Description of the main science and technology results

This final S&T report contains a summary description of the work that has been carried out during the Project Periods I - III of the Project. This report has been structured as follows:

(i) Each work package is described independently

(ii) Each work package contains brief overview, key aspects of the work and results

This will enable the reader to get an overview of the progress in the work package.

The Project has been structured into seven work packages

WP1 – Coordination (MAN)

WP2 - Automated and manual binary conversion (RTD)

WP3 – Addressable binary conversion (RTD)

WP4 – Binary DNA readout on microscope slides (RTD)

WP5 – Binary DNA readout on optical disk (RTD)

WP6 – Development and validation of an optical disk based DNA sequencing system (RTD)

WP7 – Dissemination, exploitation and IPR management (MAN)

The relation between the work packages is schematically described in Figure 1. The three central components of the Digital sequencing system are the WP2 where the DNA is converted into a binary representation (for example A → (0,0)), WP5 where the optical disk, the Biorecord, is developed; and finally WP6 where the sequencing system is developed and demonstrated. WP4 provides a stepping stone on the route by enabling use of glass slides, an established platform, for development work while the Biorecord and Sequencing system are in development. WP3 provides an indexing technology enabling in silico assembly of short DNA reads into longer, virtual DNA reads.

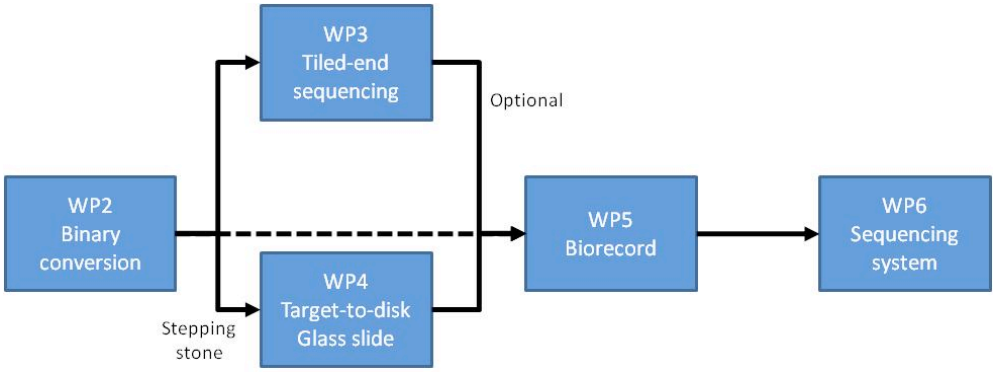


Figure 1 – Schematic representation of the work packages. Dashed line refers to a direct, but challenging, route.

The goal of the Project has been to develop a high-capacity, low-cost Digital Sequencer with ability to provide sequence data. Two target applications have been envisioned (i) global shot-gun sequencing approach using RNA or DNA as starting material (ii) target sequencing of diagnostic relevance (companion diagnostics). Important for both products is also the ability to be run by operators without significant previous training or experience on laboratory techniques; this also contributes to making sequencing a part of everyday life and enables sequencing to be carried out in settings where it is not available today. This product profile has been under continued monitoring, and has been changed to reflect the development in the DNA sequencing field.

The project has evolved along with remarkable progress in the next generation sequencing field with clear segmentation of the entire sequencing market. Among the key events during the most recent years are that one of the major manufacturers in this field, Roche, with their 454 platform has announced their discontinuation with 454 NGS technology. While Illumina, has expanded its market share even more and has presented their high-throughput 1000 USD per human genome HiSeqXTen instruments (requiring > 10 MUSD investment). Life Technologies is re-focusing, current the use of

their Ion Torrent platform is geared towards gene panels and other medium throughput applications, while the Pacific Biosystems single molecule systems (1 MUSD investment) is more addressing long read niche without high throughput ambitions.

Central and essential for the current Project is the ability to provide a Sequencer at a low cost, thereby fulfilling the aim of making DNA sequencing a part of the everyday life. This is a shared ambition with the efforts in nanopore sequencing that also has been evolved significantly over the most recent years but needs to mature significantly to be able to provide the required accuracy in for example companion diagnostics of cancer gene mutations.

Automated and manual binary conversion (WP2)

Overview of the sequencing technology

The aim of the work carried is to establish a conversion protocol that transforms DNA into a binary format. Conversion enables DNA to be read out on various DNA sequencing platforms, including nanopores and the optical disk-platform developed in this project.

