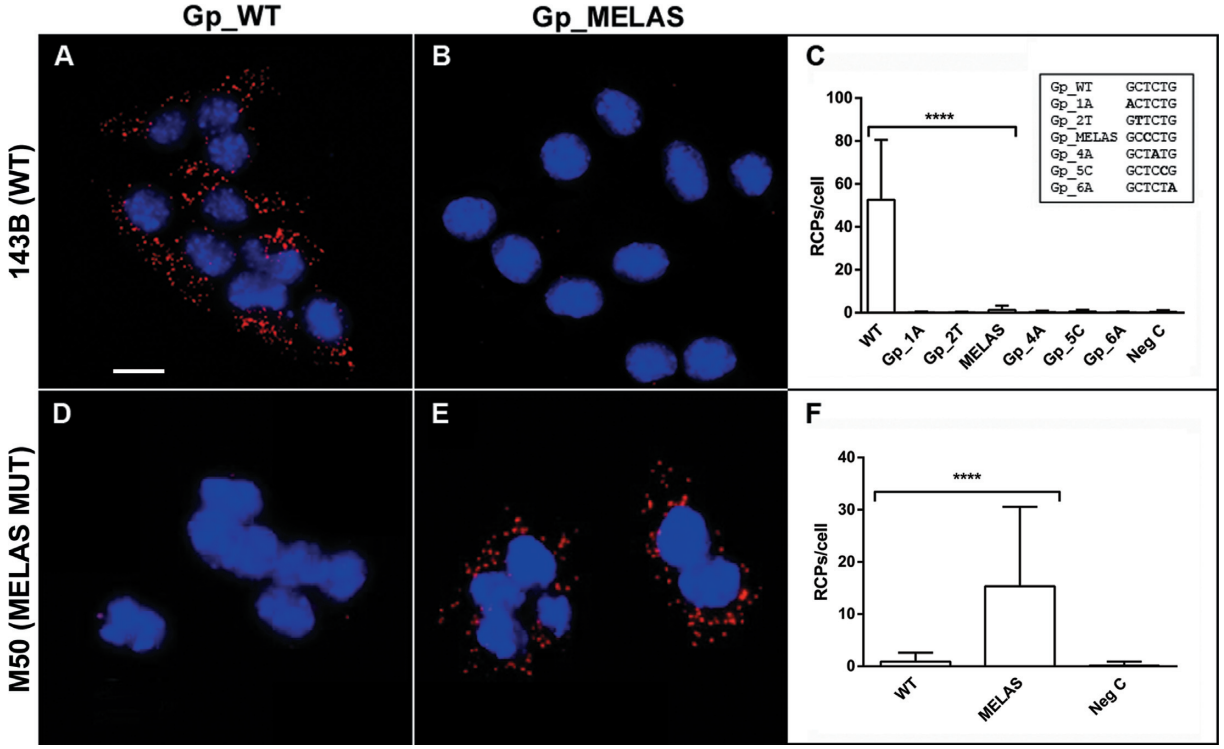

Figure 3. *In situ* mtDNA mutation detection in 143B and M50 cell lines using padlock gap-fill ligation. Six gap probes with different mutations, including the MELAS A3243G mutation, were combined with the perfect match wild-type gap probe and used to target the MELAS mtDNA mutation in wild-type 143B (A–C) and mutant M50 (D–F) cell lines. Seven reactions were performed for each cell line with only one 5′-phosphorylated hexamer included in each reaction. (**A** and **D**) Representative images for respective cell line with wild-type 5′-phosphorylated gap probe. (**B** and **E**) Representative images for respective cell line with MELAS mutation-specific 5′-phosphorylated gap probe. (**C** and **F**) Quantification of the results obtained for each parallel reaction with the phosphorylated hexamer present in each reaction indicated on the *x*-axis. The respective sequences are reported in the box with specific mutations in bold. As negative controls all hexamers included in the reactions were non-phosphorylated. Error bars = SD. *P*-values were obtained with a two-tailed Mann–Whitney U test (**** = $P < 0.0001$).

