Visualising individual sequence-specific protein–DNA interactions in situ

Irene Weibrecht¹, Milan Gavrilovic², Lena Lindbom¹, Ulf Landegren¹, Carolina Wähby³,⁴ and Ola Söderberg¹

¹Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Rudbeck Laboratory, Uppsala University, SE-75185 Uppsala, Sweden
²Department of Information Technology, Centre for Image Analysis, Uppsala University, SE-75105 Uppsala, Sweden
³Imaging Platform, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
⁴Department of Information Technology, Centre for Image Analysis, Science for Life Laboratory, Uppsala University, SE-75105 Uppsala, Sweden

Gene expression – a key feature for modulating cell fate – is regulated in part by histone modifications, which modulate accessibility of the chromatin to transcription factors. Until now, protein–DNA interactions (PDIs) have mostly been studied in bulk without retrieving spatial information from the sample or with poor sequence resolution. New tools are needed to reveal proteins interacting with specific DNA sequences in situ for further understanding of the orchestration of transcriptional control within the nucleus. We present herein an approach to visualise individual PDIs within cells, based on the in situ proximity ligation assay (PLA). This assay, previously used for the detection of protein–protein interactions in situ, was adapted for analysis of target PDIs, using padlock probes to identify unique DNA sequences in complex genomes. As a proof-of-principle we detected histone H3 interacting with a 26 bp consensus sequence of the Alu-repeat abundantly expressed in the human genome, but absent in mice. However, the mouse genome contains a highly similar sequence, providing a model system to analyse the selectivity of the developed methods. Although efficiency of detection currently is limiting, we conclude that in situ PLA can be used to achieve a highly selective analysis of PDIs in single cells.

Introduction

During recent years, it has become evident that epigenetic modifications of histone proteins bound to promoter sequences of genes contribute to the regulation of gene expression and cellular differentiation. Histone modifications also play important roles in establishing and maintaining the malignant phenotype of tumour cells [1]. Such epigenetic modifications or changes in transcription factor activity are even hypothesised to be among the initial events transforming a cell into a so-called cancer stem cell (CSC). By studying protein–DNA interactions (PDIs) such as epigenetic changes at the genomic DNA level during cancer development or activation of specific transcription factors when cancer cells undergo epithelial–mesenchymal-transition, markers for CSC and metastatic tumour colonies could be identified. Because such cells will only present a fraction of all cells comprising the tumour from a patient biopsy, they cannot be identified by bulk analyses averaging over the whole population of cells. However, PDIs have long been studied in bulk populations by methods such as electrophoretic mobility shift assays [2] and chromatin immunoprecipitation [3], providing information about proteins that bind certain DNA sequences, averaged over large numbers of cells. Therefore, new methods are needed to enable investigation of PDIs at cellular and subcellular resolution in cells and tissues. Direct investigation of PDIs in the context of the tissue, rather than monitoring transient and stochastic downstream effects like mRNA expression, would give a more coherent insight into the regulatory mechanisms that drive diseases such as cancer and could ultimately lead to new approaches for therapy.
Fluorescence resonance energy transfer has been used to investigate proximity between proteins and genomic DNA [4], with a green fluorescent protein-fusion of the heterochromatin-associated proteins HP1α and HP1β as donor and an intercalating fluorescent dye as acceptor, to demonstrate the binding of the protein to genomic DNA. While this method provides single cell analysis, it is limited by the need for transfection of genetic constructs and thus cannot be used to study endogenous protein complexes. More importantly, no information can be obtained about the bound sequence. Co-localisation of a transcription factor detected by immunofluorescence and DNA fluorescence in situ hybridisation (FISH) has also been studied [5] but the approach is limited by the low spatial resolution of confocal microscopy. In addition, although the assay offers more specific detection of DNA sequences, FISH is typically incapable of distinguishing between closely similar sequence variants.

The in situ proximity ligation assay (PLA) serves to reveal endogenous proteins and protein–protein interactions in fixed cells and tissue sections, using antibodies to target individual proteins. In this assay, dual antibody binding results in the formation of an amplifiable circular DNA molecule. After rolling circle amplification (RCA), this DNA molecule is visualised by hybridisation of fluorescence-labelled detection probes [6–7]. The potent signal amplification by RCA enables detection of single molecules, which become labelled with hundreds of fluorophores. Here we have adapted in situ PLA for analyses of PDS in situ by developing a protocol and evaluating its efficiency and selectivity by investigating the proximity of histone H3 to genomic Alu-repeats as a model system. The advantage of this model is that histone H3 is present in large numbers and thus often present in conjunction with Alu-repeats, an abundant sequence in the genome. Alu-repeats are dispersed over the whole genome, although density varies between chromosomes, preferentially present in GC-rich regions [8]. The 26 bp Alu-consensus sequence [9] that we chose as a target sequence is present in humans, but no exact copies are present in mice. We compared a design using hybridisation to a more stringent one based on ligation of padlock probes for DNA target recognition [10–11]. The work presented here provides a basis for a novel approach to study epigenetic alterations in situ.

