7. Results

We aimed to develop innovative methodologies applicable for whole genome analysis, specifically for oligonucleotide fingerprinting (OFP) and resequencing of genomes. To this end two complementary assays were developed: 1). liquid-based hybridisation assay with real-time detection in miniaturised array platforms using induced Fluorescence Resonance Energy Transfer (iFRET) technology, and 2). microarray-based hybridisation assays with LNA-modified oligoprobes

With the use of miniaturised array formats and multiplexed hybridisation, the speed and throughput of the experiment system was significantly increased. By using optical detection methods (laser excitation + CCD-camera), the hybridization assay was down-sized to nano-liter scale, thus reducing experimental costs and increasing sample number per assay.

7.1 Liquid-based hybridisation assay with real-time detection

We have developed a new iFRET protocol for hybridisation of short oligoprobes in solution. The assay is functional in nano-liter volume format and is suitable for SBH applications. The liquid-based hybridisation requires only accurate delivery of the oligoprobes and targets into individual wells or onto a surface, thus skipping procedures common to majority of conventional methods: immobilisation DNA to the surface and removal of the unhybridised oligoprobes.

Initially all the oligoprobes were examined with a model set of synthetic ssDNA targets ranging from 12-45 nt length for establishing assay conditions to be applied in the clone screening procedure. In model experiments, we achieved down-sizing of our liquid-based hybridisation assay using iFRET technology from 100 µl in a cuvette (detected using a spectrophotometer) to 2 µl droplet on glass support, and further to 30 nl in nano-well plate (measured using a CCD-camera). Through the application of nano-dispensing techniques, re-usable nano-structured high-density plates, and miniaturisation of the hybridisation process to the nanowell format, we show that the reaction volume can be reduced down to about 30 nl (Fig.11) indicating significant reduction of analysis cost per sample.

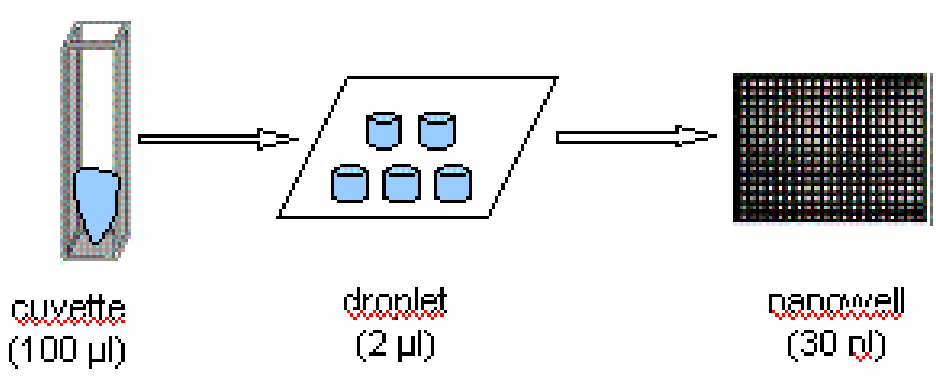


Figure 11. Down-sizing of iFRET technology.

7.1.1 Preparation of RNA templates

RNA as a hybridisation template has several advantages, such as 1). single-stranded character, 2). easily to produce large quantities of RNA 3). easy fragmentation process. In this study, RNA was generated in vitro from purified PCR product and then fragmented into short segments using 10X fragmentation buffer at high temperature. Short RNA segments reduce the potential of RNA molecules to form secondary structures, and therefore enhance the availability of binding region for hybridisation. 1 µl fragmented RNA sample was loaded into individual wells on the RNA 6000 Nano Labchip® and analyzed in the Bioanalyzer.

The Agilent 2100 Bioanalyzer provides a robust platform for the quantification and high resolution of DNA/RNA fragments via electrophoresis in microfluidics chips (Crocitto et al. 2004). Virtual gels were visually analyzed and the fluorescence data was tabulated and graphed. For the fluorescence versus time data plot below (Fig. 12), the corrected RNA area was 19.14, the RNA concentration wass 33.69 ng/µl, and the length of the RNA was below 100 nt.

7.1.2 Preparation of single-stranded DNA templates

Although advantageous as a hybridisation target, RNA is prone to degradation and has high degree of secondary structure formation, which can interfere with hybridisation. To minimise secondary structure of target, steps must be taken to reduce these effects, such as fragmentation of RNA as performed in this study.

Therefore DNA as hybridisation target was taken into consideration, however, dsDNAs do not hybridise to oligoprobes without prior denaturation by heat or alkali treatment. Moreover, these denaturation steps are difficult to control, especially when

a large number of PCR reactions have to be analysed in parallel. The efficiency of hybridisation is relatively low even after denaturation step (Holloway et al. 1993). Hence we adapted a method for preparing ssDNA templates (Nikiforov et al. 1994). It includes amplifying dsDNA templates in which one of the PCR primers contains five phosphorothioates at its 5’ end and the opposite strand primer remains unmodified. Following amplification, one strand of each amplicon, in which the degraded strand lacks the 5' terminal phosphorothioates, is degraded with T7 gene 6 exonuclease, leading to ssDNA production.

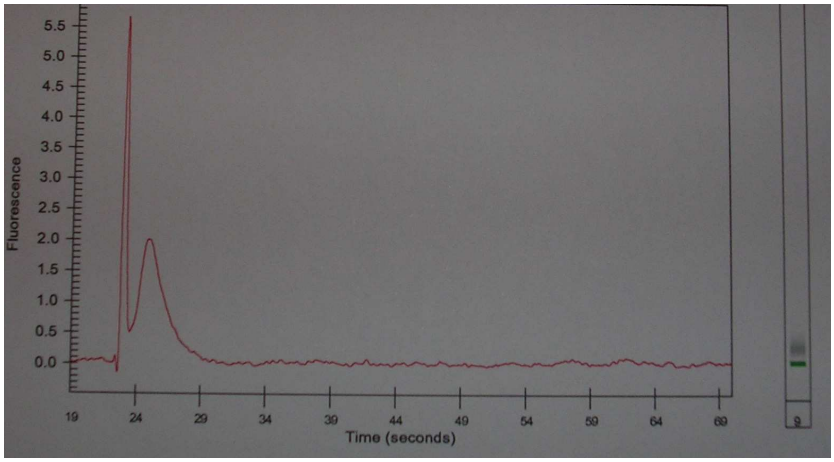


Figure 12. Electropherogram image of representative fragmented RNA (produced from clone N05). From right to left, the fluorescence peaks correspond to the digestion fragment and the front marker.

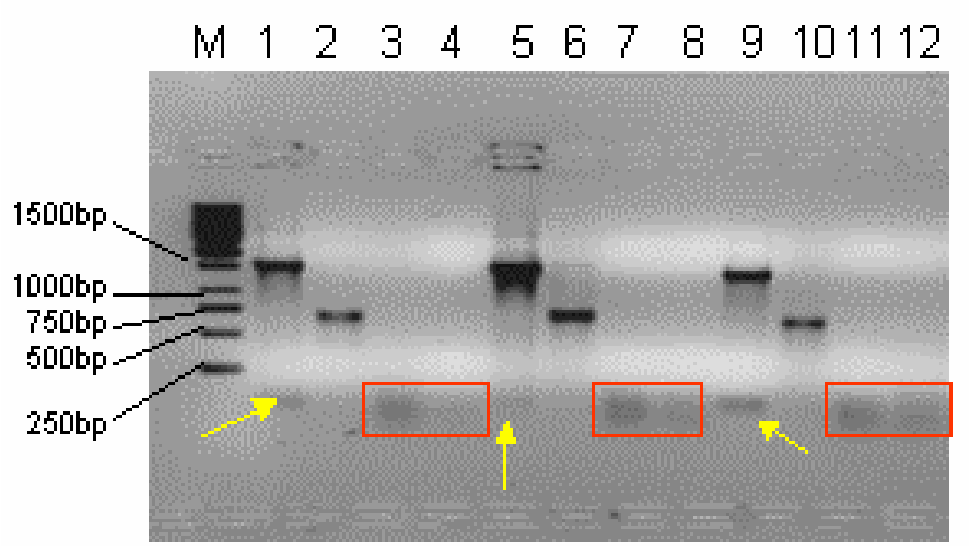


Figure 13. Gel analysis of representative PCR reactions without or with T7 Gene 6 exonuclease treatment. Lane M, O'GeneRuler^{TM} 1kb DNA ladder (5 µl). Lane 1, 5 and 9, 5 µl PCR products (dsDNAs) of clone D11, D18 and D20 before T7 Gene 6 exonuclease treatment. Lane 2, 6 and 10, 10 µl PCR products of clone D11, D18 and D20 after T7 Gene 6 exonuclease treatment (ssDNAs). Lane 3 and 4, Lane 7 and 8, Lane 11 and 12, represent 10 µl fragmented forward and reverse strand of PCR products of the clone D11, D18 and D20 upon DNase treatment (short ssDNAs), respectively.

Fig. 13 visualised the DNA products separated in the gel. The weak bands of lower molecular weight indicated by the yellow arrows represented primer artifacts. After T7 Gene 6 exonuclease treatment, the primers without phosphorothioate modification were digested, thus reducing the signals in lanes 2, 6 and 10. Afterwards the ssDNAs without primer artifacts shown in lanes 2, 6 and 10 were subjected to DNase digestion, but without any further purification procedures. The bands in the red squares correspond to the short ssDNA fragments as the consequent hybridisation targets, instead of primer dimmers. ssDNA (the bands in lanes 2, 6 and 10) migrates faster than the double-stranded counterpart (the bands in lanes 1, 5 and 9). Because ssDNA binds ethidium bromide poorly, those bands appear much more faint than the equivalent amount of dsDNA loaded on the gel.

Incorporation of phosphorothioate-linkages into one of the amplification primers blocked digestion of that strand completely, so that the single-stranded amplification product could be obtained after digestion with T7 gene 6 exonuclease. This

exonuclease had a high activity in PCR buffer conditions, therefore it was added to the reaction mixture after amplification and the exonuclease reaction was allowed to proceed for 1 hour at room temperature or for 15-30 minutes at 37°C. No further purification steps were needed before the hybridisation.

7.1.3 Proof-of-principle of iFRET in quartz cuvettes (~100 µl format)

We tested the hybridisation behaviour of a comprehensive set of dye-labeled 6-, 7- and 8-mer oligoprobes sharing the same core sequence. In the course of this work, we have shown that the hybridisation detection method worked also for LNA-modified oligoprobes, with strikingly better performance, i.e., higher signal with lower unspecific background. Reliable data could be also obtained with 6-mer LNA-modified oligoprobes. LNA molecules proved to be superior for hybridisation by forming exceedingly stable duplexes with complementary nucleic acid target.

iFRET is a variation of FRET that is particularly well-suited for the detection of nucleic acid hybridisation. The underlying mechanism involves monitoring changes in fluorescence that are the result of an energy transfer reaction between a specific pair of donor and acceptor moieties. The iFRET entails using a dsDNA-specific intercalating dye (e.g., SYBR Green I) as the donor, with a conventional acceptor (e.g., Bodipy, Rox or Cy5) affixed to the short oligoprobes. Hybridisation of the oligoprobe to its complement induces excitement of the donor dye and subsequent energy transfer to the acceptor dye. The energy transfer reaction (and concomitant hybridisation status) can easily be detected by monitoring the fluorescence output of the acceptor dye.

7.1.3.1 Lower hybridisation background for LNA-modified oligoprobes

Background measurements were carried out in quartz cuvettes in 100 µl buffer-2 solution containing Sybr Green I at a 1:1000 dilution and 1 µM oligoprobes (non-modified 7-mer DNA oligoprobe OP-17 and LNA-modified oligoprobes with the same core sequence but different length, OP-17L, OP-17L-6nt and OP-17L-5nt), without targets. The hybridisations were detected by thermo-controllable spectrophotometer and processed using Avasoft 5.1 software.

A background comparison between LNA-containing and non-modified DNA oligoprobes was shown in Fig. 14. The signal intensities were measured at the wavelength of 670 nm. It was observed that LNA-modified oligoprobes (OP-17L, OP-

17L-6nt and OP-17L-5nt) delivered much lower background than 7-mer non-modified DNA oligoprobe (OP-17) of the same core sequence. Tests with 7-, 6- and 5-mer oligoprobes sharing the same 5-mer core sequence showed a clear correlation of background reduction with the oligoprobe length.

iFRET signal of Cy5 fluorescence was detected (maximum emission at 670 nm) as a representation of the hybridisation events only in the presence of both oligoprobes and complementary RNA or DNA molecule targets. Comparising with the iFRET hybridisation signals in Fig. 17, when the target molecules were absent, the background fluorescence (direct fluorescence of acceptor) of non-hybridised oligoprobes was minor when excited directly by the excitation with light suitable for the donor fluorophore. In all experiments the background fluorescence of oligoprobes in buffer-2 was marginal (data not shown).