Table 2. *In situ* sequencing results on human and mouse fibroblasts cell lines

Condition[a]	Cells	Tot RCPs	Sequenced RCPs (%)	RCPs/Cell	G (Hum)	A (Mus)	Error rate (%)
Sequencing of human BJ-hTERT cells							
Hum-P/Mus	1093	34726	99.8	31.7	34624	38	0.1
Hum/Mus-P	738	380	100	0.5	371	9	NA
Hum-P/Mus-P	579	15026	99.3	25.8	14893	39	0.3
Neg C	440	11	100	0.025	10	1	NA
Sequencing of mouse embryonic fibroblast cells							
Hum-P/Mus	415	342	100	0.8	302	40	NA
Hum/Mus-P	505	16386	99.8	32.4	51	16308	0.3
Hum-P/Mus-P	606	14567	99.6	23.9	144	14370	1
Neg C	473	24	100	0.05	16	8	NA

[a]hexamers present in the reaction; -P = 5′-phosphorylated hexamer.

(Table 3 and Supplementary Figure S3). In order to identify the mutant hexamer ligated to the padlock gap probe we sequenced three bases of the mutant hexamer incorporated into the RCPs of the *EGFR* mutation positive sample. More than 90% of the previously detected RCPs were sequenced. 98.5% of the reads mapped to the expected 2573 T>G mutation with a sequencing error rate of 0.5%, confirming the specificity of the method (Table 4). These data demonstrate the applicability of the assay for highly specific mutation detection in patient tissues for a potential diagnostic use.

DISCUSSION

We have investigated the use of *in situ* oligonucleotide gap-fill ligation as a method to reveal presence of mutations in mitochondrial DNA and mRNA molecules. In this method, cocktails of gap probes are used to compete for hybridization on the target region and ligation to the padlock gap probe. Combined with RCA, the detection events can be visualized and sequenced *in situ*. We showed that this method can efficiently reveal the presence of mutations in the gap region to a high degree of specificity and we applied it to screen for the MELAS point mutation directly on mtDNA

Table 3. *In situ* detection of *EGFR* mutations in lung tissue

Sample type	Status	Condition[a]	RCPs/mm^2
Fresh Frozen	Normal	2573 T>G-P/WT	1.3
		2573 T>G/WT-P	5.1
Fresh Frozen	L858R	2573 T>G-P/WT	40.3
		2573 T>G/WT-P	4.5
Fresh Frozen	Normal	no gap oligo	0.7
FFPE	L858R	MUT-P/WT	12.7
		MUT/WT-P	7.5
		Neg ctr	0.6
FFPE	*EGFR* WT	MUT-P/WT	0.4
		MUT/WT-P	1.7

[a]hexamers present in the reaction; -P = 5′-phosphorylated hexamer; MUT = six gap probes coding for reported mutations in codon 858 and 859 in exon 21 in *EGFR*; Neg ctr = no cDNA primer.

Table 4. *In situ* sequencing of *EGFR* mutations in FFPE lung tissue

Mutation	Sequence[a]	Reads	% of tot. reads
2573 T>G	C**G**G	333	98.5
2575 G>A	CT**G**	2	0.6
2572 C>A	A**T**G	0	0
2572 C>T	T**T**G	0	0
2573 T>A	CA**G**	0	0
2574 G>A	CT**A**	0	0
Unexpected	GGT	1	0.3
Unexpected	GTG	1	0.3
Unexpected	TGG	1	0.3
Tot. reads		338	100

[a]Letters in bold indicate mutations.