Methods

Cell culture

Human TERT immortalised fibroblasts (BJTert) and mouse fibroblasts (NIH3T3) were seeded on 8-well collagen-coated chamber slides (BD Bioscience), 33,000 cells per well in modified Eagle’s medium (Gibco, for BJTert cells) or Dulbecco’s modified Eagle’s medium (Sigma, for NIH3T3 cells) + 10% foetal calf serum (heat inactivated; Sigma) and incubated overnight at 37°C.

Fixation and permeabilisation

Cells were put on ice for 10 min and subsequently washed twice with ice cold phosphate buffered saline (PBS), 5 min each (900 μl per well). The cells were fixed with 2% (w/v) p-formaldehyde (PFA; Sigma) in PBS for 30 min on ice. After a short wash with 900 μl ice cold PBS per well, the cells were permeabilised with ice cold 70% ethanol for 30 min on ice. Next, the silicon masks of the chamber slides were removed and the slides were dried at room temperature. Subsequently, the slides were rehydrated in PBS for 10 min and cells were permeabilised in 2.9 μM pepsin (Sigma, freshly diluted from 580 μM stock) in 37°C warm 0.1 M HCl (Sigma) for 65 s at 37°C. Before and subsequently to a wash in 1 x NaCl, 0.1 x Tris–HCl (pH 7.5) for 60 s at 37°C, the slides were washed twice for approximately 1 min in PBS and finally dried by an ethanol series (70%/85%/99.6%, 2 min each) and a short spin in a table-top centrifuge. A hydrophobic barrier pen was applied to the borders of the wells and 8-chamber secure seals (9 mm in diameter, 0.8 mm deep; Grace Bio-Labs) were attached to the slides. The cells were rehydrated in washing buffer (100 x Tris–HCl (pH 7.5), 150 x NaCl, 0.05% Tween 20 (Sigma)) for 10 min at 37°C. Thereafter, all incubations were done in a moisture chamber, and the wells were sealed with a PCR tape when incubation time exceeded 30 min. Genomic DNA was digested with 0.5 U/μl Alul (New England Biolabs (NEB)) in 1 x NEB-buffer-4 and 0.2 μg/μl BSA (NEB) for 10 min at 37°C. All subsequent washes were performed by removing the incubation solution from the wells first, and then flushing the wells with approximately 500 μl of the appropriate washing buffer. After the Alul-treatment, the cells were washed in washing buffer before the DNA was made partially single stranded by treatment with 0.2 μg/μl lambda-exonuclease (NEB) in 1 x lambda-exonuclease buffer (NEB), 0.2 μg/μl BSA and 10% glycerol (Sigma) for 30 min at 37°C. The cells were washed in washing buffer before and subsequently to a post-fixation step with 1% (w/v) PFA in PBS for 5 min on ice.

Immunofluorescence

Cells were treated as described above. The cells were blocked in Starting Block T20 PBS (Pierce) containing 2.5 mM L-cysteine (Sigma) and 2.5 ng/μl sonicated salmon sperm DNA (Invitrogen) for 1 h at 37°C. Afterwards, 2.5 ng/μl of the primary rabbit-anti-histone H3 antibody (#1791, Abcam) was applied in blocking buffer at 4°C overnight. The wells were washed in tris buffered saline (TBS) (10 x Tris–HCl, 150 mm NaCl, pH 7.7) with 0.05% Tween 20 (TBS + T, 1000 μl per well) before 7.5 ng/μl donkey-anti-rabbit-FITC (Jackson ImmunoResearch) was added in 2 x saline sodium citrate buffer (SSC), 0.25 μg/μl BSA, 7.5 ng/μl poly(A) (Sigma), 0.05% Tween 20 and incubated for 30 min at 37°C. Finally, the slides were rinsed once in TBS, secure seals were removed and the slides were washed 2 x for 10 min in TBS before they were spun dry and mounted in Vectashield mounting medium (Vector) containing 100 ng/ml 4',6-diamidino-2-phenylindole (DAPI).