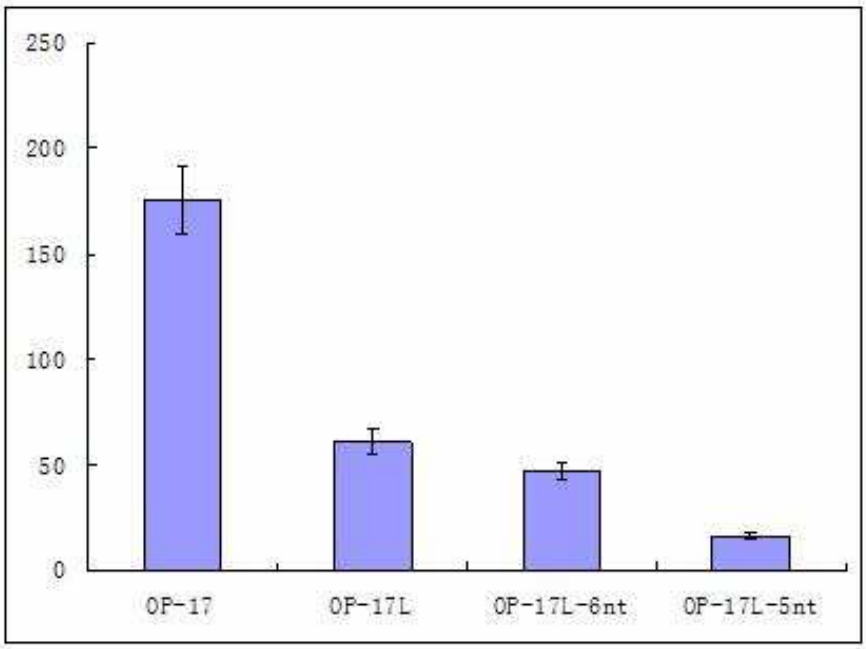


Figure 14. Background comparison of a selected oligoprobe set sharing the same 5-mer core sequence. The Y-axis represents relative signal intensity units. Data were plotted as a mean and SD from three independent measurements.

7.1.3.2 Multiplexed detection of iFRET signals

In order to examine the influence of the dye on signal generation and signal intensity, oligoprobes with different fluorescent dyes attached were tested. Three acceptor dyes (Bodipy, Rox and Cy5) were selected due to the existing overlaps between the fluorescence emission spectrum of these three acceptors and the absorption spectrum of the donor (SYBR Green I). We revealed that fluorescent donor-receptor pair, such as Bodipy-SYBR or Rox-SYBR, for generation of specific FRET signal. However, the best signal-to-background ratio was achieved when Cy5-labeled oligoprobes were used. Multiplexing, i.e., using at least two oligoprobes labeled with different dyes was also possible.

In principle, appropriate filtering should permit detection of more than one acceptor dye, thus allowing for multiplexing of hybridisation oligoprobes. This possibility was examined for three fluorescent dyes (Bodipy, ROX and Cy5). Emission filters with a maximum at 560 nm for Bodipy TMR fluorescence, at 630 nm for the ROX filter and at about 670 nm for Cy5 were used. Fig. 15 shows the detection of hybridisations using three oligoprobes simultaneously. Synthetic DNA oligonucleotides that are specific for distinct oligoprobes were probed with either an equal molar mixture of three (Fig. 15) or of two oligonucleotides (Fig. 16), or separately with each oligoprobe (Fig. 18). As the emission of Bodipy dye is close to the emission of Sybr Green I, it was difficult to distinguish between these two signals, making signal measurements from Bodipy-labeled oligoprobes undiscriminative. Our results suggest that most reliable data could be obtained using a combination of ROX- and Cy5-labeled oligoprobes (Fig. 16, curve marked with star).

7.1.3.3 Quantification of iFRET signals

A 100 µl buffer-2 solution containing Sybr Green I at 1:1000 dilution, 100 pmol oligoprobe OP-04L-6nt (sequence and positions of LNA modification are shown in Table 2) was hybridised in quartz cuvettes with different amount (0 pmol, 5 pmol, 10 pmol, 20 pmol, 30 pmol and 40 pmol) of full matching short synthetic single-stranded target (DNA-10 in Table 4), respectively. Quantification of iFRET signals is shown in Fig. 17. Despite low concentrations of target DNA (~50 nM), the hybridisation signals were still distinguishable, thus making miniaturisation to nanoliter volumes feasible.

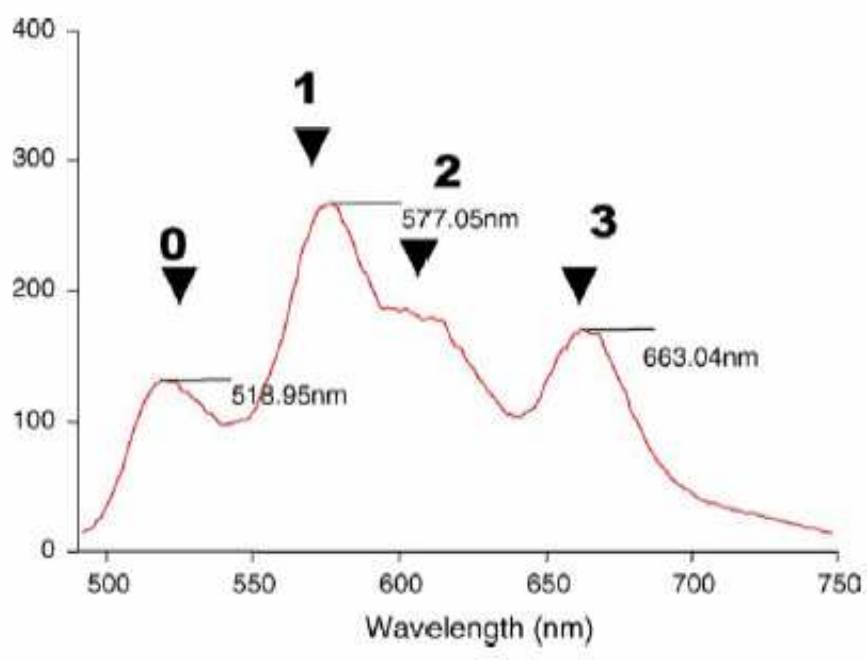


Figure 15. Hybridisation of three pooled 7-mer oligoprobes labeled with Cy5, ROX and Bodipy fluorescent dyes (OP-13-8nt, OP-32 and OP-33, Table 3) to two synthetic DNA targets (DNA-2 and DNA-3, Table 4) in the same cuvette. Oligoprobes OP-13-8nt and OP-33 match to DNA-2, and oligoprobe OP-32 match to DNA-3. Oligoprobes were used at equal concentrations, and in 10-fold excess to each DNA target. Hybridisation was performed in 100µl volume in a cuvette and detected by UV-Vis-NIR Spectrometers (Varian, USA). The observed peaks are: 0 (518.95 nm) from intercalating dye Sybr Green I; 1 (577.05 nm) from Bodipy-oligo; 2 (not identified, as it overlaps with the Bodipy-labeled oligoprobe emission) ROX-labeled oligoprobe; and 3 (663.04 nm) from Cy5-labeled oligoprobe hybridisation. The Y-axis represents relative signal intensity units.

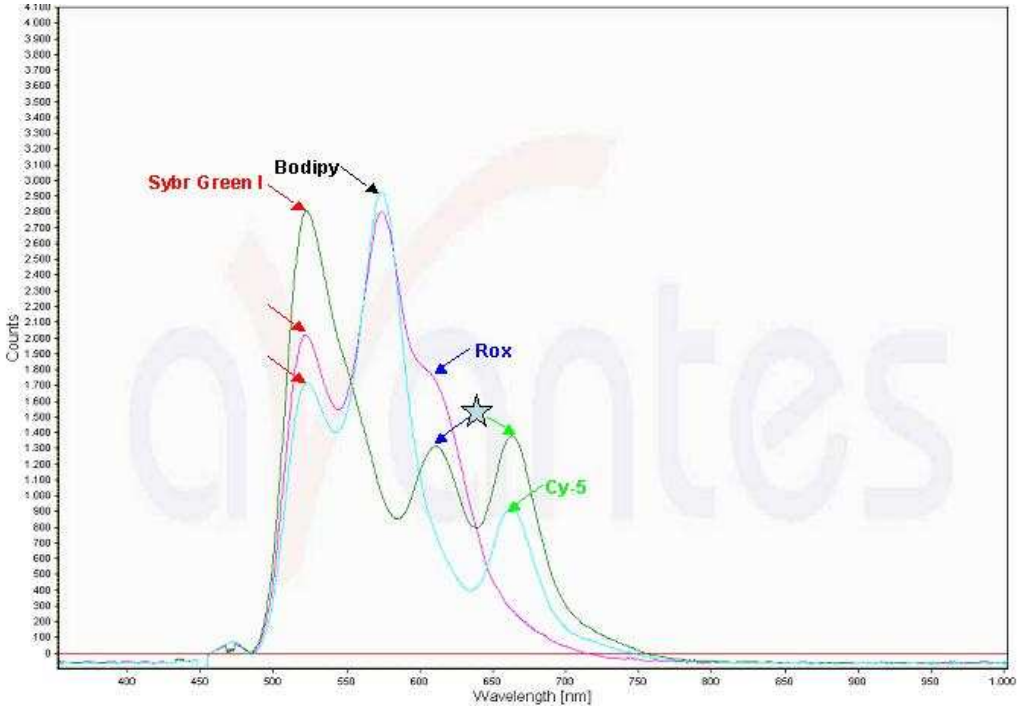
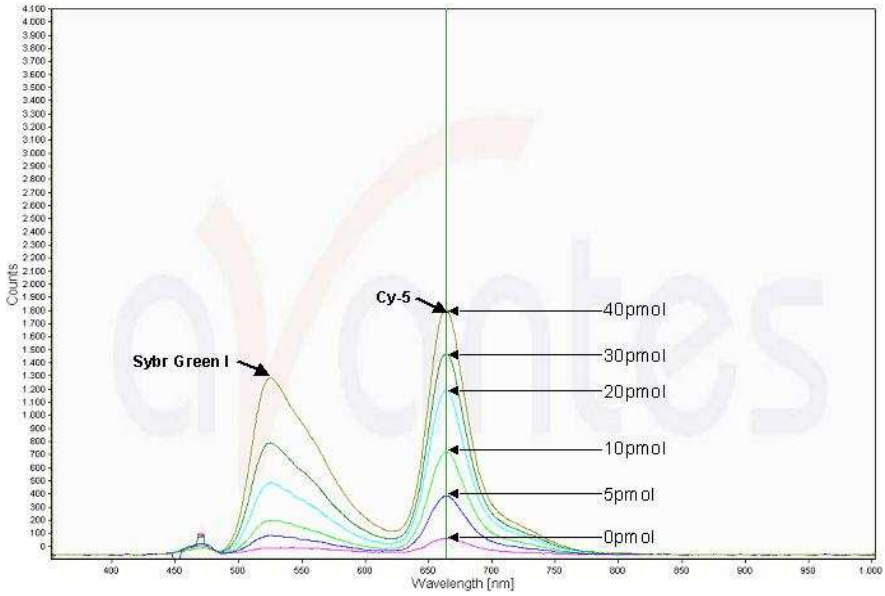


Figure 16. Hybridisation of two pooled 7-mer oligoprobes labeled with two of the three fluorescent dyes (Cy5, ROX or Bodipy) (OP-17, OP32 and OP-33, Table 3) to two fully matched synthetic DNA targets of the three (DNA-2, DNA-3 and DNA-7, Table 4). Oligoprobe OP-17 matches to DNA-7, OP-33 matches to DNA-2, while OP-32 matches to DNA-3. Green curve: ROX and Cy5 labeled oligoprobes OP-32 and OP-17, together with DNA-7 and DNA-3. Pink curve: Bodipy and ROX labeled oligoprobes OP-33 and OP-32, together with DNA-2 and DNA-3. Blue curve: Bodipy and Cy5 labeled oligoprobes OP-33 and OP-17, together with DNA-2 and DNA-7. Oligoprobes were used at equal concentrations, and in 10-fold excess compared with DNA targets. Hybridisation was performed in a 100 µl volume in a cuvette and detected by a Thermo-controllable spectrophotometer (Avantes Fiberoptic, The Netherlands). The observed peaks are: red arrows (518.95 nm) from intercalating dye Sybr Green I; black arrow (577.05 nm) from Bodipy labeled oligoprobe hybridisation; blue arrows (620 nm) from ROX labeled oligoprobe hybridisation; and green arrows (663.04 nm) from Cy5 labeled oligoprobe hybridisation. The curve marked with star denotes the setup with best separated signals from pooled oligoprobes using a combination of ROX- and Cy5-labeled oligoprobes. The Y-axis represents signal intensity units.

A.



B.

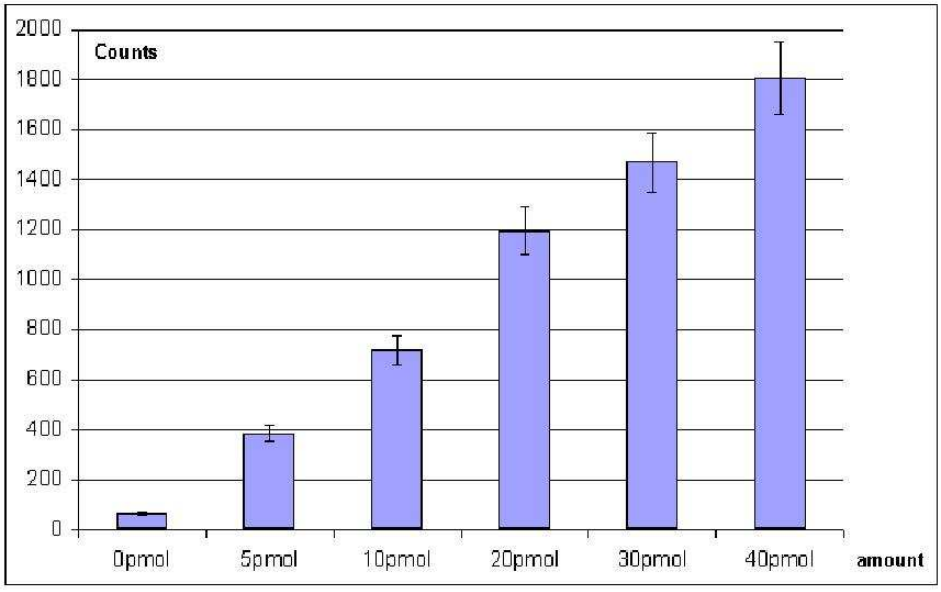


Figure 17. Quantification of iFRET signals. **A**. iFRET signals measured in quartz cuvettes by thermo-controllable spectrophotometer. Hybridisation of target DNA-10 at different comcentrations (0 pmol, 5 pmol, 10 pmol, 20 pmol, 30 pmol and 40 pmol) with matching oligoprobe OP-04L-6nt was performed. The black arrows indicate the signals from Sybr Green I (left) and Cy-5 (right), respectively. **B**. Quantification of the hybridisation signals. The Y-axis represents relative signal intensity units. Data were plotted as a mean and SD from three independent measurements.