In a converted DNA sequence the four bases A, C, G and T are replaced by binary tags corresponding to (0,0), (0,1), (1,0), and (1,1). The binary value can be defined through various mechanisms, including different sequence composition (for example, 0 = AAA, 1 = TTT), length (0 = 8 nt, 1 = 12 nt), presence of a label (0 = Cy5, 1 = Cy3, or alternatively 0 = no biotin, 1 = biotin), or through various other means. Preferably, the conversion protocol is flexible in terms of the approach chosen, thereby facilitating the use of protocol on more than one intended sequence readout platform.

A

	Odd	Even
A	1	1
T	1	0
G	0	0
C	0	1

B

```
1 2 3  4 5 6  7 8 9  (nt position)
A G T  G G C  A G T  (sequence)
110010 000001 110010 (binary code)
OEOEOE OEOEOE OEOEOE (odd/even bit)
```

Figure 1 – A) Example of a coding scheme of the binary sequencing concept demonstrating that the DNA sequence can be decoded by a combination of two bits (0,1). Here colour is used to denote bits with value '1'. The different combinations of '0' and '1' can be assigned in any way to the four nucleotides. B) Example conversion of a 9-nt sequence read (grouped into sets of three for clarity).

The strategy has been to develop several conversion protocols in parallel to both reduce overall risk, as well as to ensure compatibility with different readout platforms. The first strategy is based on further improvement of the design polymer concept (Lexow (Sequencing method using magnifying tags, EP 1141399 B1)) and automation of the conversion process on a lab robot. The second strategy utilises a binary tag concept, where the target to be sequenced is split into multiple tag molecules, each carrying a label indicating its binary value. At the molecular level the two approaches express similarities in for example the enzymes and molecular biology steps. On the other hand, the two approaches are different in that the former generates a long, physically-linked molecule (the design polymer), while the latter generates a large number of tags that are linked to each other through physical or temporal compartimentalisation.

The positioning step is the key step of the protocol. In the *majority* of the sequencing protocols used today, the sequencing is carried out in an iterative manner; nucleotides are identified one-by-one using the previously identified nucleotide as a starting point for the next step. The 454 chemistry (Roche) uses pyrosequencing chemistry, Illumina's technology uses reversible terminators on the labelled nucleotides, SOLiD technology (Life Technologies) uses chemically cleavable adapters. The binary tag protocol is conceptually different from the three mentioned above; here the identification ('readout') of all the nucleotides to be sequenced are carried out in a parallel manner in steps 2) and 3) of the protocol. To achieve this, the amplified template molecule is 'split'/indexed into a number of aliquots, where each aliquot will interrogate a unique bit ('position') of the sequence read. Consequently, the iterative approaches require multiple readout steps, while the Digital Sequencing

approach in its simplest form only requires one readout step. The benefit of this is a rapid sequencing reaction, which is significantly faster than most other currently available sequencing technologies.

Key aspects of work

(i) Design polymer-based conversion

The concept of binary conversion of target sequences using the Design Polymer concept has previously been successfully demonstrated using the lambda genome as starting material. Those efforts were based on a PCR-based procedure that is not well suited for automation and high throughput sample preparations. This work of this WP aims to address these at two levels. Firstly, develop a non-PCR based approach for binary conversion, and secondly develop an automated procedure.

(ii) Binary tag-based conversion

On the binary tag-based conversion protocol a simplified proof-of-concept demonstration had been achieved prior to the start of the Digital Sequencing project. However, the detailed optimisation and troubleshooting of the protocol for removal of background signals observed in the results was included as starting point for this work in the Project.