Conjugation of anti-rabbit PLA probe

One hundred μg donkey-anti-rabbit (Jackson ImmunoResearch) was dialysed against PBS in a dialysis cup (Slide-A-lyzer Mini dialysis units 7000 MWCO; Pierce) overnight at 4°C. The antibody was subsequently concentrated to approximately 20 μl by centrifugation in Amicon Ultra 0.5 ml 10K Ultracell 10K Membrane (Millipore), prewashed in PBS. Then sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC, Pierce) freshly dissolved in dimethyl sulfoxide (DMSO, Sigma) was added in ~25 x molar access over antibody for 2 h at room temperature. The PLA oligonucleotide (5’-thiol-GTC TTA ACT ATT AGC GAT ACG GTCTC CGA ATC GCT GAC AGA ACT AGA CAC-3’, HPLC pure; Biomers) was degassed for 5 min at 80°C, chilled and reduced by incubation with 50 μl dithiothreitol (DTT, (50 μM final concentration of the PLA oligonucleotide) for 1 h at 37°C. Afterwards the volume was filled up to approximately 50 μl through the
FIG. 1
Hybridisation-based in situ PLA. (A) The DNA-binding PLA probe is hybridised to single-stranded genomic DNA. (Ai) Histone H3 is detected by a primary antibody and a secondary PLA probe. (Aii) Two circularisation oligonucleotides (one containing a hybridisation site for fluorescence detection, red) are hybridised to both PLA probes, which are (Aiv) subsequently ligated forming a circular DNA molecule. (Av) This circular DNA molecule is amplified by RCA and detected through hybridisation with fluorescence-labelled oligonucleotides. (Bi) Immunofluorescence detection of histone H3 in human (BJhTert) and mouse (NIH3T3) fibroblasts (red). Nuclei are depicted in white and scale bars represent 10 μm. (Bii) RCA-mediated detection of hybridised DNA-binding PLA probe in human and mouse cells (red). Nuclei are depicted in white and scale bars represent 10 μm. (C) Detection of individual histone H3 proteins in proximity to an Alu-repeat (RCA products seen as red spots). All images shown here are enhanced for better photographic representation, inserts show a higher magnification of the same field of view, nuclei are depicted in blue, scale bars represent 10 μm. (D) Quantification of histone H3-Alu-repeat interactions in human and mouse cells. Results are shown from automated quantification of RCA products in 33 (human) and 36 (mouse) cells from one experiment out of three replicates. Each dot represents the number of detected interactions in an individual cell. Median (blue line), 25th and 75th percentile (grey box) are shown.
**FIG. 2**

Genomic DNA-templated in situ PLA. (Ai) Histone H3 is detected by a primary antibody and a secondary PLA probe. (Aii) Two circularisation oligonucleotides together covering the whole target sequence (one containing a hybridisation site for fluorescence detection, red) are hybridised to the genomic DNA forming a single G/A mismatch with the genomic DNA. (Aiii) The two circularisation oligonucleotides are subsequently ligated. The ligation step increases the detection.
addition of buffer A (55 mM phosphate buffer, 150 mM NaCl, 20 mM EDTA, pH 7.0). Buffer A was also used to prewash G-50 columns (GE Healthcare). Both antibody and oligonucleotide were then purified over three such columns to remove activating reagents. Immediately afterwards, antibody and oligonucleotide were mixed and dialysed against PBS overnight at 4°C.

Hybridisation-based in situ PLA

Cells fixed and permeabilised as described above were incubated with 200 nM Hyb-oligonucleotide (Fig. 1a) (5'-GCC TCC CAA AGT GCT GGG ATT ACA GGA AAA AAC ATG GAT GTT CTT GAC ATG GCA ATG ACG CTA A-3', PAGE pure; Integrated DNA Technologies (IDT), target complementary parts are shown in italic, mismatches to the mouse sequence are shown in bold) in 1× T4 DNA-ligation buffer (10 mM Tris-acetate, pH 7.5, 10 mM magnesium acetate, 50 mM potassium acetate), 250 mM NaCl, 0.25 μg/μl BSA and 0.05% Tween 20 for 30 min at 37°C. Slides were then washed with washing buffer and subsequently in 2× SSC + 0.1% Tween 20 for 5 min at 37°C. Afterwards the buffer was changed to TBS + T. The cells were blocked in Blocking Block T20 PBS containing 2.5 mg/mL-cysteine and 2.5 ng/μl sonicated salmon sperm DNA for 1 h at 37°C. Next, 2.5 ng/μl of the primary rabbit-anti-histone H3 antibody or rabbit IgG (Jackson Immunoresearch) were applied in blocking buffer at 4°C overnight. The wells were washed in TBS + T (1000 μl per well) and 2.5 ng/μl secondary anti-rabbit-PLA probe was preincubated in blocking buffer for 30 min at room temperature and then applied to the slides for 1 h at 37°C. After a wash in 10 mM Tris–HCl (pH 7.5), 0.1% Tween 20 for 5 min at 37°C and a short rinse in TBS + T, two connector oligonucleotides (5’-phosphate–AGC GAT CTG CGA GAC AGT GAA TGC GAC TGC GTC TAA GAG AGT ACA GCA GCC GTC TTA GGC TGC TCA TGG CCA T-3’, PAGE pure, IDT; 5’-phosphate–CTG AAG ATC CAT GAA AGT GTC GTC TAG TTC TGT C-3’, PAGE pure, IDT) were applied at a final concentration of 125 mM for 30 min at 37°C in 1× T4 DNA-ligation buffer, 0.05 U/μl T4 DNA ligase (Fermentas), 1 mM ATP (Fermentas), 0.25 μg/μl BSA, 250 mM NaCl and 0.05% Tween 20. The wells were washed with TBS + T before RCA with 1 U/μl phi29 DNA polymerase (Fermentas), 1× phi29 buffer (Fermentas), 0.25 mM dNTP (Fermentas), 0.2 μg/μl BSA, 5% glycerol and 1:1000 mouse-anti-histone H3 (phospho-S10) (Abcam) for 1.5 h at 37°C. Detection was performed after a wash with TBS + T with 250 nm detection oligonucleotide (5’-Alexa Fluor 555–CAG TGA ATG CGA GTC CTG CT-3’, HPLC pure; TriLink) in 2× SSC, 0.25 μg/μl BSA, 7.5 ng/μl poly(A), 0.05% Tween 20 and 6.5 ng/μl donkey-anti-mouse-FITC F(ab')2-fragment (Jackson Immunoresearch) for 30 min at 37°C. Finally, the slides were rinsed once in TBS, secure seals were removed and the slides were washed 2× 10 min in TBS before they were spun dry and mounted in Vectashield mounting medium containing 100 ng/ml DAPI.