7.1.4 Verification of iFRET in 2 µl droplet format

In standard ‘Sandwich’ set-up we have used hybridisation volumes of 2 µl and glass as surface to downsize the reaction volume. The ‘sandwich’ scheme of the droplet experiment is demonstrated in Fig. 5. 1 µl short ssDNA or ssRNA target and then 1 µl oligoprobe in buffer-2 solution including 1:1000 diluted SYBR Green I were dispensed onto the glass slide surface. The slide was then covered with a second glass slide, spaced with an adhesive tape providing a distance of 255 µm between the two slides. The final concentrations in the hybridisation mixture were 0.5–1.5 µM for oligoprobes and 0.15–0.5 µM for the targets.

7.1.4.1 Effects of oligoprobe length on hybridisation performance

For OFP, we aimed at employing as short oligoprobes as possible and tested some oligoprobes of different length (pentamer, haxamer, heptamer and octamer), but our results suggested that signals obtained with pentamers (LNA-modified or non-modified pure DNA) were unacceptably low (not shown), so that the more appropriate oligoprobe length was shown to be 6- or 7-mers.

Hybridisation results with 8-, 7- and 6-mer oligoprobes sharing the same 6-mer core sequence showed a clear correlation of signal reduction with the decrease in oligoprobe length (Fig. 18). A typical signal comparison between LNA-modified and non-modified DNA oligoprobes is shown in Fig. 18. It was observed that even 6-mer oligoprobes, modified with only single LNA base, delivered much stronger signal (e.g., two times higher) than non-modified 6-mer DNA oligoprobe of the same sequence and was comparable to that of a 7-mer DNA oligoprobe. Therefore in further experiments 7- or 6-mer oligoprobes were used.

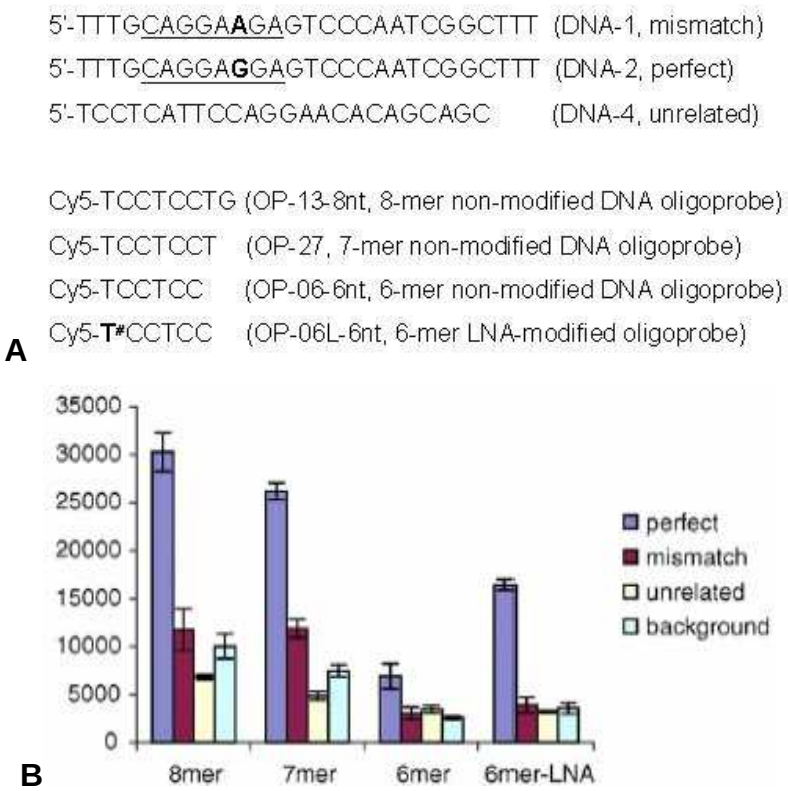


Figure 18. Assessment of minimal oligoprobe length suitable for iFRET detection. **A**. Oligoprobes of different length but with the same 6-mer core were used for hybridisation with three DNA targets: perfectly matched, with a single mismatch (indicated in bold) and unrelated ones (no match). Sequence fragment complementary to the oligoprobes is underlined. **B**. Hybridisation results for oligoprobes with different length. DNA hybridisation was performed in droplets. For each oligoprobe, a background signal only from the oligoprobe and Sybr Green I in solution without DNA target was also measured and subtracted from the signal of the oligoprobe–target mixture. Data are presented from three independent experiments. Data were plotted as a mean and SD from three independent measurements.

7.1.4.2 Stringency of hybridisation conditions

Routinely, hybridisations in a 2 µl micro-droplet volume placed on a glass slide were performed (set-up shown in Fig. 5), with each sample in duplicate. The standard hybridisation assays using Cy5-labeled oligoprobes were performed by tracking the changes of iFRET signals using both excitation and emission wavelengths of 540 and 670 nm, respectively. The slides were placed onto a thermal cycler with a heating increment of 3 ℃/min. The im ages were captured at different temperature points (10℃, 20℃, 30℃ and 40℃) of a pplied temperature ramp.

Fig. 19 shows the monitoring of oligoprobe hybridisation with different short synthetic ssDNA targets. For most of the oligoprobes, the results of hybridisation should be negative (no matching sites), but for the oligoprobe OP-17, DNA targets "a" and "c" should be positive, whereas DNA target "b" is a single mismatch. Signals from perfect and mismatched duplexes (1 and 2, as indicated by the arrows) were better discriminated at higher temperatures (e.g. 40 ℃).

Melting profiles obtained via repeated heating and cooling of the RNA:oligoprobe duplexes are shown in Fig. 20 for three different oligoprobes (OP-15, -28 and -30). It was shown that the hybridisation signals were very stable in the course of temperature cycling. Taking the hybridisation signals from oligoprobe OP-15 and its full-matching RNA target as an example, the signals obtained at 10℃ with heating and cooling cycles were always at the same level of 4000. Measurement of melting profiles offered the possibility of dynamic measurements of hybridisation events to discriminate the mismatching targets.

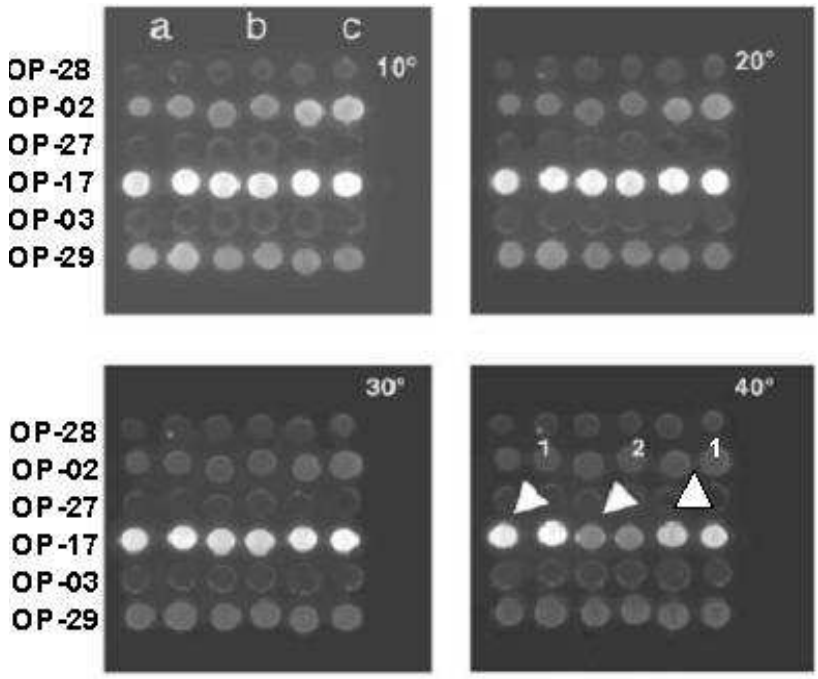


Figure 19. Monitoring of the hybridisation experiment following heating. Six 7-mer Cy5-labeled non-modified DNA oligoprobes (in Y-direction, OP-28, OP-02, OP-27, OP-17, OP-03 and OP-29; sequence shown in Table 3) were hybridised with three synthetic ssDNA targets (in X-direction, "a" (DNA-7), "b" (DNA-8) and "c" (DNA-9), see Table 4) in duplicates. Image acquisition was performed at the different temperatures between 10℃ and 40 ℃).

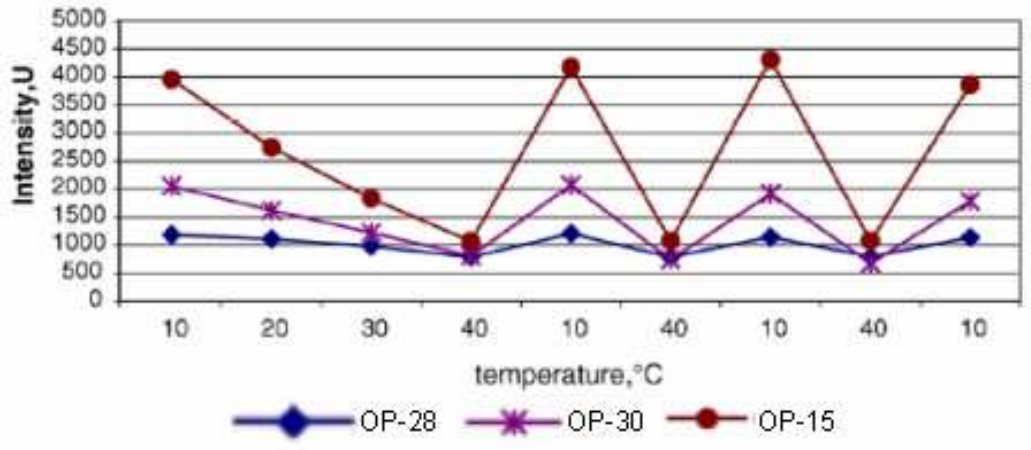


Figure 20. Fluorescence-monitored melting profiles of duplexes formed by the same fragmented RNA template (from clone B03) with three non-modified heptamers (OP-15, -28 and -30). Individual hybridisations with three oligoprobes were performed in droplets. The starting temperature of hybridisation was 10℃ and then it underwent cyclic changes to 40℃ (rate of 1.5℃/min) and back to 10℃. The data were collected for 10, 20, 30 and 40℃ temperature points. The hig hest signal intensity (Y-axis) was obtained with the fully matched oligoprobe OP-15.

7.1.4.3 Sequence-dependent hybridisation profiles

In order to evaluate the applicability of our method for OFP analysis, a pilot study with selected genomic DNA clones was performed. Fragmented RNA templates have been generated from genomic DNA clones of known sequences, which were described previously by Bauer et al. (2004). RNAs were generated from three DNA clones (D11, G11 and N05) by in vitro transcription from both ends of each respective clone, resulting in six RNA templates. Thus, additional control of hybridisation specificity was provided.

The RNA samples were hybridised with a selected set of oligoprobes. Graphical representation of different RNA hybridisation profiles with three oligoprobes is shown in Fig. 21. Clearly, it was possible to distinguish between templates originating from the same (such as RNA-G11-F and RNA-N05-F or RNA-G11-R and RNA-N05-R) or different (such as RNA-D11-F and RNA-G11-F) sequences. In this way, RNA-G11 and RNA-N05 (originating from DNA clones G11 and N05 of 1001 and 1344 nucleotides in length) belong to the same cluster, whereas the sequence of RNA-D11 (1259 nucleotides) was not related to RNA-G11 or RNA-N05.

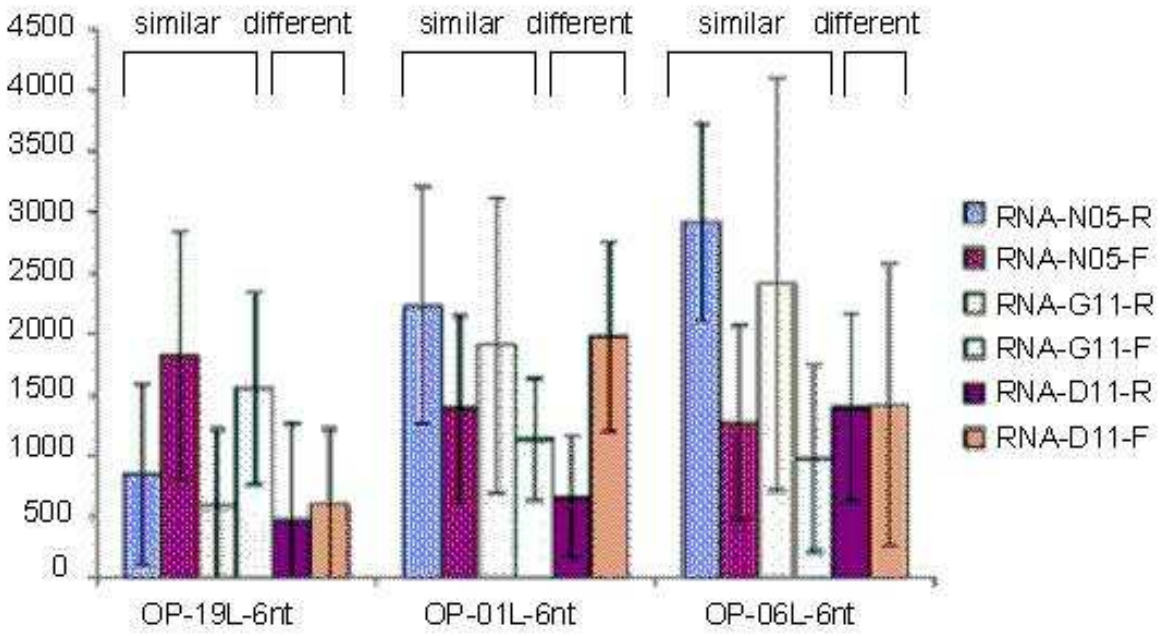


Figure 21. Sequence-dependent hybridisation profiles obtained for six RNAs with three 6-mer LNA-modified oligoprobes (OP-19L-6nt, OP-01L-6nt and OP-06L-6nt; sequences shown in Table 2). The RNA-D11-F and RNA-D11-R were produced from the same original DNA clone D11 via in vitro transcription from both the 5′- and the 3′-end, respectively. Hence, six different RNAs (RNA-D11-F and RNA-D11-R, RNA-G11-F and RNA-G11-R, RNA-N05-F and RNA-N05-R) originating from three DNA clones (D11, G11 and N05) were analysed. Hybridisations were performed in droplets, as described in the text. The background fluorescence of the oligoprobe and Sybr Green I alone was subtracted. Average values of intensities are shown with standard deviations.