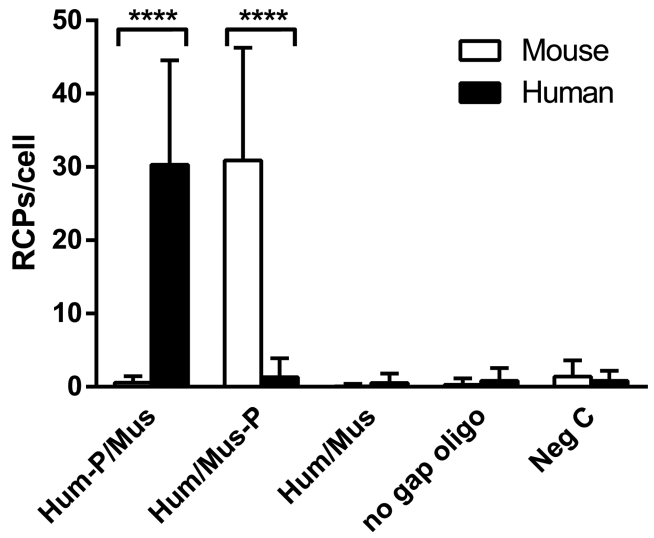

Figure 4. *In situ* genotyping of *ACTB* mRNA in human BJ-hTERT and mouse embryonic fibroblasts. Human and mouse fibroblasts were cultured and separately tested for a species-specific SNV on *ACTB* mRNA. Two gap probes, human and mouse specific, were combined in two reactions where only one of the two was 5′-phosphorylated as indicated on the *x*-axis. Error bars = SD. *P*-values were obtained with a two-tailed Mann–Whitney U test (**** = $P < 0.0001$).

of fixed cells. We furthermore established the gap-fill protocol combined with *in situ* sequencing for multiplexed readout by genotyping a SNV on *ACTB* mRNA and to detect the L858R point-mutation (c.2573T>G) in the *EGFR* gene in fresh frozen and FFPE tumor samples with high specificity. This single mutation is the most prevalent point mutation making up approximately 40% of all mutations in *EGFR*-mutated non-small cell lung cancer (21). *EGFR* mutations in non-small cell lung cancer are a positive predictor for the efficacy of *EGFR* tyrosine kinase inhibitors and are used in clinical diagnosis for patient stratification (22). This is the first report performing *in situ* sequencing of RCPs generated in FFPE tissues which allows simultaneous mutation detection and demonstrates the potential of this assay for detection of clinically relevant point mutations directly in routine tissue sections. Padlock probe-based assays are well recognized for their high specificity, which is determined by the fidelity of the ligation reaction. The oligonucleotide gap-fill ligation presented herein is dependent on the DNA ligase to be able to discriminate between mismatched nucleotides not only at the position immediately adjacent to the ligation site, but also in the middle of the gap probe. A systematic analysis of the ligation rate of all possible single nucleotide mismatched gap probes would require comparison of 18 mutant hexamer sequences against the wild-type. Under the assumption that two or more mutations in the same hexamer represent a less favorable condition for misligation to happen, we only investigated the frequency of nonspecific ligation for the three possible mismatches at the most distant position from the two ligation sites in single probe reactions. As shown in Figure 2, the A::G mismatch at the third position from the 5′-end of the gap probe showed the highest frequency of ligation among the three mismatches, but was still 16-fold lower than the perfect match (A::T) gap probe measured in the same experiment. This result is in agreement with literature showing that C::A and T::G misligation can occur, although at a reduced ligation rate, be-