Detection of hybridisation oligonucleotides by padlock probes

Cells were treated as described for Hybridisation-based in situ PLA until the slides had been blocked. Then, 125 nm padlock oligonucleotide (5’-phosphate–TGT CAA GAA CAT CCA TGT CAG TGA ATG CGA GTC CTG CTA AGA GAG TAG TAC AGC AGC GTT TTA GGC TCA TGG CC-3’, HPLC pure, IDT) – a padlock probe directed against the free part of the Hyb-oligonucleotide – was applied in 1× T4 DNA-ligation buffer, 0.25 μg/μl BSA, 0.05 U/μl T4 DNA ligase, 0.05% Tween 20, 1 mM ATP and 250 mM NaCl and incubated for 30 min at 37°C. RCA, detection and mounting were performed as described above.

Genomic DNA-templated in situ PLA

Cells were fixed and permeabilised before the slides were blocked and incubated with primary antibody and secondary PLA probes (Fig. 2a), as described for Hybridisation-based in situ PLA After washing, 200 nM GenomeTemplate-oligonucleotide 1 (5’-phosphate–CT G GGA TTA CAG GAA AAA AAG AGT AGT TAG TTC TGT C-3’, PAGE pure, IDT, target complementary parts are shown in italic, mismatches to the mouse sequence are shown in bold) and 200 nM Genome-Template-oligonucleotide 2 (5’-phosphate-CGC TAA TAG TTA AGA AGC TCA TGT GAT AAC GGT AGT CCG TCT AAA AAA AGC CGC CCA AAG TG-3’, PAGE pure, IDT, target complementary parts are shown in italic, mismatches to the mouse sequence are shown in bold, the G/A mismatch utilised for MutY-cleavage is underlined) were applied in 1× T4 DNA-ligation buffer, 0.25 μg/μl BSA, 0.05% Tween 20 and 250 mM NaCl and incubated for 30 min at 37°C. After a washing step with washing buffer, 0.05 U/μl T4 DNA ligase in T4 DNA-ligation buffer, 1 mM ATP, 0.25 μg/μl BSA, 250 mM NaCl and 0.05% Tween 20 was applied for 30 min at 37°C. The slide was washed in 2× SSC + 0.05% Tween 20 for 5 min at 37°C and subsequently in TBS + T. Next, the probes were digested with MutY/endonuclease IV (EndoIV) using 0.64 μM MutY protein (USB), 0.1 U/μl EndoIV (Fermentas) in MutY buffer (20 mM Tris–HCl, pH 7.6, 30 mM NaCl, 1 mM EDTA, 100 mM KCl, 1 mM DTT) and 0.5 μg/μl BSA for 30 min at 37°C. After an additional washing step in TBS + T 125 nm of the additional sequence oligonucleotide (5’-phosphate–AGC GAT CTG CGA GAC GTG AT-3’, HPLC pure; Biomers) in 0.05 U/μl T4 DNA ligase, 1× T4-ligation buffer, 1 mM ATP, 0.25 μg/μl BSA, 0.05% Tween 20 and 250 mM NaCl were incubated for 30 min at 37°C. RCA, detection and mounting were performed as described above, except that 250 nm of an additional detection oligonucleotide (5’-Cys-5’-AGC GAT CTG CGA GAC GTG AT-3’, HPLC pure, IDT) were added to the mix.