7.1.4.4 Short ssDNA from PCR product as hybridisation target

We exploited the use of fragmented ssDNA as the hybridisation target. During our tests we realised that unfragmented long ssDNA can not be used for hybridisation with short oligoprobes (data not shown). We found that it was necessary to fragment the template into pieces of about ~100 nt in length, which render them suitable for interaction with short oligoprobes.

As illustrated in Fig. 22 we observed the iFRET signal when short single-stranded PCR products were complementary to the oligoprobe. Significant hybridisation signals were observed from the combinations of OP-08L with short forward ssDNA from clone D11, and OP-14L with short forward ssDNA from clone

D20. Since short reverse ssDNAs from clones D11 and D20 did not match OP-08L and OP-14L, respectively, no obvious iFRET signals were observed. Therefore, use of T7 Gene 6 exonuclease pretreatment of single-stranded amplification products with one strand phosphorothioated and the other unmodified, can efficiently eliminate by-products (e.g. remaining primer), which otherwise impede the consequential hybridisation process, and produce highly efficient quality of hybridisation signals.

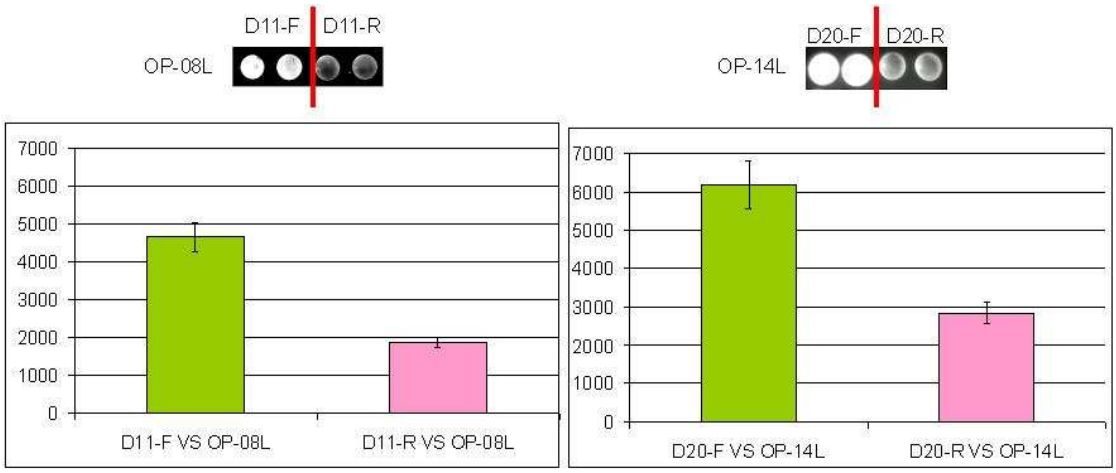


Figure 22. iFRET signals using short ssDNAs as targets. The images on the top panel represent the iFRET signals between two hybridisation pairs, OP-08L with fragmented forward (D11-F) and reverse (D11-R) ssDNA product from clone D11, OP-14L with fragmented forward (D20-F) and reverse (D20-R) ssDNA derived from clone D20. The hybridisation signals were quantified with results shown in the diagram at the bottom. Average values of intensities are shown with standard deviations.

7.1.5 Potential large-scale application of iFRET in nano-well (~30 nl format)

The micro-droplet format allowed for reduction of the reaction volume below 100 nl when using rapid dispensing techniques. For high-density arrays, the nano-well format offers a platform in nanolitre volume range. The method has been verified for nano-well plates using 500 µm pitch (400 nano-wells/cm^2) and 36 nl total volume per well. Fig. 23 presents an image acquired by CCD at 10ºC for fluorescence hybridisation detection of five synthetic short ssDNA as templates with five 7-mer non-modified DNA oligoprobes, each sample in duplicates, dispensed in two directions into a nano-well plate (experimental scheme shown in Fig. 7). Hybridisation

conditions were the same as in the 2 µl droplet experiments. iFRET signals indicated by green squares in Fig. 23 were generated by the hybridisation between oligo OP-17 and target DNA-9 (a), oligo OP-03 and target DNA-9 (b), oligo OP-31 and target DNA-9 (c), oligo OP-17 and target DNA-5 (d), oligo OP-03 and target DNA-5 (e).

Along with application in OFP, the assay described here provides an option for micro-volume processing for applications, such as PCR, towards cost reduction. Furthermore, it could potentially be used to create ultra-high-throughput screening, which is necessary for applications such as SNP detection or resequencing. In conclusion, the proposed method can be used as a robust platform for in-solution miniaturised sequence-specific clone characterization, suitable for automation in high-throughput applications.

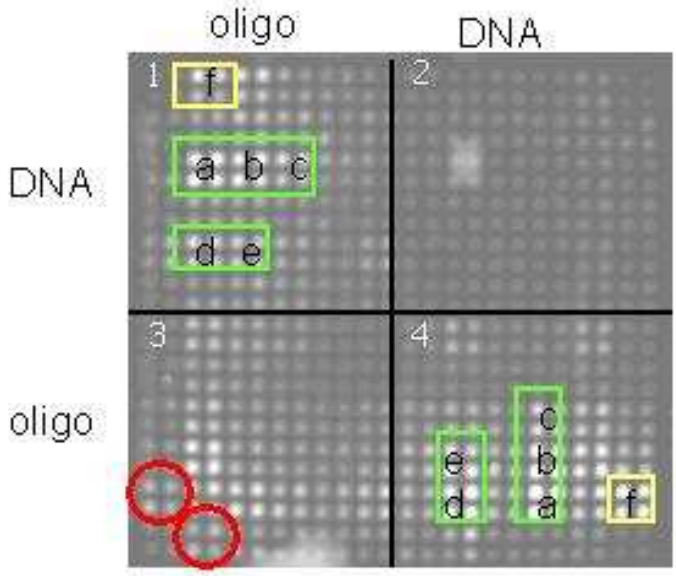


Figure 23. An image of hybridisation in a 24 × 24 nano-well plate. DNA targets (short synthetic ssDNAs such as, DNA-5, DNA-6, DNA-9, DNA-4, DNA-8) and oligoprobes (such as OP-17, OP-03, OP-31, OP-15, OP-32) were dispensed with 96 channels using a TopSpot dispensing head as shown in Fig. 7. In fields 1 and 4, both DNA and oligoprobes (whole hybridisation mixture) are dispensed, whereas in fields 2 and 3, only oligoprobes or DNAs (together with Sybr Green I), respectively. The first two columns in both fields 1 and 2 were buffer only. For normalisation of positive hybridisation signals (fields 1 and 4, signal locations indicated by green squares a~e), the background signal was substracted from the samples containing no target DNA (sample position marked with red circles, background of oligoprobe OP-17). iFRET signals indicated by yellow squares were generated by the hybridisation between oligoprobe OP-17 and its mismatch target DNA-8 (f).

7.2 Microarray-based hybridisation assays with LNA-modified oligoprobes

A robust microarray-based method was developed to detect the hybridisation events on a nanoporous membrane platform. The technique combines the advantages of non-radioactively labeled oligoprobes and the new chemistry LNA. The hybridisation procedure itself takes less than 1 hour. Clearly distinguishable hybridisation signals were achieved using as little as 4 fmol of DNA target when hybridised with Cy5-labeled, 7-mer LNA-modified oligoprobes. The described approach could dramatically reduce the number of oligoprobes needed for OFP analysis to a half, e.g., from 200-300 octamers to 100-150 heptamers, and for SBH the reduction would be to one fourth from 32,768 octamers to 8,192 heptamers.

Furthermore, by introducing the LNA into oligoprobes, the hybridisation stability of heptamer non-modified DNA oligoprobes was significantly improved. Analysis of receiver operating characteristic (ROC) curves revealed that the LNA-modification clearly improves the sensitivity and specificity of hybridisation experiment (See section 7.2.12). Clustering analysis demonstrated the application of LNA-modified oligoprobes for reliable clustering of DNA sequences to characterise unknown DNA libraries (See section 7.2.14).

7.2.1 Usage of a nanoporous membrane

We utilised a nanoporous membrane for fluorescence-based high-throughput hybridisation assays. Three-dimensional nanoporous substrates offer a higher capacity of potential immobilisation space than modified glass does. Hence, larger amounts of DNA can be retained on the membrane, resulting in stronger signal intensities and a better signal dynamic range. Furthermore, strong non-covalent attachment of DNAs to the membrane leads to better retention of the arrayed DNA. More consistent data than the ‘two-dimensional’ surface can thus be delivered (Stillman and Tonkinson, 2001). The stable attachment of soft membrane onto glass slide facilitates automation of the process.

In our studies, the experimental verification of this concept was carried out by using Pall Life Science Gene Array Slides, where a soft nano-membrane layer is bound to a glass substrate, in the format of a microscopic slide (75×25 mm^2). The Pall hybridisation membrane is a positively charged nanoporous membrane for rapid ionic binding of nucleic acids and is widely applied to macro- and micro-arrays.

Previously published analysis of the hybridisation signals demonstrated that the Pall nanoporous membrane delivers the highest signal-to-noise ratio compared to Amersham Hybond N+, Schleider and Schuell Nitran supercharged and Boehringer Nylon membranes (Alberola et al. 2004). It was also shown to be flexible enough to withstand extensive handling, to offer superior performance with non-radioactive detection systems, to be reusable and not to crack, shrink, or tear during the multiple procedures.

7.2.2 PCR amplification

After ethanol precipitation, the products from 50 µl PCR reaction were re-suspended using 25 µl water. Following the electrophoresis of 2 µl precipitated PCR products on 1.2% agarose gel contaning 1 µg/ml ethidium bromide, under UV light, the dsDNAs amplification revealed to be very specific (Fig. 24).

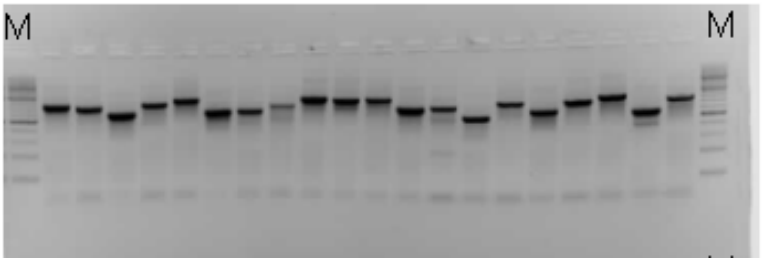


Figure 24. Gel analysis of 20 representative PCR reactions from the 66 genomic DNA clones. Lane M: O'GeneRuler^{TM} 1kb DNA ladder (2 µl). The lanes between the left and right ladder lanes represent the PCR products.

7.2.3 Hybridisation comparison of non-modified and LNA-modified oligoprobes

Non-modified hexamer DNA and LNA-modified oligoprobes gave rather weak or no hybridisation signals (data not shown). However, hybridisations of LNA-modified 7-mer oligoprobes (listed in Table 1) with the arrayed dsDNA resulted in significantly higher hybridisation signals than that derived from hybridisations with the respective non-modified DNA oligoprobes (Fig. 25). This behaviour can probably be attributed to the high affinity of LNA to DNA. One of the most important features of LNA-modified oligoprobes is high thermal stability of their duplexes with complementary DNA or RNA. Insertion of a single LNA base into oligoprobe can elevate its melting temperature (Tm) by ca. 10℃. This effect is most p rofound for oligoprobes below 10nt (Braasch and Corey, 2001). From these reasons, heptamer LNA-modified oligoprobes were used in future experiments.