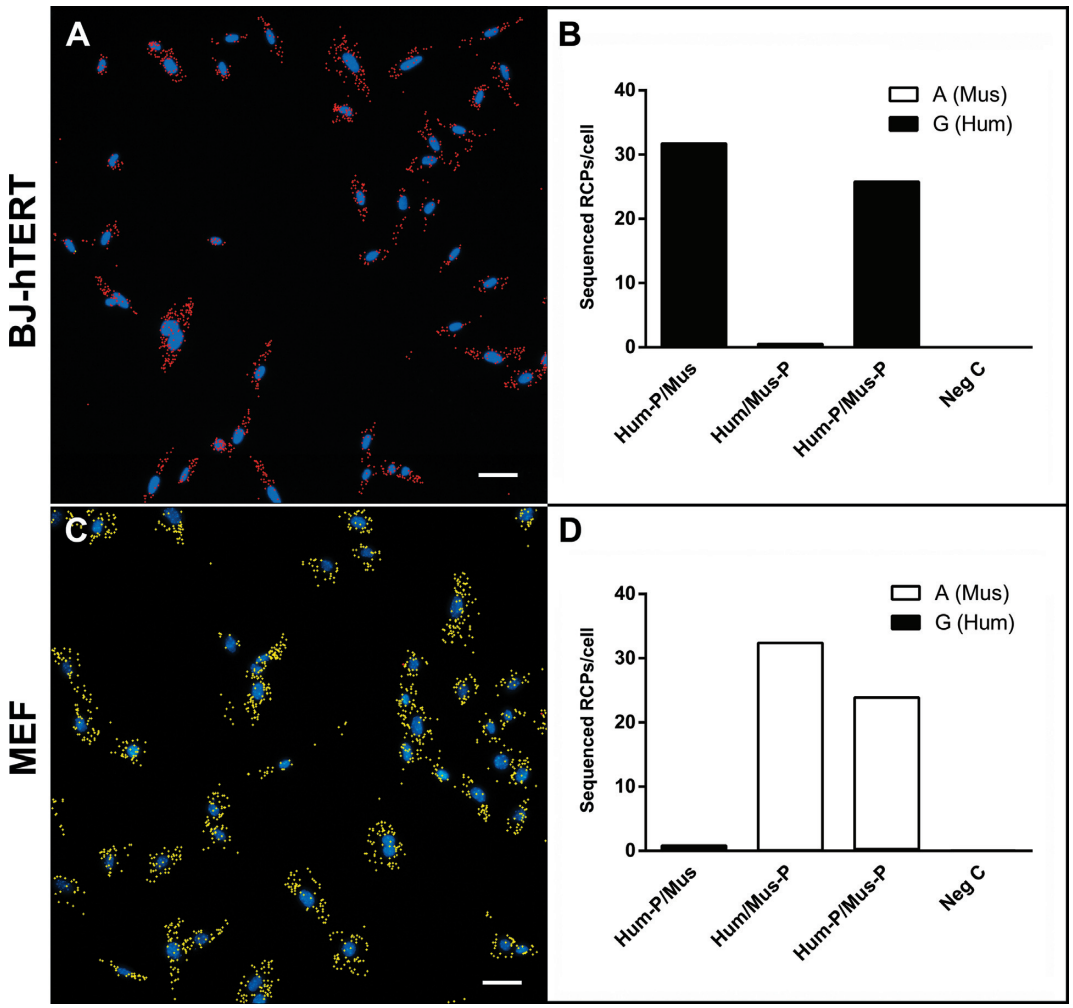

Figure 5. *In situ* sequencing of an expressed SNV in human and mouse fibroblasts. Human BJ-hTERT and mouse embryonic fibroblasts were subjected to *in situ* sequencing of RCPs to detect a SNV in *ACTB* mRNA. Human and mouse specific hexamers were combined in four reactions where only one gap probe or both of the hexamers were 5′-phosphorylated, or where the primers for cDNA synthesis were omitted. (**A** and **C**) Representative images for respective cell lines of the reactions where both gap probes (hum and mus) were 5′-phosphorylated. Cell nuclei are shown in blue and the base called for each RCP is plotted on the cell images in red for human *ACTB* and yellow for mouse *ACTB*. Scale bars 20 μm. (**B** and **D**) Quantification of the four reactions showing the mean number of identified RCPs for the two *ACTB* variants represented for each cell line.