Padlock probe-based in situ PLA

The padlock probe was made by ligation of two parts: A (5’-phosphate–CT G GGA TTA CAG GAA AAA AAG AGT AGT TAG TTC TGT C-3’, HPLC pure, IDT) and B (5’-TCA TGG CC-3’, HPLC pure, IDT). The A part was extended with dATP and dTTP in the presence of T4 DNA ligase in the presence of B part. The product was then treated with alkaline phosphatase and the Dig-labelled product was used as probe.
FIG. 3
Padlock probe-based in situ PLA. (Ai) Alu-repeats are detected by hybridisation of the hairpin-containing padlock probe (also containing a hybridisation site for fluorescence detection, red) and (Aii) ligated by Amp-ligase. (Aiii) Histone H3 is detected by the primary antibody and the secondary PLA probe. (Aiv) The hairpin structure is opened by UNG/EndoIV treatment to free the complementary ends (blue). In the same step MutY/EndoIV enzymes are utilised to create a free 3’ end of genomic DNA by cleaving the G/A mismatch formed by the padlock probe. (Av) The protein-bound PLA probe hybridises to the opened padlock probe incorporating an additional splint oligonucleotide (green). (Avi) The former padlock probe is religated to form a circular DNA molecule. (Avii) This is amplified by RCA and detected through hybridisation of fluorescence-labelled oligonucleotides to the two detection sites (red, green) resulting in two-coloured signals appearing yellow.
AAA AAG CGT CUT AA-3’ – PAGE pure, IDT, target complementary parts are shown in italic, mismatches to the mouse sequence are shown in bold) and B (5'-phosphate–CTA UTA GCG ACA AAA AAG UCG CTA ATA GTT AAG ACG CTC AGT GAA TGC GAG TCC GTC TAA AAA AA G CCG CCC AAA GTG-3’ – PAGE pure, IDT, target complementary parts are shown in italic, mismatches to the mouse sequence are shown in bold, the G/A mismatch utilised for MutY-cleavage is shown with an underline). The two oligonucleotides were mixed at a final concentration of 4.4 μM, 1 x T4 DNA-ligation buffer and 2 mM ATP were added. The oligonucleotide mix was first incubated at 90°C, then 65°C for 10 min each. After cooling to room temperature, 0.025 U/μl T4 DNA ligase was added. The sample was ligated for 30 min at 37°C and then incubated for 10 min at 65°C to inactivate the ligase.

Cells were fixed and permeabilised as described above, before incubation with 200 nm padlock probe (Fig. 3a) in 1 x T4 DNA-ligation buffer, 250 mM NaCl, 0.25 μg/μl BSA and 0.05% Tween 20 for 30 min at 37°C. Afterwards the slides were washed in washing buffer and the probe ligated with 0.25 U/μl Amp-ligase ( Epicentre) in 1 x Amp-ligase buffer (20 mM Tris–HCl (pH 8.3), 75 mM KCl, 10 mM MgCl2, 0.5 mM NAD (Sigma), 0.01% Triton X-100 (PlusOne)), 0.25 μg/μl BSA and 5% glycerol for 30 min at 45°C. Slides were then washed and treated with primary antibody and secondary PLA probe as described above. Next, probes were digested with MutY/EndoIV and uracil–DNA glycosylase (UNG) for PDI detection with 0.64 μM MutY glycosylase, 0.1 U/μl EndoIV, 0.05 U/μl UNG (Fermentas) in MutY buffer and 0.5 μg/μl BSA for 30 min at 37°C. For samples where the padlock probe should be applied as DNA-detection probe only, UNG was omitted from this step. All subsequent steps were performed as described above for genomic DNA-templated in situ PLA.

Image acquisition and preprocessing
Images were acquired using a Zeiss Axioplan 2 epifluorescence microscope, the AxioCam MRm CCD sensor and a 40 x/1.3 Oil PlanNeofluar objective together with filters for DAPI, FITC, Cy3 and Cy5. Imaging positions were chosen using the FITC channel, avoiding cells showing staining for phospho-histone H3 S10, a mitotic marker. For each condition, z-stacks of seven images, 0.275 μm apart, were acquired at four positions of the well. As cells in interphase are comparably flat, z-stacks were collapsed by taking maximum intensity projections of individual z-slices, while the purpose of imaging more than one focal plane was to avoid longitudinal chromatic aberrations. Lateral chromatic aberrations were corrected using a rigid geometrical transformation. Background illumination was reduced by subtracting the median intensity from Cy3 and Cy5 channels, respectively.

Image analysis
To quantify PDIs on a per cell basis, we identified individual cells as well as point-source signals. Cell nuclei imaged in DAPI were segmented from the image background by intensity thresholding and touching nuclei masks were separated based on shape [12]. Individual signals were thereafter detected in the Cy3 (red) and Cy5 (green) channels separately using a point-source signal detection method [13]. The method consists of two parts, a detector, which is a cosine filter to enhance the signals, and a verifier, which is a sine filter to validate the result from the detector [13]. The results of signal detection in the Cy3 and Cy5 channels were combined, and the ratio of fluorescence intensity from the two colour dyes was extracted from each detected signal. Signals were thereafter classified as red, green or two-coloured based on the ratio distribution of green to red intensity [14–15]. When compared with classification methods based on intensity thresholding [16–17], this type of ratio-based classification ensures that strong and weak signals are assigned to the same class. For hybridisation-based in situ PLA single coloured Cy3 signals were regarded as detected PDIs, while for genomic DNA-templated in situ PLA and for padlock probe-based in situ PLA only two-coloured (Cy3 and Cy5) signals were regarded as true PDIs. For each of the experiments performed under new conditions, image data were divided into a training and a test set. The training sets were used for fine tuning parameters related to signal size and intensity in the image analysis pipeline, followed by application of the algorithms to the larger test sets to present results per cell population.