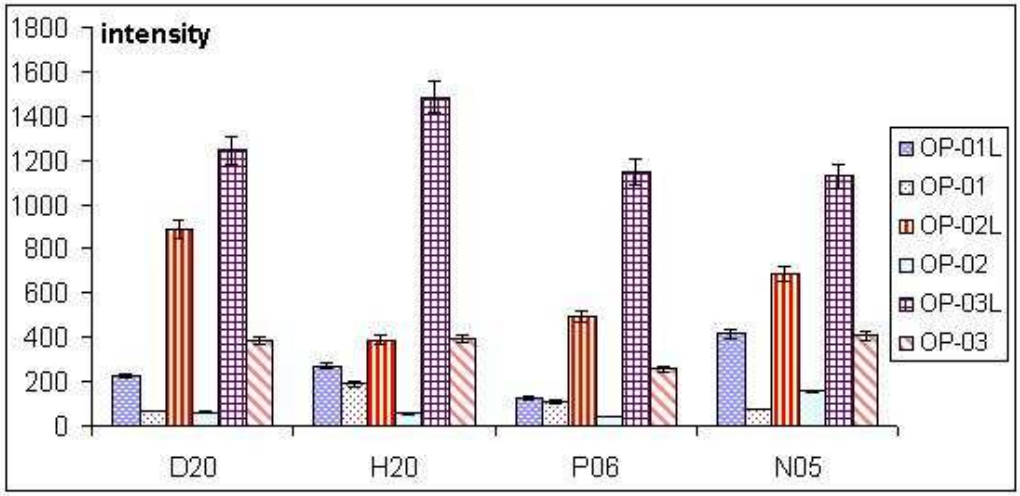


Figure 25. Hybridisation comparison of LNA-modified and non-modified oligoprobes using dsDNAs as templates. 100 droplets (40 nl) of 0.3 µM dsDNAs, produced from clones D20, H20, P06 and N05, were spotted in duplicates onto nanoporous membrane and hybridised with 7-mer LNA-modified and non-modified oligoprobes (OP-01, -01L, -02, -02A, -03, -03L), respectively. The background (fluorescent intensity around the hybridisation spots) was subtracted. Data were plotted as a mean and SD from three independent measurements.

7.2.4 Effect of the number of LNA modifications on the hybridisation behaviour

We examined the influence of the LNA modification number on the hybridisation signal intensity and determined that more than two LNA modifications per oligoprobe may result in deterioration of the signal. The higher number of modifications did, indeed, lead to a stronger signal. This was, however, accompanied by an increase in non-specific oligoprobe binding. In our experimental conditions, 7-mer oligoprobes with more than two LNA modifications had higher affinities to the DNA target and exhibit a much higher rate of false-positives (Fig. 26). For such oligoprobes it was therefore difficult to precisely discriminate mismatch in the target sequence under the short washing time and moderate washing conditions used in this study. LNA hybridises very tightly when several consecutive residues are substituted with LNA. Unless very short probes are required, stretches of more than four LNAs should be avoided. Therefore we used only one to three LNA modifications per oligoprobe in further experiments.

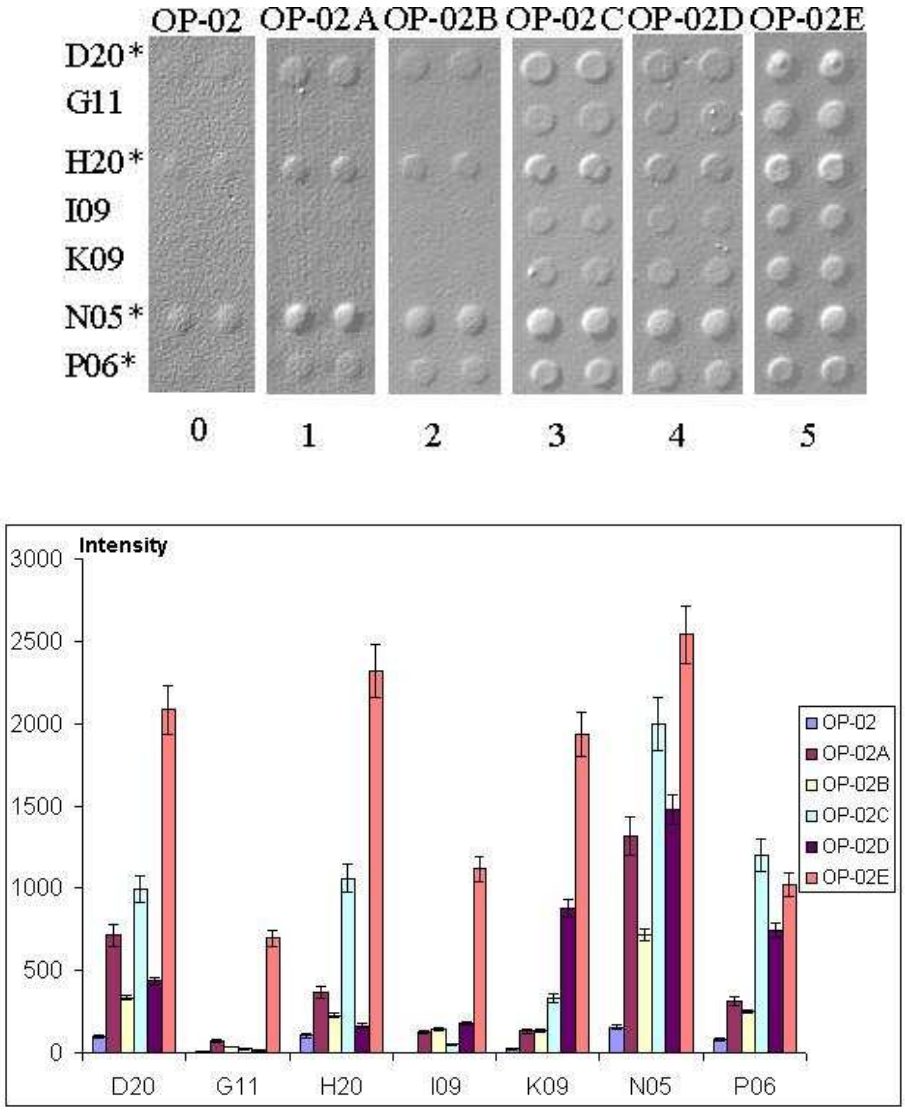


Figure 26. Effect of LNA modification number on hybridisation efficiency. 100 droplets (40 nl) of 0.4 µM dsDNAs amplified from seven clones (D20, G11, H20, I09, K09, N05 and P06) were spotted in duplicates onto nanoporous membranes and then hybridised with OP-02 with different LNA modifications (OP-02, OP-02A-E). The numbers under the images indicate the LNA modification number. The images show the hybridisation signals from seven dsDNAs, of which four (from clones D20, H20, N05 and P06, indicated by asterisk) match to OP-02, whereas the other three (from clones G11, I09 and K09) do not. The background was subtracted. The hybridisation signals were quantified as shown in the diagram in the lower panel. Data were plotted as a mean and SD from three independent experiments. Fluorescence intensity was presented in arbitrary units as described in Methods.

7.2.5 Signal detection limit and signal linearity

At the minimum dispensing volume of 100 pl of the SciFlexArrayer, the amount of DNA spotted in pl volume remained below the signal detection limit (data not shown). In this study, the volume of DNA solution at the concentration of 1 µM could be scaled down to about 20 nl or even 4 nl (Fig. 27). These volumes correspond to the immobilised DNA amounts as low as 20fmol or 4fmol, respectively. Despite such low amounts of target DNA immobilised on the nanoporous membrane, the hybridisation signals were still distinguishable from the background, thus making miniaturisation in nanoliter volumes feasible. Taken together, the detection limit in our system was found to be 4 nl of the droplet volume (Fig. 27), containing 4 fmol of the target DNA. The signal intensities have increased with increasing amount of the spotted DNA molecules. This is because nanoporous membrane has the capacity to linearly absorb a high amount of nucleic acid material.

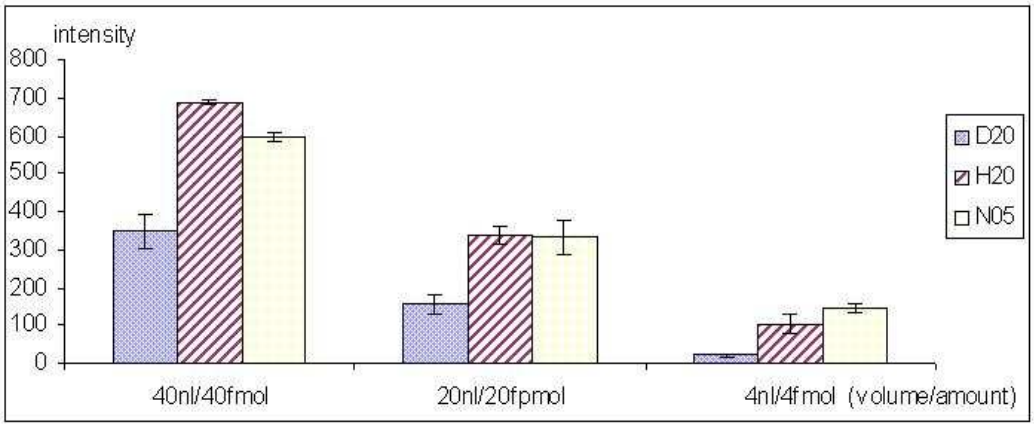


Figure 27. Determination of signal detection limit and signal linearity. Three dsDNAs produced from clones D20, H20 and N05 at a concentration of 1 µM were dispensed in duplicates onto the nanoporous membrane at different volumes (i.e. 100 droplets, 50 droplets and 10 droplets). Subsequently they were hybridised with the matching oligoprobe OP-03L. Data were plotted as a mean and SD from three independent experiments.

7.2.6 Re-utilisation of the nanoporous membrane slide

In order to compare multiple sets of microarray data, we first attempted to increase the lifespan of the nanoporous membranes. The degree to which a nanoporous membrane binds and retains arrayed DNA after subsequent stripping cycles is an important economic aspect of the membrane array utility. Hence, the

number of membrane stripping and re-probing cycles that can be performed without significant decreases in hybridisation signal quality is one of the crucial selection criteria when considering a membrane substrate for high-throughput applications.

One imperative objective of our work was the re-usability of the DNA-arrays. In this study, Pall VIVID nanoporous membrane slides could be successfully stripped and re-probed for at least 19 times without significant loss of signal intensity (Fig. 28). The slides can be stored at room temperature and used for at least 6 months, without significant reduction in signal quality (data not shown). Another important parameter for high-throughput applications is the cycling time when re-using the slides. We were able to shorten the hybridisation protocol below 60 min, without losing efficiency of the oligoprobe binding process. The hybridisation step itself is a very rapid process (30 min here vs. 3–16 hours in the classical radioactive hybridisation approach). Furthermore, the solid slide format is an important factor for subsequent automation of the process.

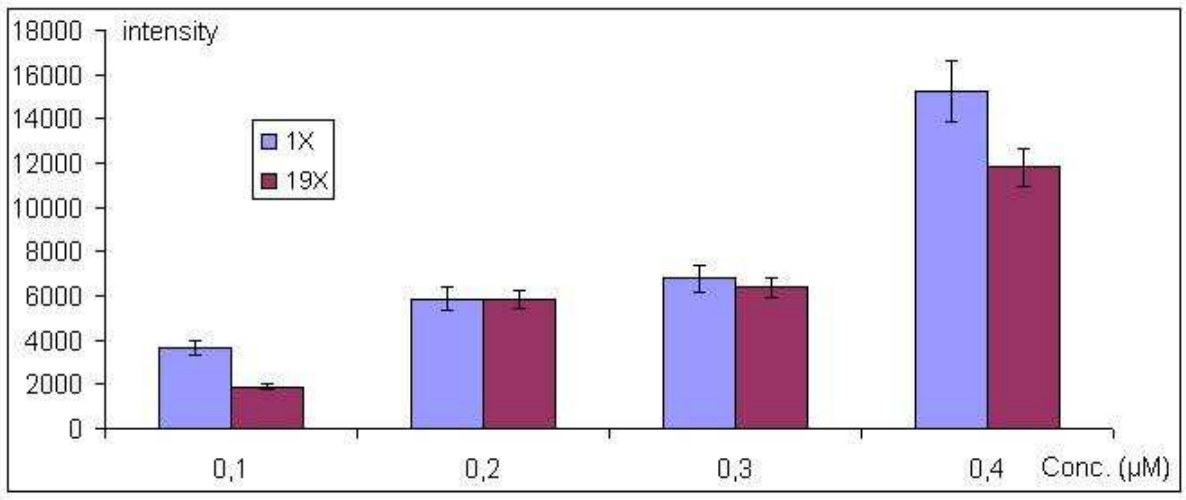


Figure 28. Re-utilisation of the nanoporous membrane slide. 100 droplets (40 nl) of synthetic ssDNA (DNA-10, sequence in Table 4) at concentrations of 0.1–0.4 µM was spotted onto the membrane slide and then hybridised with the full-match OP-02L (sequence shown in Table 1). The signal intensities following hybridisation with the same membrane for single stripping (indicated as 1X) and for 19 stripping rounds (indicated as 19X) are presented. Data were plotted as a mean and SD from three independent experiments. Fluorescence intensity was presented in arbitrary units as described in Methods.

7.2.7 Verification of OFP

For a model OFP validation, 8 genomic DNA clones (D20, G11, H20, I09, K09, N05, P06 and O20) with known sequence were selected from a genomic library (Radelof et al. 1998). After PCR reaction and ethanol-precipitation, the concentration of the eight products were adjusted to 0.4 µM and dispensed in duplicates onto nanoporous membrane as described above. Four LNA-modified 7-mer oligoprobes (OP-1L, OP-4L, OP-5L and OP-6L) were applied to subsequent hybridisation.

Fig. 29 and Fig. 30 show an example of sequence-specific signatures for these eight clones with four oligoprobes. Since the dynamic range of signal varies (e.g. OP-01L vs. OP-05L in Fig. 29 and Fig. 30), there is a demand for setting a threshold to distinguish between 'positive' and 'negative' hybridisations. Nevertheless, the positive signals are clearly distinguishable from the negative ones.