cause their pair geometry causes the smallest disruption to the double stranded structure among all possible base pairs (23,24). Increasing the complexity of the hybridization reactions to six mutant gap probes competing with the wild-type for the target, did not result in a higher rate of misligation, as shown in Figure 3. Furthermore, whether performed on mtDNA or mRNA in cells or on tissue, strong significant differences were observed between results of wild-type and mutant-specific gap probes, indicating that the assay is specific in detecting the latter condition (Figures 3 and 4, and Table 3). The overall high specificity of our assay may be explained by the high temperature at which ligation was performed (45°C), well above the T_m of the gap probes, and by the fact that two ligation events need to happen to circularize the padlock gap probe (12). Specificity is a desirable characteristic particularly for clinical applications, but it comes to the price of lower efficiency compared to the performance of a regular padlock probe. We estimated the decrease in efficiency to about 20% when only the perfect match gap probe was added to the ligation reaction. This difference was however less in the latter experiments where competing mismatched gap probes were included in the reaction together with the matched gap probe, for reasons that we do not yet understand. This caused us to try to use fully degenerated 5′-phosphorylated hexamers to circularize the padlock gap probe, which did however not result in generation of any detectable signal (data not shown). This failure may be due to incomplete ligation of gap probes which are complementary only at one end, i.e. 5′-end, but containing mismatches at the 3′-end, thus inhibiting full circularization of the padlock probe and decreasing the overall efficiency, but further experiments are required to confirm this hypothesis. Assessing the sensitivity of this method in detecting point mutations at transcript level is technically challenging because of the biological variation in mRNA expression at single cell level. To give an estimation of the lower limit of detection of this method, we decided to target a point mutation in the *KRAS* gene in the heterozygous cell line