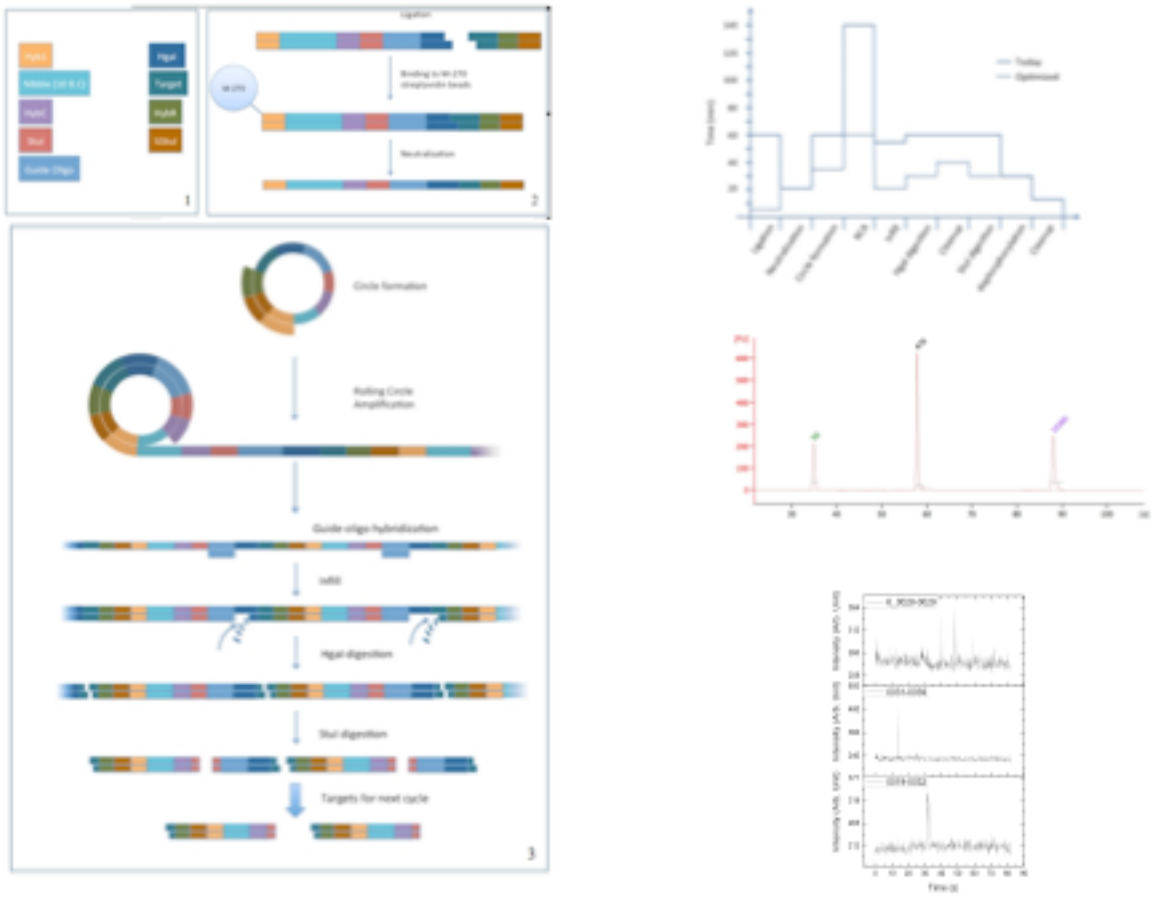


Figure 2 - Schematic overview. Left panel: 1 Key to the colour-coded segments in 2 and 3. (2) The capture adapter, containing the binary information is ligated to the target sequence. The ligated fragment is then immobilized through biotin-streptavidin interaction on M-270 streptavidin covered magnetic beads and the biotinylated strand is removed by treatment with NaOH. (3) The productive strand is amplified through RCA producing a long single stranded fragment of concatemers to which guide oligonucleotides are hybridized. The product is made double stranded by filling in the gaps between the oligonucleotides. The long concatemer is digested with HgaI, which cuts into the target sequence, and StuI, which removes the part that is not needed for the next cycles. Right panel (top): The automated binary conversion protocol shown with the time required for each step for the current version and a potential optimized version. Right panel (middle): Electropherogram of the PCR products from the design polymer of 450 bp segment. Right panel (down): Extracted emission intensities as a function of time from three different nanopores in a DNA translocation experiment using a design polymer. This demonstrates single-molecule sensitivity of the optical system

Results

Using the Design Polymer technology, we have developed an integrated and automated system for conversion of target genomic material using type II restriction enzymes, ligation and Phi 29 amplification (Figure 2). We addressed low efficiency in enzymatic steps by reducing the number of steps by using 5 bp sticky ends instead of 2 bp. All the steps of the conversion was automated on a Magnatrix robotics workstation and sample preparations of DNA and RNA was developed in parallel to be able to feed into the conversion workflow (Borgström et al. PLoS One. 2011 Apr 27;6(4):e19119; Stranneheim et al , PLoS One. 2011;6(7):e21910). As outlined previously, the design polymer-based approach can provide a binary conversion enabling early generations of nanopore sequencing to be brought onto market. A collaborative pilot experiment with nanopores was recently done with a design polymer (10 bp bit size; >400 bp) covering the BRAF hot spot region. The experiments were done together with Prof Jan Linnros, KTH using his expertise in solid state physics.