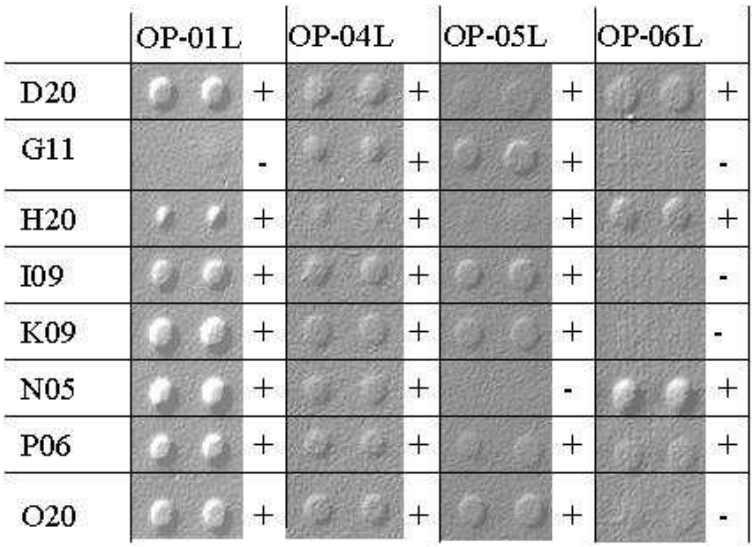


Figure 29. Image of model OFP for eight dsDNA clones using four LNA-modified oligoprobes. 100 droplets (40 nl) of eight dsDNAs, amplified from clones D20, G11, H20, I09, K09, N05, P06 and O20, at a concentration of 0.4 µM, were spotted in duplicates onto nanoporous membranes, denatured and hybridised with four LNA-modified 7-mer oligoprobes (OP-1L, OP-4L, OP-5L and OP-6L). Sequence match between the oligoprobe and the dsDNA was indicated as '+', and no match as '–'. The images were acquired by CCD camera after a single washing of the membranes.

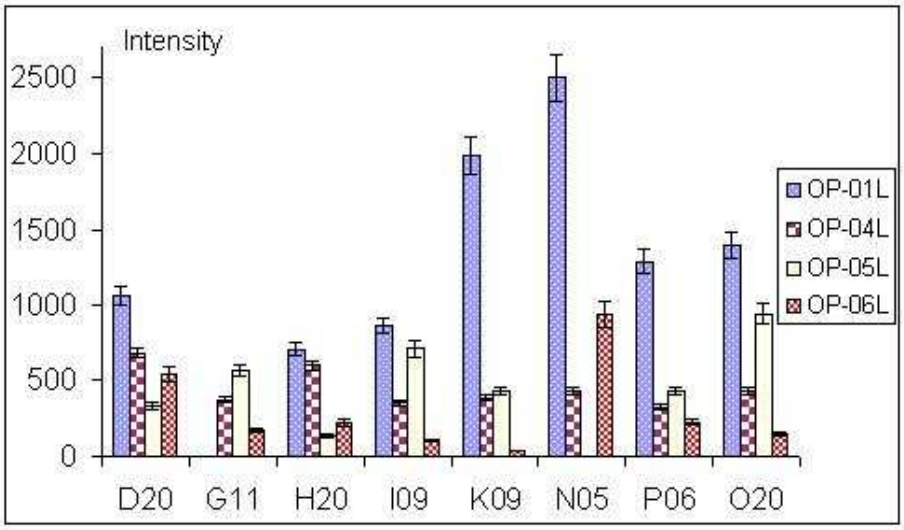


Figure 30. Quantification of hybridisation signal intensities of the model OFP (from Fig. 29) for eight dsDNA clones using four LNA oligoprobes. The background was subtracted. Data were plotted as a mean and SD from three independent experiments. Fluorescence intensity was presented in arbitrary units as described in Methods.

7.2.8 OFP using medium array density

For further verification, 66 sequence-verified genomic DNA clones with a size range of 250 bp to 2 kb were randomly selected from the genomic library constructed previously (Radelof et al. 1998). The selected genomic DNA clones have overlapping sequences and are devoid of repeat regions, and can be grouped into different sequence-clusters.

The nanoporous membrane was divided into 3 fields (Field-1, -2 and –3, shown in Fig. 31A). Different volumes of PCR products of 66 genomic DNA clones were dispensed onto the three fields, so that it could be determined at how low concentration of dsDNAs the hybridisation signals could be detected. In our case, Field 1 to 3 corresponded to 10 nl, 20 nl and 40 nl of PCR products, respectively. In each field 96 samples were spotted in quadruplicate (each letter on the left side of Fig. 31C stands for a four time duplicates in column), thus each field contained 384 spots and the whole membrane 1, 152 spots. The spot distance was 800 µm.

The purified dsDNAs from the 66 genomic DNA clones at concentration of 0.1 µM were spotted in column 1-11 as shown in Fig. 31B. After dispensing and consequent blocking treatment, the nanoporous membranes were hybridised with the 26 selected 7-mer LNA-modified oligoprobes. In our experimental set-up, we found

that with 20 nl (Field 2 in Fig. 31A), distinct and reproducible hybridisation signals could be obtained. An example of that is shown in Fig. 31B, where specific hybridisation signals from OP-15L could be observed.

The hybridisation signals and background vary across the hybridisations with different oligoprobes, mostly due to the different hybridisation performance of these oligoprobes (e.g., OP-15L, -25L and –26L) with difference sequences and LNA modifications. Furthermore, according to the experimental steps listed in Methods there are other sources of signal fluctuations including fluctuations in target and array preparation, hybridisation process, background effects and image processing. The non-specific background is of homogenous nature across the images (Fig. 31B), instead of being restricted to the hybridisation area, which means no significant deterioration to the subsequent data analysis.

7.2.9 Specificity of hybridisation

DNA microarray is a powerful technology for parallel high-throughput analysis of gene expression and genotyping. Hybridisations specificity should discriminate between samples containing duplexes with a perfect match and those containing mismatched base. The most favourable conditions for successful discrimination in hybridisation experiments are low washing temperatures, long washing times and an excess of initial hybrids (Drmanac et al. 1990). Moreover LNAs can be used to improve the design of microarray oligoprobes because they allow application of shorter and therefore better discriminating sequences.

In this setting, two 7-mer LNA-modified oligos (OP-24L, Cy5-gTcTgg**a** and OP-25L, Cy5-gTcTgg**c**) with one end base difference (indicated in red and bold) were designed for the application of genotyping discrimination. Fig. 32 shows the hybridisation images of these two oligoprobes with 36 of the 66 purified PCR products. The spotting set-up indicated in green color, described the theoretical matching conditions between the two oligoprobes and the 36 dsDNAs. The experimental hybridisation signals very well conform to the in-silica predictions. The hybridisation signal intensities reflected accurate discrimination between full-match and single-base mismatch oligoprobe/target pairs.

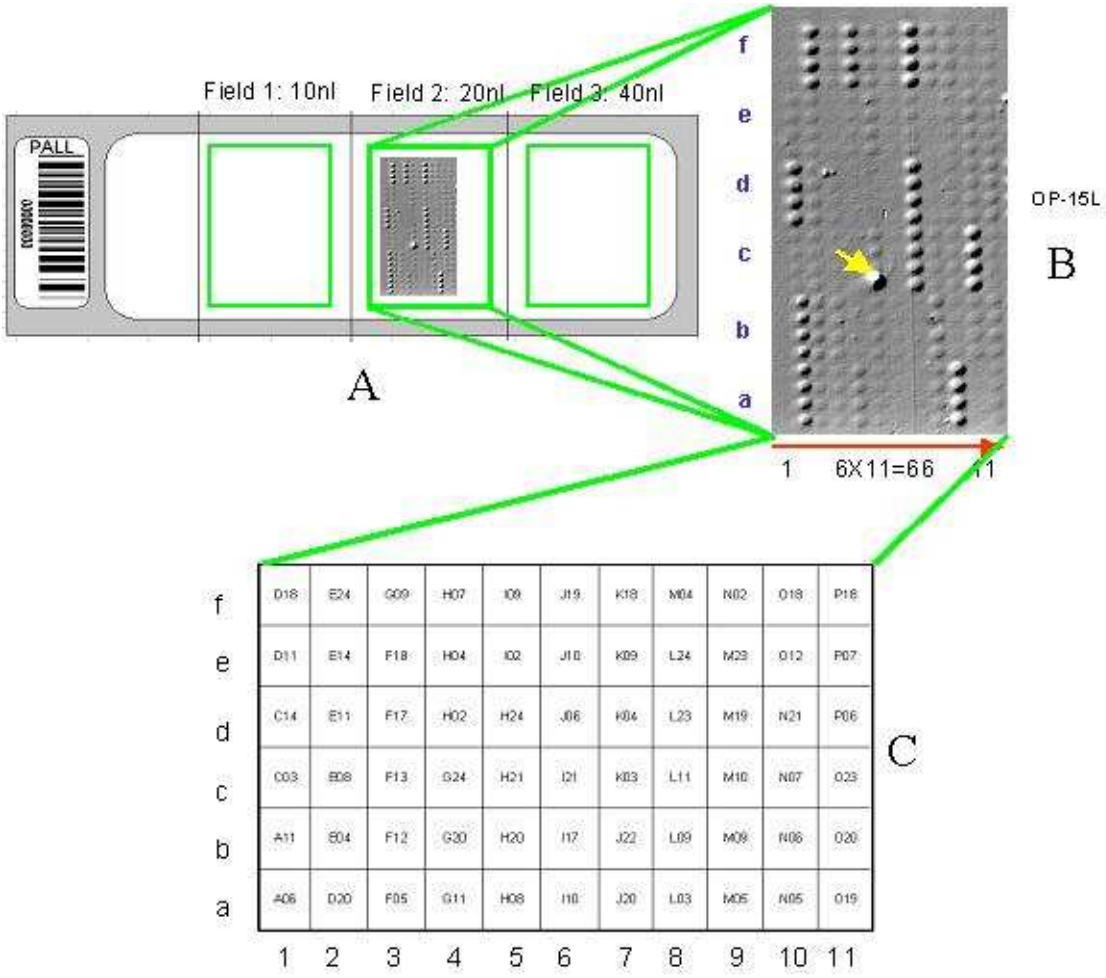


Figure 31. dsDNA dispensing set-up with an example of hybridisation image. **A**. Three fields on the nanoporous membrane with different dilution series of the spotted DNA (10 nl, 20 nl and 40 nl in corresponding fields). **B**. A hybridisation image between the DNA samples spotted on field 2 and oligoprobe OP-15L. The single signal indicated by an arrow comes from Cy-5 dye contamination. **C**. The arrangement of the DNA samples in the single field. The purified dsDNAs from the 66 genomic DNA clones at the concentration of 0.1 µM were spotted in column 1-11.

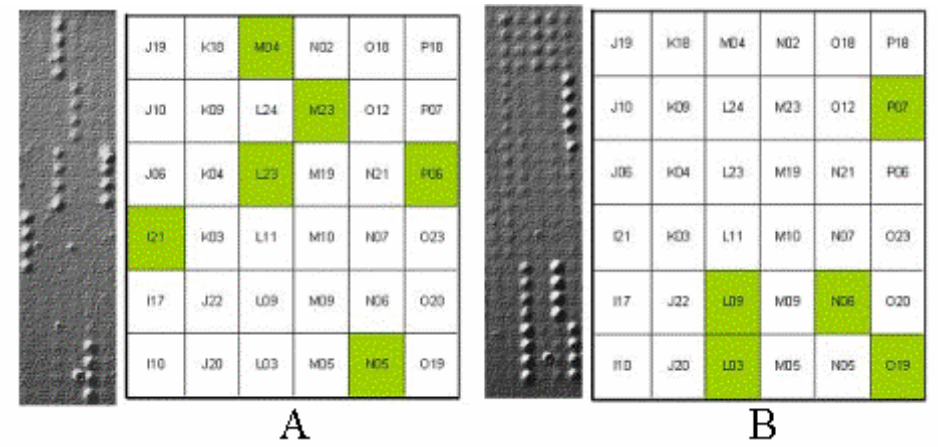


Figure 32. Specificity of hybridisation using two 7-mer LNA-modified oligoprobes with one nucleotide difference at 3'-end. **A**. The hybridisation image on the left panel and matching conditions on the right panel of oligoprobe OP-24L (sequence: gTcTgga) and selected 36 dsDNAs. The dsDNAs in green indicate full match to OP-24L. **B**. The hybridisation image on the left panel and matching conditions on the right panel of oligoprobe OP-25L (sequence: gTcTggc) and selected 36 dsDNAs (same as in A.). The dsDNAs in green indicate full match to OP-25L.

7.2.10 Manufacturing DNA microarrays from unpurified PCR products

PCR and spotting of the amplified DNA fragments form the work-intensive core of microarray production. The time and cost involved are considerable. And the purification process itself needs to be well established and quality controlled in order to produce reliable results. In order to circumvent this problem, we investigated the difference of hybridisation when using purified PCR products and direct PCR products without purification from the same clones. We tried to evaluate the possibility of usage of unpurified PCR product in this novel hybridisation system. Fig. 33A shows the hybridisation image pairs of OP-04L, -15L and –17L with purified matching PCR products and the unpurified counterparts. In general, signal intensity on arrays made of unpurified PCR products is about 48-93% of the intensity obtained with the respective purified molecules (Fig. 33B), which is slightly lower than the previously reported (Diehl et al. 2002).