Results
We developed means to use proximity ligation to detect the colocalisation of a DNA-binding protein and a specific short genomic DNA sequence in individual fixed cells, and evaluated different approaches (Figs 1a, 2a and 3a) with regard to their selectivity.

The cells to be investigated for specific PDIs were PFA-fixed and permeabilised with ethanol, followed by a pepsin treatment and a wash in high salt to render the nuclei more accessible for probing. Then genomic DNA was digested by Alu restriction enzyme treatment. As genomic DNA cannot be heat denatured for sequence detection without compromising the protein of interest, the target sequence had to be made accessible for hybridisation by treatment with lambda-exonuclease, to use its S–3’ activity to generate single stranded overhangs with free 3’ ends at the sites of digestion. As proof-of-principle we chose to investigate the proximity between histone H3 and a 26 bp Alu-consensus sequence (5’-CTCGTAATCCGGACGACTTTTGAGGGAAGC-3’) [9], present in approximately 60,000 copies per cell in humans as determined by searching perfect sequence matches in the human genome sequence (NCBI 36, March 2006 assembly). Although the exact sequence is not present in mice, their genomes contain a sequence similar to the human 26 bp Alu-consensus sequence (5’-CCCTTTAATCCGGACGACTTTTGAGGGAAGC-3’, differences to the human sequence are indicated in bold) [18] in approximately 6000 copies (perfect sequence matches per mouse genome, NCBI 36, October 2006 assembly), which is not present in humans. This provided a good...
model for studying the selectivity of probing. We confirmed that the antibody used to target the histone H3 stains human and mouse nuclei approximately equally using immunofluorescence on human (BJhTert) and mouse (NIH3T3) fibroblasts (Fig. 1bi).

Hybridisation-based in situ PLA

In situ PLA has previously been used for the detection of individual protein–protein interactions [6]. To extend the application to specific PDIs, we first investigated a simple hybridisation approach to detect the genomic DNA sequences. In this design, we used an anti-rabbit immunoglobulin (Ig) antibody conjugated with an oligonucleotide as a PLA probe to detect a primary rabbit antibody directed against histone H3. The other PLA probe was an oligonucleotide, complementary to the target sequence and extended with a sequence required for the PLA reaction (Fig. 1a). To investigate the selectivity of the DNA-directed PLA probe we used it for hybridisation to human cells carrying the target sequence, and to mouse cells that carry a similar but non-identical sequence in approximately tenfold fewer copies. The bound probes were then visualised by RCA, using a padlock probe directed to the free 3′ end of the hybridisation probe. Padlock probes are single-stranded oligonucleotides that hybridise to their target sequences with their 5′ and 3′ end facing each other. Upon perfect hybridisation the ends can be joined by ligation, creating a circular DNA molecule, which may then be locally amplified by phi29 DNA polymerase through RCA [10,19]. The resulting single-stranded DNA molecule, consisting of approximately 1000 complementary repeats of the original circle, forms a sub-μm sized bundle of DNA. We detected these RCA products by hybridisation of fluorescence-labelled oligonucleotides. As the fluorophores are concentrated in a very small volume they were easily distinguished from background fluorescence and appeared as bright spots by fluorescence microscopy. Every detected RCA product gives rise to one distinct spot and enumeration of these enabled quantification in single cells.

Prominent detection signals confirmed that the probe indeed hybridised to DNA in the nuclei of the human cells. However, signal concentration was so high that individual RCA products could not be distinguished. Reaction products were also observed in the mouse cells, despite their lack of the exact target sequence (the mouse sequence differed from the human sequence in c.4G > C, c.18deT and c.19T > C positions in the region targeted by the probe) (Fig. 1bi).

To detect the colocalisation of histone H3 with Alu-repeats in genomic DNA, the DNA-directed PLA probe was used together with the primary antibody and the secondary PLA probe described above. Proximal binding of these two probes templated the hybridisation and ligation of two subsequently added oligonucleotides, creating a circular DNA molecule, which was then replicated by RCA and detected using fluorescent hybridisation probes as above (Fig. 1c and d). Hybridisation-based in situ PLA produced ~160 signals per cell in human cells, a more than fivefold increase when compared with ~30 signals per cell observed in mouse cells lacking this precise genomic sequence (Fig. 4). When an irrelevant rabbit IgG was substituted for the primary antibody, or when the primary antibody or the DNA-binding PLA probe was omitted negligible numbers of signals were detected (median = 0, upper percentile ≤1), regardless of which cell line was used (data not shown).