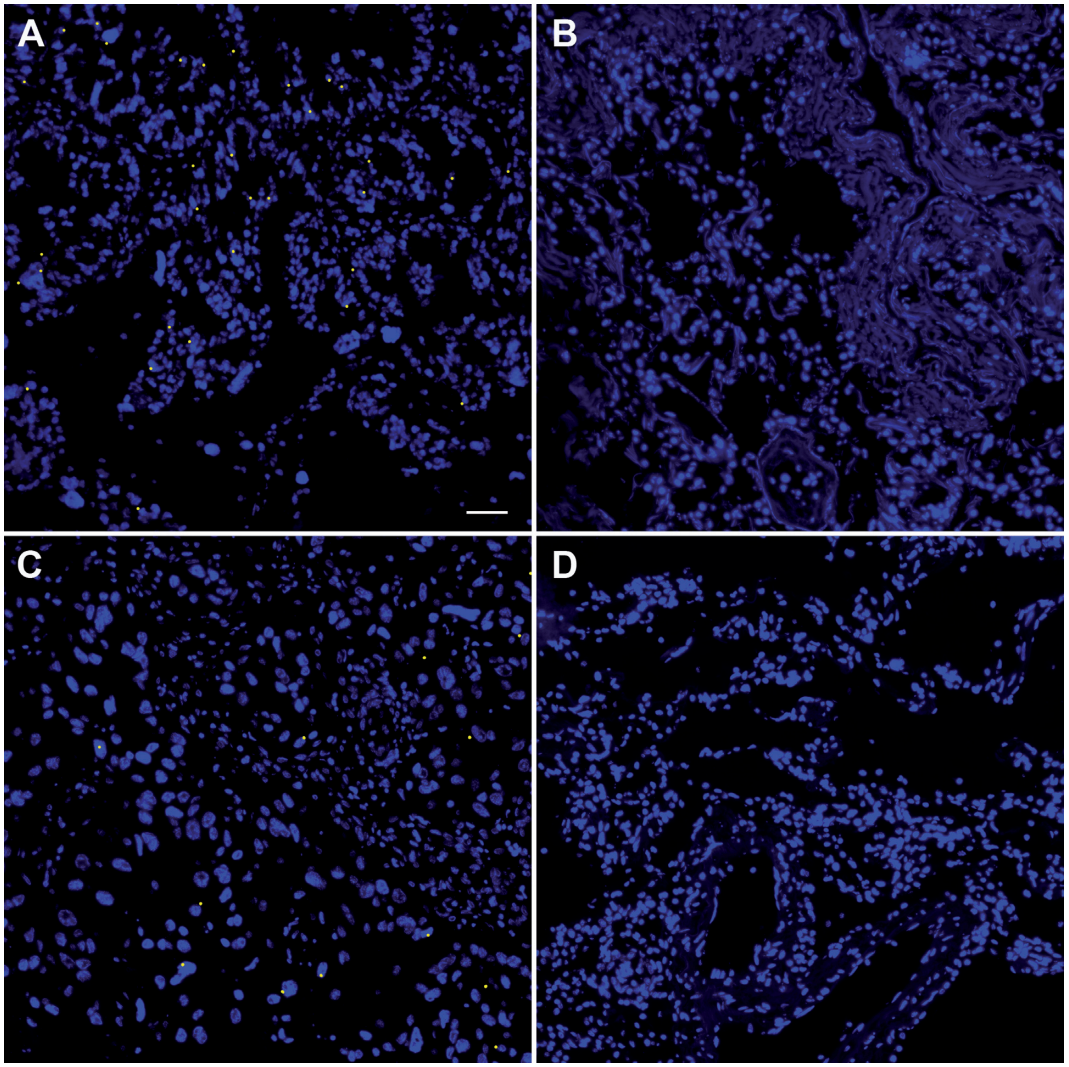

Figure 6. *In situ* mutation detection in fresh frozen and FFPE lung tissue sections. Mutation detection in fresh frozen lung tissues was performed with two gap probes, only one being 5′-phosphorylated, targeting the L858R mutation in the *EGFR* gene and the corresponding wild-type sequence. Representative images of a fresh frozen, L858R positive, tumor sample (**A**) and a normal tissue (**B**) tested with the mutant hexamer (2573 T>G) 5′-phosphorylated. FFPE tissue sections were tested with a pool of six gap probes coding for reported mutations in codon 858 and 859 in exon 21 in *EGFR* and the corresponding wild-type sequence. Representative images of FFPE, L858R positive (**C**) and *EGFR* wild-type (**D**) tumor samples tissue tested with the pool of mutant hexamers 5′-phosphorylated. Cell nuclei are shown in blue and the detected RCPs in yellow. Scale bar 50 μm.

HCT 116. This cell line has been previously characterized by RNA sequencing analysis and low RPKM values for the *KRAS* gene (3.8) compared to *ACTB* and *EGFR* (756 and 14, respectively) have been reported in the GEO database (accession GSE34791) (25,26). For *KRAS*, we were able to detect one mutant RCP in about 10 cells and thus, when detecting low abundant variants more cells have to be counted in order to detect the mutant target. However, for clinical diagnostics applications, high resolution, e.g. signal detected in every cell, may not be necessary although desirable.

Most of the previous studies conducted to assess the mechanism and specificity of the ligation process of mesophilic and thermostable ligases were conducted on systems where two oligonucleotides were joined on a DNA template requiring one single ligation event. Although not the aim of the presented work, an interesting observation emerges from our experiments. It has been previously reported that the thermostable enzyme *Tth* ligase is unable to efficiently join oligonucleotides of equal or shorter length than six bases at the 5′-end of the ligation site, requiring contacts up to the 9th base pair 5′ upstream of the ligation point (23). Nevertheless, we show in this work that short hexamer probes can be efficiently ligated. The property of a thermostable ligase to direct ligation in the 3′ to 5′-direction has been demonstrated before (27), and accordingly we hypothesize that a first ligation event takes place at the 5′-end of the padlock gap probe, followed by the sealing of the nick at the 3′-end of the padlock gap probe. Regardless of whether such an order of ligation exists, our results show that Ampligase can tolerate a nick in the footprint region of the DNA sequence, albeit at 20% lower efficiency of the assay compared to a single ligation event. It is also worth mentioning that hexamer oligonucleotides are the shortest gap probes that we were able to generate signals with using

the protocol presented herein (data not shown). The maximum length of the gap probe that can be used in the assay is limited by the ability of the DNA ligase to tolerate mismatches far from the ligation sites and remains to be tested.

The padlock gap-fill ligation assay may prove an advantageous strategy for the simultaneous assessment of mutational status at many loci of known sequence. It has been shown previously that multiple padlock probes can be used to assess the presence of different mutations in the same sample (28). However, the number of distinguishable padlock probes is limited by the number of fluorophores that can be spectrally separated in the sample. The padlock gap probe approach offers a simple design to be able to increase the number of investigated loci. Only one padlock gap probe is required for each targeted locus, offering the possibility to screen for many mutational variants at each locus in each reaction by addition of short, inexpensive sequence-specific gap probes.