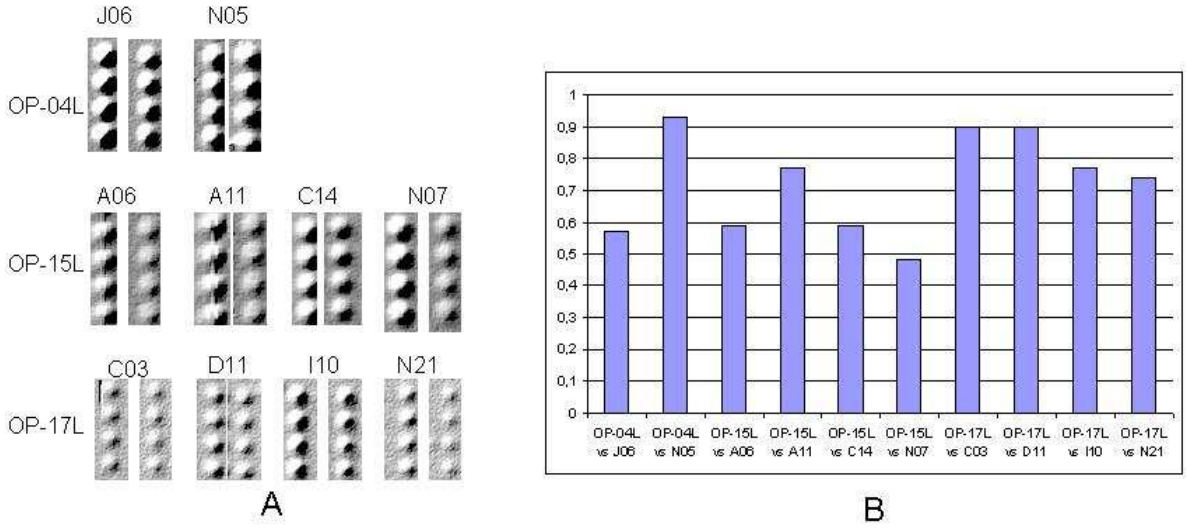


Figure 33. Comparison of hybridisation signals between purified and unpurified PCR products using the same LNA-modified oligoprobes. **A.** Hybridisation images of selected 7-mer LNA-modified oligoprobes (OP-04L, -15L and –17L) with their matching dsDNAs after or without purification. **B.** The percentages of the hybridisation signal intensities from unpurified dsDNA targets comparing to purified ones as references.

7.2.11 Microarray image evaluation

The microarray image homogeneity is crucial in the process of spot detection and quantification. Image analysis is a major part of data processing for array hybridisation experiments. Only very few images are of perfect quality. For example from the hybridisation images of oligoprobe OP-15L (Fig. 31B), dye contamination could be observed. In this case the signals from contamination are not in the positions of the hybridisation signals so they could be manually corrected it, thus enabling further analysis process. In case where contamination signals are overlapped with the hybridisation signals, especially where more than one signal of the quadruplicate of one sample is overlaid with the contamination signal, the down-stream analysis could not be performed. Fig. 34 shows few examples of image problems.

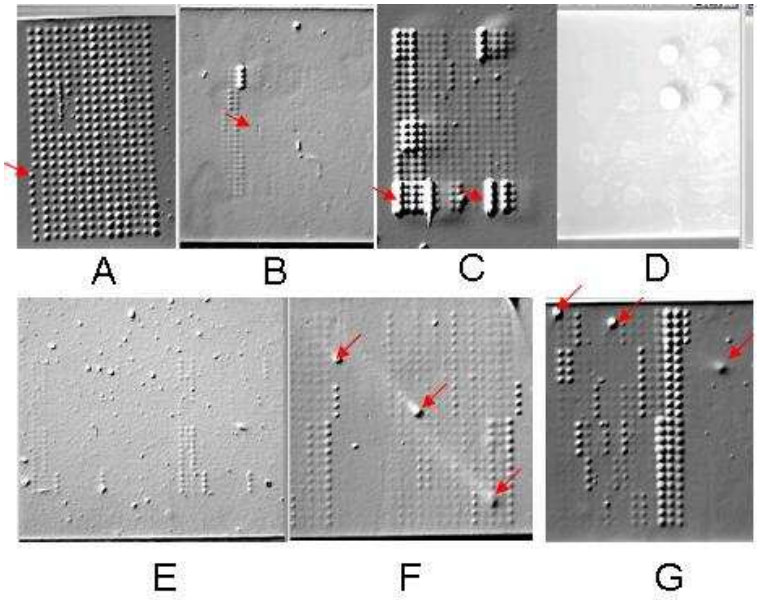


Figure 34. Image problems. **A.** Spot problem. The spots which should in one column are not in the right positions, which will result in difficulty for automatic image analysis **B.** physic scratch of the nylon membrane. **C.** Overshining due to the very high signals from the neighbouring spots or due to the too close spots. **D.** High background probably resulting from the improper blocking procedure. **E.** Chemical precipitation or dirty slide. **F.** Comet-tails leading to disastrous artificial signals. **G.** Dye or oligoprobe contamination which may bring some trouble for signal evaluation.

7.2.12 Receiver operating characteristic (ROC) analysis

In order to evaluate the performance of the LNA-modified oligoprobe set in hybridisation experiments, we compared the experimentally derived signals from genomic DNAs containing the oligoprobe sequence (or its reverse complementary sequence) with genomic DNAs that without containing the oligoprobe sequences.

It is important for this study that the oligoprobe sequences could be randomly selected, since it was aimed to avoid individual oligoprobe bias resulting from unusually high or low matching rates. The matching rate of an oligoprobe is dependent on the length of the oligoprobe and the average length of the genomic DNAs under consideration. The practical performance of the hybridisation experiment is a trade-off between oligoprobe length (shorter oligoprobes match to more sequences) and stability of the match (shorter oligoprobes show less stability). The 66 human genomic DNA sequences have a mean length of 1,322 bp with standard deviation of 239 bp. Theoretically, the expected number of matches of a specific 7-

mer oligoprobe within a set of N sequences of that length is equal to N*0.15, i.e. the oligoprobe will match to 15% of the clones (Herwig et al. 2000). The hit rates of the selected LNA-modified oligoprobes in our study ranged from 4.5% to 60.6% with a median hit rate of 19% what comes close to a randomly selected oligoprobe. Therefore, we concluded that there is no sequence bias in our data set. Individual characteristics of the LNA-modified oligoprobes are shown in Table 1.

A receiver operating characteristic (ROC) is a graphical plot of the sensitivity versus (1 - specificity) for a binary classifier system as its discrimination threshold is varied. The ROC can also be represented equivalently by plotting the fraction of true positives (TP = true positive) vs. the fraction of false positives (FP = false positive). ROC curves are shown in Fig. 35. For each hybridisation experiment we computed the average signal for each genomic DNA across its four replicate values. This signal was used as a threshold that judges the amount of false positive (X-axis) and true positive (Y-axis) genomic DNAs. True positive signals were defined as signals above the threshold derived from a genomic DNA that contains the oligoprobe sequence. False positive signals are those that are above the threshold derived from a genomic DNA that does not contain the probe sequence. In our case, the AUC measured the ability of the individual oligoprobe to distinct signal from noise. Most LNA-modified oligoprobes showed good performance in separation of the signal from noise indicated by the value for the integral of the ROC curve (green line, maximum equal to 1, Fig. 35). However, for some LNA-modified oligoprobes (OP-04L, OP-09L and -18L), the performance droped below the 0.9 limit. The performance measure used here was the Area Under the ROC Curve (AUC) indicated by the green lines. In classification theory, this measure is usually known as discrimination and can be interpreted as the ability of the classifier to correctly classify an unknown sample. The lack of performance of these oligoprobes cannot be explained straightforwardly by the individual oligoprobe statistics (Table 1), neither by an unusually high (or low) matching rate or position of LNA modification in the oligoprobes, nor any differences with respect to the melting temperature or the GC content.

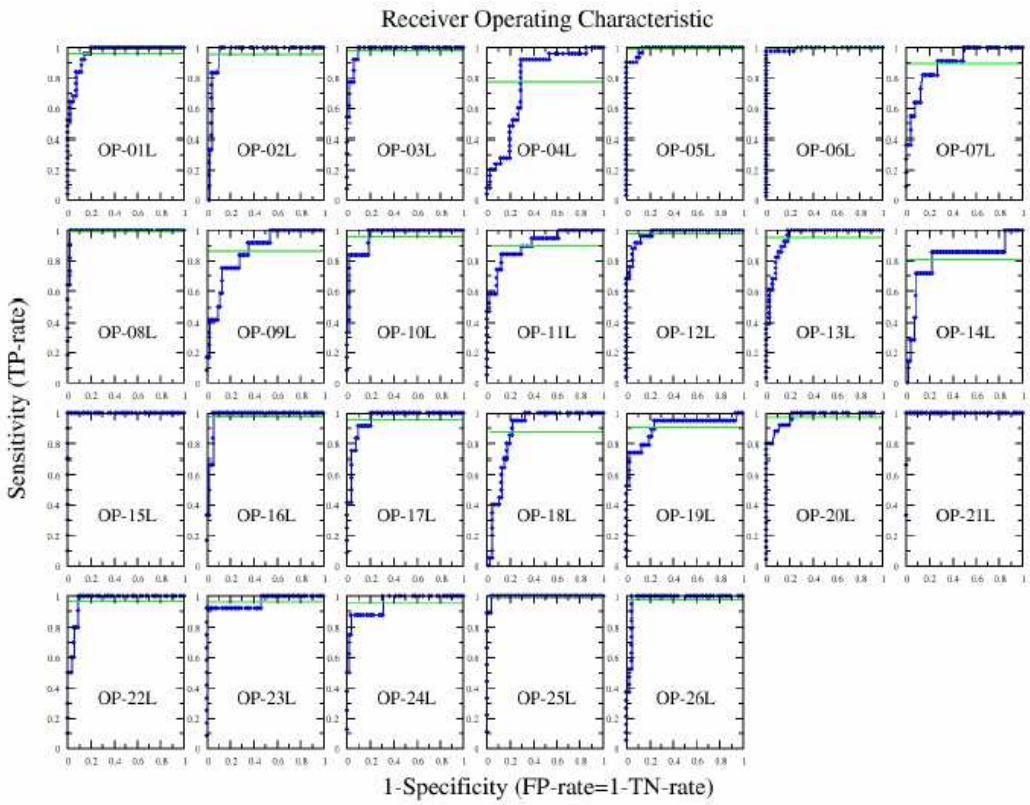


Figure 35. Receiver Operating Characteristic (ROC) of all 26 LNA-modified oligoprobe hybridisation experiments. X-axis shows the false positive rate, Y-axis the true positive rate. The green lines indicate the area under the ROC curve (AUC, maximum equal to one). TP: true positive, FP: false positive, TN: true negative.

7.2.13 Hybridisation comparison of LNA- and non-modified oligoprobes

In order to further demonstrate the gain in hybridisation specificity with respect to LNA modification of the oligoprobe we compared the ROC curves for selected non-modified oligoprobes with the same sequences of the corresponding LNA-modified oligoprobes. The ROC curves of all the six pairs of non-modified and LNA-modified oligoprobes are shown in Fig. 36. All LNA-modified oligoprobes showed an increase in performance compared to the respective DNA sequence. Moreover, for only one of the non-modified oligoprobes (OP-15), the ROC value exceeded the quality level (ROC curve integral, green line) of 0.9, whereas ROC value for all oligoprobes containing LNAs was above 0.9. A possible reason for the increase of performance is the increase in melting temperature (Tm) when switching from non-modified to LNA (comparison in Table 1). The average increase for the six oligoprobes under consideration was 11.0 ℃ with a range from 5 to 20℃.

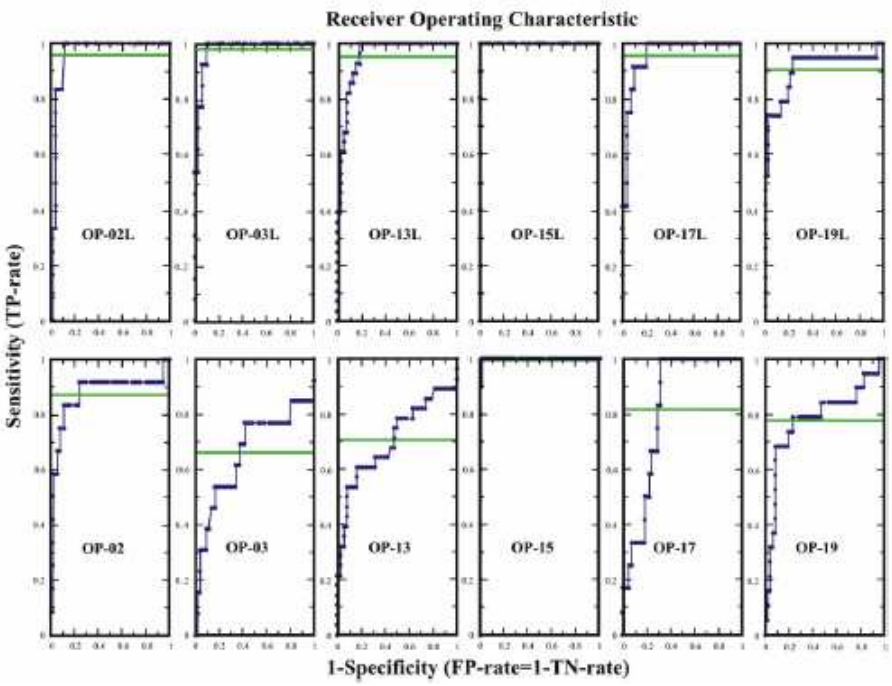


Figure 36. ROC curve comparison of LNA-modified (upper panel) with non-modified oligoprobe hybridisations (lower panel) using dsDNAs as templates. X-axis shows the false positive rate, Y-axis the true positive rate. The green lines indicate the area under the ROC curve (maximum equal to one). TP: true positive, FP: false positive, TN: true negative.

Since the CVs (coefficients of variation) compute the ratios of the standard deviations and the means across the replicates, they can be used to compare the reproducibility in the different hybridisation experiments. CVs range from zero (perfect reproducibility) to higher values. CVs above one indicate complete irreproducibility of hybridisation signals.

Higher thermodynamic stability of duplex pairing between oligoprobe and genomic DNA sequence was also demonstrated through the reproducibility of the hybridisation signals. Fig. 37 shows the histograms of the CVs derived from the four replicates signals for each genomic DNA, in each hybridisation experiment for all 26 LNA-modified oligoprobes and all selected 6 non-modified oligoprobes. The blue line shows the CVs for all 26 LNA hybridisations and the green histogram shows the CVs for all six non-modified oligoprobe hybridisations. While the peak of the LNA histogram (0.034) and the DNA histogram (0.045) are comparable, the histograms of the non-modified oligoprobes show a much broader tail indicating a high proportion of signals with low reproducibility.