Genomic DNA-templated in situ PLA

To improve the selectivity of detection of the targeted genomic DNA sequence, we decided to utilise the genomic DNA sequence itself to template the ligase-dependent circularisation of oligonucleotide probes. In this design the single-stranded genomic DNA will thus act as a PLA probe. For this second design, the primary antibody and secondary PLA probe were applied to the cells with fixed and partially digested DNAs. This genomic DNA provided one of the two templates required for circularisation of the pair of oligonucleotides, the other template being the oligonucleotide attached to the PLA probe (Fig. 2a). An intentional G/A mismatch in the probe–target hybrid was used to enzymatically cleave the genomic sequence at the site of probe binding using the mismatch-specific MutY glycosylase (MutY) and EndoIV [20]. This gave rise to a free 3′ end in the genomic DNA sequence that was used to prime the RCA. To ensure that the signals indeed reflect bona fide PDIs, we designed the reactions so that a short oligonucleotide was included in the circular DNA molecule, templated by the oligonucleotide attached to the anti-rabbit Ig antibody. This sequence provided a second detection site in the RCA products in addition to the one already incorporated in the circularisation oligonucleotide. The inclusion of this extra DNA segment allowed us to distinguish any probes that had been circularised without the participation of the anti-histone antibody. We ensured that both segments were reflected in the RCA product by detecting these two motifs using hybridisation probes with distinct fluorescence. Thus,
only RCA products detectable with probes directed against both segments were considered indicative of PDI (Fig. 2b and c). As expected, the use of genomic DNA as a PLA probe increased the proportion of signals in human cells 13-fold when compared with mouse cells for which signals decreased to 5 signals per cell (Fig. 4). However, with this more stringent probing, the total number of signals found in human cells decreased to ~80 signals per cell. A median of 0 signals (upper percentile ≤ 1) was observed in both cell lines when an irrelevant rabbit IgG was applied instead of the primary antibody, or when no primary antibody or circularisation oligonucleotides were added.

**Padlock probe-based in situ PLA**

Even though genomic DNA-templated in situ PLA improved signal detection selectivity, it still produced false positive signals in the mouse cells despite the absence of the exact target sequence. As discussed above, hybridisation of the circularisation oligonucleotides to similar but not identical sequences might allow ligation and we thus wanted to replace T4 DNA ligase with a more stringent enzyme, that is, Amp-ligase, which tolerates higher temperatures and provides more specific target recognition. However, antibodies are likely to be sensitive to temperatures substantially exceeding 37°C. Hence, to use the more specific Amp-ligase the reaction scheme had to be changed. The DNA-binding probe was first hybridised and ligated at the higher temperature of 45°C before antibody incubation at 37°C. The DNA circle was formed from one padlock probe instead of using two circularisation oligonucleotides; thereby, both target complementary sequences were linked into one molecule [10,19]. However, the circularisation probes have to be partially complementary to the PLA probe used for detecting the protein component of the PDI, but these reagents must not be brought together because of this complementarity during the incubation with the PLA probe.

To overcome this obstacle, we have introduced a general approach where one of the DNA reagents, here the padlock probe, includes two hairpin structures shielding the complementarity to the PLA probe used for protein detection. Several uracil bases were incorporated into one strand of each hairpin structure to provide a substrate for enzymatic digestion. This allowed us to liberate the sequences of the padlock probe that are complementary to the PLA probes after both probes had independently bound their targets (Fig. 3a). As a consequence, the padlock probe could be hybridised onto the genomic DNA and ligated by Amp-ligase without interfering with the subsequent primary antibody and the secondary PLA probe incubations. Only when all probes had bound were the complements in the hairpin structures removed by digestion with UNG, to remove uracil-bases, and the DNA-backbone was cut at the abasic sites by EndoIV treatment [20]. This liberated hybridisation templates for adjacent PLA probes that templated the ligation of the remaining parts of the padlock probes to recreate circular, now somewhat smaller, DNA molecules. In the same step the G/A-mismatches between genomic DNA and circularised probes were MutY/EndoIV cleaved to create a free 3’ end in the genomic DNA at the site where the circularised probe had bound. As described for genomic DNA-templated in situ PLA, incorporation of an additional sequence element in this circle was required to distinguish correctly reacted probes. This time the incorporation of the additional oligonucleotide – resulting in two-coloured RCA products – also served to control for the UNG/EndoIV cleavage because the hairpin-padlock probe alone also could be amplified by RCA, if not digested by UNG/EndoIV. In that case it served as a regular padlock probe, detecting genomic DNA but independent of protein binding (Fig. 3b), while two-coloured signals represented proximity between the DNA sequence and the target protein (Fig. 3b).

In this manner, the proximity between Alu sequences and histone proteins interacting with the DNA was visualised with high selectivity, producing ~25 PDI signals per human cell, while negligible numbers of signals were observed in the mouse cells, rendering the detected signal in human cells 500 times higher than in mouse cells (Figs 3c and 4). As above none of the technical controls utilising irrelevant rabbit IgG, omitting the primary antibody or the DNA-binding PLA probe resulted in signals (both median and upper percentile equals zero) in either cell types (data not shown). Results were similar when detection of genomic DNA using the hairpin-containing padlock probe was compared with those using a regular padlock probe without hairpin segments, hence the hairpin structures did not impair detection or amplification (Supplementary Fig. 1).