In conclusion, we have developed a targeted method for mutation detection based on padlock gap probes and *in situ* RCA and showed the high specificity of the method by successfully screening for the MELAS mutation A3243G in the mitochondrial DNA of two cell lines with known mutational status. Furthermore, we genotyped a single nucleotide variant in *ACTB* mRNA transcripts performing *in situ* sequencing of the targeted SNV. To demonstrate the potential of this assay for use in clinical diagnostics, we targeted a point mutation in *EGFR*, indicative of treatment response, in tissue samples and applied *in situ* sequencing to confirm the detected mutation. We propose that *in situ* padlock gap-fill ligation provides an inexpensive assay for screening for the presence of multiple mutations in mutational hotspots *in situ,* and the possibility to combine the assay with *in situ* sequencing provides a new method to capture and sequence genetic material directly in cells and tissues.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Swedish research council; VINNOVA; the European Community's 7th Framework Program (FP7/2007–2013) DiaTools [259796]; CareMore project FP7-HEALTH-2013-INNOVATION-2 [601760]. Funding for open access charge: CareMore project.
Conflict of interest statement. None declared.

REFERENCES

1. Landegren,U., Kaiser,R., Sanders,J. and Hood,L. (1988) A ligase-mediated gene detection technique. *Science*, **241**, 1077–1080.
2. Jackson,D.A., Symons,R.H. and Berg,P. (1972) Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 2904–2909.
3. Shendure,J., Porreca,G.J., Reppas,N.B., Lin,X., McCutcheon,J.P., Rosenbaum,A.M., Wang,M.D., Zhang,K., Mitra,R.D. and Church,G.M. (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, **309**, 1728–1732.
4. Conze,T., Shetye,A., Tanaka,Y., Gu,J., Larsson,C., Goransson,J., Tavoosidana,G., Soderberg,O., Nilsson,M. and Landegren,U. (2009) Analysis of genes, transcripts, and proteins via DNA ligation. *Annu. Rev. Anal. Chem.*, **2**, 215–239.
5. Weiss,B. and Richardson,C.C. (1967) Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. *Proc. Natl. Acad. Sci. U.S.A.*, **57**, 1021–1028.
6. Larsson,C., Koch,J., Nygren,A., Janssen,G., Raap,A.K., Landegren,U. and Nilsson,M. (2004) In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nat. Methods*, **1**, 227–232.
7. Andersson,C., Henriksson,S., Magnusson,K.E., Nilsson,M. and Mirazimi,A. (2012) In situ rolling circle amplification detection of Crimean Congo hemorrhagic fever virus (CCHFV) complementary and viral RNA. *Virology*, **426**, 87–92.
8. Shaposhnikov,S., Larsson,C., Henriksson,S., Collins,A. and Nilsson,M. (2006) Detection of Alu sequences and mtDNA in comets using padlock probes. *Mutagenesis*, **21**, 243–247.
9. Larsson,C., Grundberg,I., Soderberg,O. and Nilsson,M. (2010) In situ detection and genotyping of individual mRNA molecules. *Nat. Methods*, **7**, 395–397.
10. Hardenbol,P., Baner,J., Jain,M., Nilsson,M., Namsaraev,E.A., Karlin-Neumann,G.A., Fakhrai-Rad,H., Ronaghi,M., Willis,T.D., Landegren,U. *et al.* (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.*, **21**, 673–678.
11. Porreca,G.J., Zhang,K., Li,J.B., Xie,B., Austin,D., Vassallo,S.L., LeProust,E.M., Peck,B.J., Emig,C.J., Dahl,F. *et al.* (2007) Multiplex amplification of large sets of human exons. *Nat. Methods*, **4**, 931–936.
12. Lizardi,P.M., Huang,X., Zhu,Z., Bray-Ward,P., Thomas,D.C. and Ward,D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat. Genet.*, **19**, 225–232.
13. Ke,R., Mignardi,M., Pacureanu,A., Svedlund,J., Botling,J., Wahlby,C. and Nilsson,M. (2013) In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods*, **10**, 857–860.
14. Goto,Y., Nonaka,I. and Horai,S. (1990) A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature*, **348**, 651–653.
15. Botling,J., Edlund,K., Lohr,M., Hellwig,B., Holmberg,L., Lambe,M., Berglund,A., Ekman,S., Bergqvist,M., Ponten,F. *et al.* (2013) Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis, and tissue microarray validation. *Clin. Cancer Res.*, **19**, 194–204.
16. Carpenter,A.E., Jones,T.R., Lamprecht,M.R., Clarke,C., Kang,I.H., Friman,O., Guertin,D.A., Chang,J.H., Lindquist,R.A., Moffat,J. *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.***7**, R100.
17. Szollosi,J., Lockett,S.J., Balazs,M. and Waldman,F.M. (1995) Autofluorescence correction for fluorescence in situ hybridization. *Cytometry*, **20**, 356–361.
18. Luo,J., Bergstrom,D.E. and Barany,F. (1996) Improving the fidelity of Thermus thermophilus DNA ligase. *Nucleic Acids Res.*, **24**, 3071–3078.
19. van den Ouweland,J.M., Maechler,P., Wollheim,C.B., Attardi,G. and Maassen,J.A. (1999) Functional and morphological abnormalities of mitochondria harbouring the tRNA(Leu)(UUR) mutation in mitochondrial DNA derived from patients with maternally inherited diabetes and deafness (MIDD) and progressive kidney disease. *Diabetologia*, **42**, 485–492.
20. Forbes,S.A., Beare,D., Gunasekaran,P., Leung,K., Bindal,N., Boutselakis,H., Ding,M., Bamford,S., Cole,C., Ward,S. *et al.* (2015) COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.*, **43**, D805–D811.
21. Sharma,S.V., Bell,D.W., Settleman,J. and Haber,D.A. (2007) Epidermal growth factor receptor mutations in lung cancer. *Nat. Rev. Cancer*, **7**, 169–181.
22. Hirsch,F.R., Janne,P.A., Eberhardt,W.E., Cappuzzo,F., Thatcher,N., Pirker,R., Choy,H., Kim,E.S., Paz-Ares,L., Gandara,D.R. *et al.* (2013) Epidermal growth factor receptor inhibition in lung cancer: status 2012. *J. Thorac. Oncol.*, **8**, 373–384.