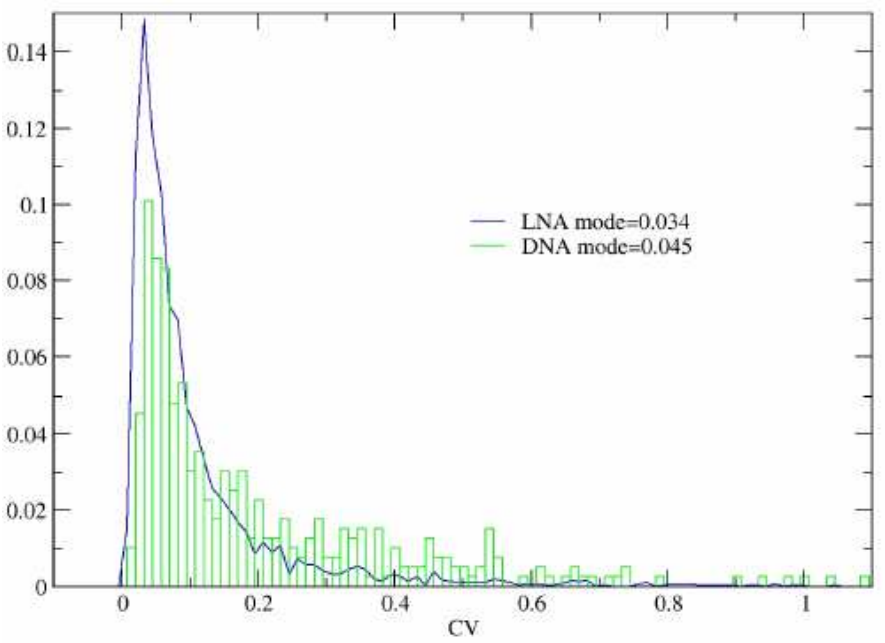


Figure 37. Histogram of coefficients of variations (CVs). Blue line shows the histogram of CVs derived from the four replicate genomic DNA signals of all LNA hybridisation experiments. The green histogram shows the CVs derived from all non-modified hybridisation experiments. X-axis shows the CV (standard deviation divided by mean), Y-axis the relative frequency.

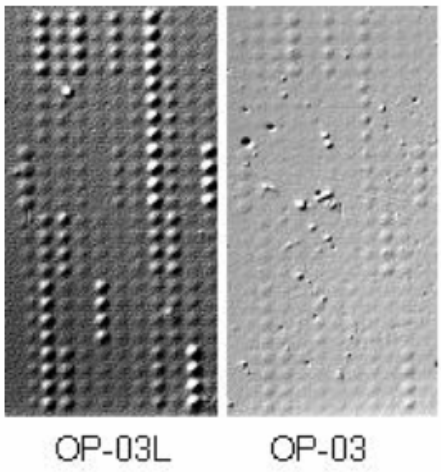


Fig. 38. Hybridisation images of LNA-modified and non-modified oligoprobes

Fig. 38 as an example presents the hybridisation images of a pair of non-modified and LNA-modified oligoprobes (OP-03 and OP-03L) for comparison. Based on the image comparison we conclude that LNA-modified oligoprobes confer more significant hybridisation signals. In conclusion modification of DNA oligoprobes with LNA leads to higher reproducibility and stability of hybridisation.

7.2.14 Hybridisation pattern of the model clone set

Clone clustering based on OFP is a very efficient tool for library normalisation. Because the raw hybridisation intensities are influenced by a number of factors such as specific activity of the oligoprobe, yields of PCR and variation in dispensing, the intensity levels have to be normalised. Normalised intensities are then used to compute fingerprints, which are clustered using an algorithm (Herwig et al. 1999) that groups together clones with similar fingerprints. The basis of this approach is that the pattern of hybridisation of oligoprobes to the cDNA reflects its sequence and can therefore be used as a `fingerprint' for its identification. Clustering by defined set of oligoprobes has been applied previously in order to identify the set of genes and their abundance in large cDNA libraries (Herwig et al. 1999 and 2002).

The overall distribution of LNA-modified oligoprobes or their reverse complementary sequences in the genomic DNA clone set of 66 is presented in Fig. 39. It can be observed that the genomic DNA sequences can be partitioned in several fragments of overlapping sequence clusters (indicated by black circles) of different sizes ranging from 1 to 15 members. In order to test the practical performance of LNA hybridisation experiments in recognising similarity of DNA sequences derived from a shotgun genomic library (Radelof et al. 1998), we used a clustering approach. The goal of the clustering approach is to reconstruct this partition using the fingerprints obtained by the analysis of the signal intensities across the hybridisation of the 66 clones with the 26 LNA-modified oligoprobes. For clustering we used a hierarchical clustering method (Herwig et al. 1999) with Pearson Correlation as a pairwise similarity measure and average linkage as an update rule applying the J-Express Pro V2.7 software package. The results shown in Fig. 40 indicate shows that the OFP approach reproduces the sequence clusters very well. Separated and highly overlapping clusters such as sequence clusters 3, 6, and 8 are recovered accurately, and clusters 4, 5 and 7 are reconstructed with only one clone excluded from the group for each cluster. Sequence clusters that are more distinct such as cluster 1 and 2 are spread in subgroups within the histogram.

Our results show that LNA-modified oligoprobes can be used effectively to reveal sequence similarity of genomic DNA sequences and thus, to characterise the content of unknown genomic DNA libraries. In practice, the number of oligoprobes is typically much smaller compared to the number of genomic DNAs (in the order of hundreds of thousands of sequences). The number of LNA-modified oligoprobes

needed to characterise such libraries can be extrapolated from OFP experiments using DNA oligoprobes. In a previous study approximately 250 octamer oligoprobes were used to derive a low-redundant set of sequences from approximately 250,000 sugar beet cDNAs (Herwig et al. 2002). With this number of oligoprobes it is possible to identify even small sequence clusters (corresponding to lowly expressed genes) with high accuracy. Since hexamer oligoprobes hybridise with much larger fractions of sequences than octamers and since the stability of LNA-modified oligoprobes is much higher than that of DNA oligoprobes this number is rather an upper bound and the effective number of LNA-modified oligoprobes needed to characterise such large number of sequences would be in the range of 100-150.

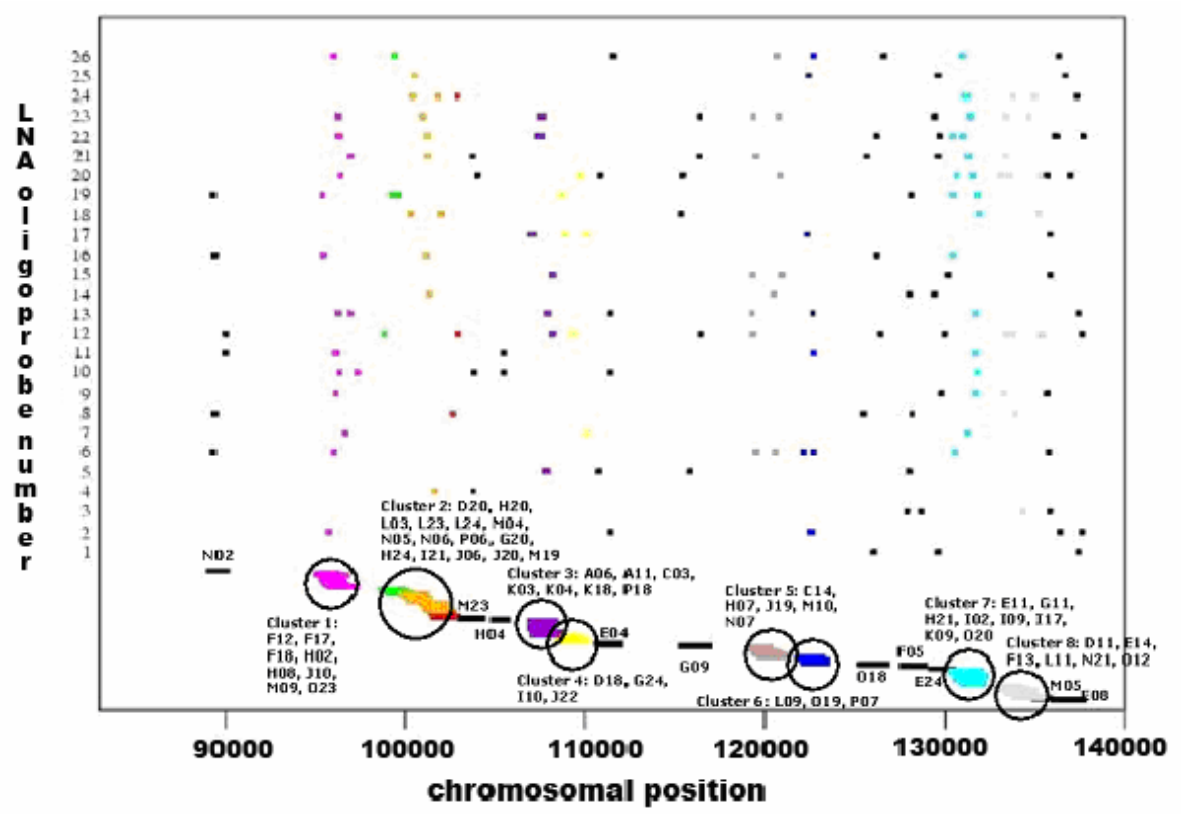


Figure 39. Oligoprobe positions in the human chromosome Xq28 cosmid clone range 89182-137664, which comprises all the 66 genomic DNA clones used in this study. Clusters of overlapping sequences are marked with circles.

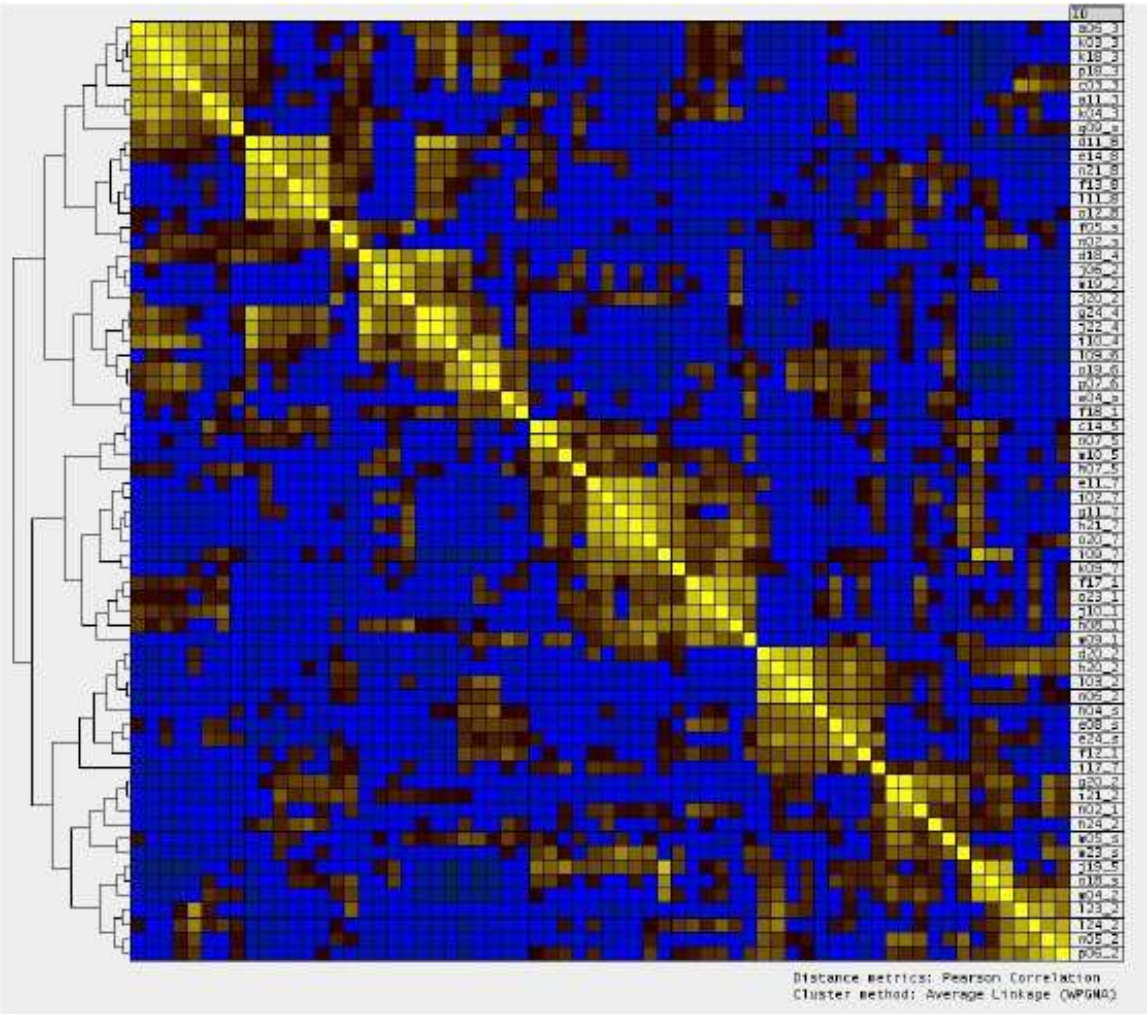


Figure 40. Hierarchical clustering of genomic DNA clones using Pearson correlation as similarity measure and average linkage as an update rule (J-Express Pro V2.7, www.molmine.com). Similarity of the genomic DNA clones is visualised with coloured boxes (yellow = high similarity, blue = low similarity). The panel on the right shows the genomic DNA clone identifier and the sequence cluster they are associated with (s = singletons).