**Discussion**

The Alu sequence targeted in the assay described is present in ~60,000 copies in the human genome and is thus presumably often located in proximity to histone H3, a component of the nucleosome. To render the nucleus accessible to all reagents required for detection, the nuclear membrane had to be permeabilised and some proteins removed. Here we had to find a balance between gaining access to the target sequence and retaining proteins of interest in place. Therefore, the fixation and permeabilisation steps required careful titrations that may need to be redone for every new protein and sequence of interest.

We chose to exclude cells stained by the mitotic marker anti-histone H3 (phospho-S10) from our analyses to avoid a source of variation among the investigated cells. When Alu-sequences were targeted with padlock probes, the detection efficiency of ~0.13% implies that a great part of the DNA is not accessible for probing under the conditions used. Further work to increase the accessibility of the DNA, without loss of bound proteins, will be required to improve the efficiency of the method for the detection of proteins binding single copy genes.

As expected, the hybridisation-based approach for detecting the target DNA sequence was the least selective, producing ~160 signals per human cell, but also ~30 signals per mouse cell. Using genomic DNA directly to template one of the two ligation reactions required to form the amplifiable DNA circles, instead of hybridising an oligonucleotide to the target sequence and utilising this as template for ligation, increased the ratio between signals detected in human and mouse cells to ~13. At the same time the number of signals in human cells dropped by approximately 50% compared with hybridisation-based in situ PLA. We assume that most of the signals lost were false positives, as those from the negative control mouse cells dropped by 80%. As restriction cleavage and exonucleolysis produce a substantial amount of ssDNA, increasing the risk of cross-hybridisation of the circularisation probes, we ensured that only RCA products resulting from recognition of both the genomic target sequence and the bound protein were scored. This was achieved by allowing the histone
H3-bound PLA probe to guide the incorporation of an additional oligonucleotide in the circularised DNA molecules. Thereby, two probes labelled with distinct fluorophores could be used to detect RCA products. The first one recognised a motif from the circularisation probe specific for genomic DNA, and the other revealed the presence of the additional oligonucleotide contributed by the anti-histone antibody. In the data analysis, only two-coloured RCA products were regarded as true signals. Signals found outside the nucleus, in the cytoplasm or outside the cells were also discarded from the analyses.

To further improve the selectivity of the assay and discriminate between closely related sequences [9,17], we switched to the thermostable Amp-ligase, necessitating a more elaborate probe design. Amp-ligase requires a higher temperature and is more sensitive to mismatches at the ligation site than T4 DNA ligase. As a consequence, ligation of the DNA-binding probe would need to occur before the protein is detected because the antibodies do not endure incubations at temperatures substantially higher than 37 °C. However, if the circularisation oligonucleotides were applied and ligated before binding of the PLA probe, then the oligonucleotide carried by the PLA probe intended for protein detection could hybridise to the DNA-bound circularisation oligonucleotides and be brought into proximity independently of the presence of the appropriate target protein. Hence, we had to mask the PLA probe complementary parts of the DNA-binding probe by introducing two hairpin structures, which hide the complementary parts of the oligonucleotide. With this padlock probe-based in situ PLA, we observed a 500-fold increase in signal count in human cells when compared with mouse cells. Comparing the number of PDI signals and Alu-sequences detected by padlock probe-based in situ PLA, 40% of all detected Alu-sequences appear to be in proximity to a histone H3. This may reflect that not all Alu-sequences are in proximity to a histone H3 or that the decrease in detection is because of the fact that neither antibody binding nor enzymatic reactions are 100% efficient.

The combination of padlock probes for DNA detection with oligonucleotide-conjugated antibodies for protein detection thus resulted in selective detection of protein–DNA complexes via a proximity ligation reaction. Further optimisation will be required to enable robust detection of proteins binding promoters of single copy genes. Such methods will be helpful to identify cells with specific PIDs and epigenetic changes at certain sequences, and their location within a tissue section. Combined with related methods to detect genes, transcripts and proteins at the single cell level, much more detailed analyses of cellular function will become possible. In a further perspective new diagnostic opportunities may arise as studies of the regulation of gene expression at the single cell level are enabled.

Funding
This work was supported by grants from the Knut and Alice Wallenberg Foundation, the Beijer foundation, The Swedish Foundation for Strategic Research; the European Union Sixth Framework Programme project Enlight [037545] and the European Union Seventh Framework Programme project DiaTools [259796].

Acknowledgements
We would like to thank Elin Falk Söqvist for help with counting the Alu-repeats in the human and mouse genome, and Sara Henriksson, Ida Grundberg and Mats Nilsson for discussions about padlock probes.

Appendix A. Supplementary data
Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.nbt.2011.08.002.

References
18. Kravev, A.S. et al. (1980) The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. Nucleic Acids Res. 8, 1201–1215