23. Pritchard,C.E. and Southern,E.M. (1997) Effects of base mismatches on joining of short oligodeoxynucleotides by DNA ligases. *Nucleic Acids Res.*, **25**, 3403–3407.
24. Khanna,M., Cao,W., Zirvi,M., Paty,P. and Barany,F. (1999) Ligase detection reaction for identification of low abundance mutations. *Clin. Biochem.*, **32**, 287–290.
25. Barrett,T., Wilhite,S.E., Ledoux,P., Evangelista,C., Kim,I.F., Tomashevsky,M., Marshall,K.A., Phillippy,K.H., Sherman,P.M., Holko,M. *et al.* (2013) NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.*, **41**, D991–D995.
26. Wang,H., Mayhew,D., Chen,X., Johnston,M. and Mitra,R.D. (2012) 'Calling cards' for DNA-binding proteins in mammalian cells. *Genetics*, **190**, 941–949.
27. Housby,J.N. and Southern,E.M. (1998) Fidelity of DNA ligation: a novel experimental approach based on the polymerisation of libraries of oligonucleotides. *Nucleic Acids Res.*, **26**, 4259–4266.
28. Grundberg,I., Kiflemariam,S., Mignardi,M., Imgenberg-Kreuz,J., Edlund,K., Micke,P., Sundstrom,M., Sjoblom,T., Botling,J. and Nilsson,M. (2013) In situ mutation detection and visualization of intratumor heterogeneity for cancer research and diagnostics. *Oncotarget*, **4**, 2407–2418.