The Project made the strategic choice for the Project Period III to focus on the alternate binary tag based approaches (Figure 3), as these are more amenable for decentralised settings and is better suited for more targeted sequencing of clinical markers (companion diagnostics) and represent a unique market niche. The work carried out in WP 4-6 therefore presents collectively development of both design polymer and the binary tag approach. The binary tag conversion protocol based on Klenow and apyrase has been experimentally demonstrated. Estimates indicate total sequencing run time of 1-3 hours for a 50-nt read, with further potential for significant reduction. Cost estimates indicate extremely low costs due to use of natural components.

In the developed binary tag-based protocol we used two labeling mixtures to achieve the parallel read-out for odd/even signals. Here we employed a first mixture that contained biotinylated dATP and dUTP bases, the two remaining bases being non-labeled terminating dideoxy bases. In the second mixture dCTP and dUTP are labeled and the other two bases were non-labeled terminating bases. The biotin offers the possibility for both fluorescens detection through labelled streptavidin binding and more importantly bead detection using streptavidin coated magnetic beads.

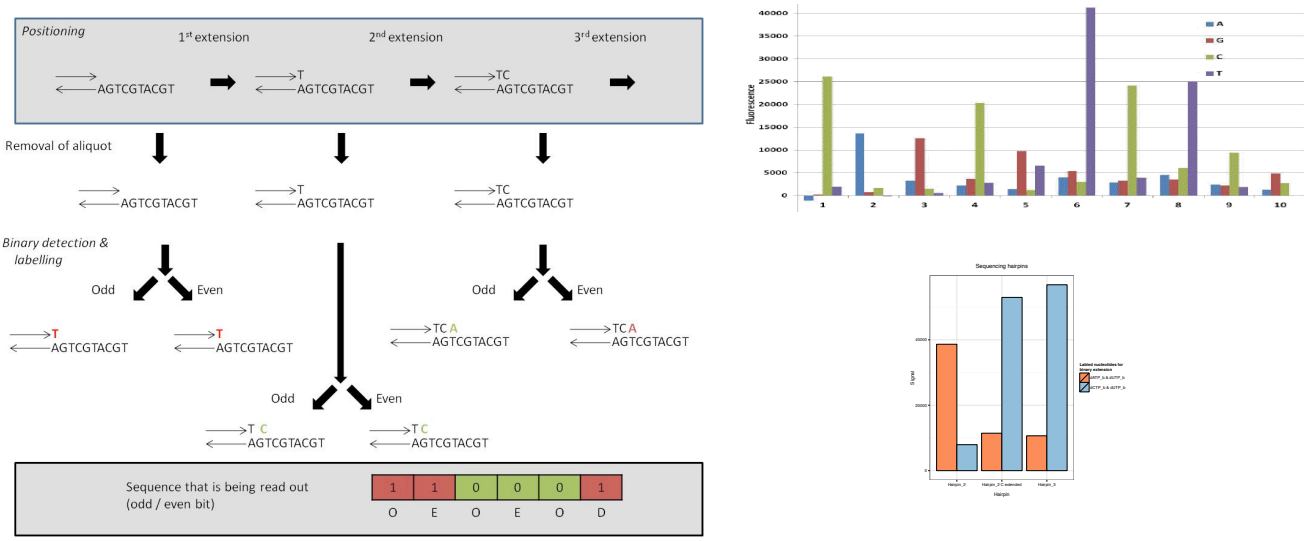


Figure 3 - Schematic principle of the extension-based binary tag conversion protocols. The numbers on the left refer to the unit operations of the protocol. Red and green represent binary units of value 1 and 0, respectively. Coding scheme (Figure 2.1) used is arbitrary. The single-base extensions (on top) can be achieved using any of the three principles outlined in the text: reversible terminator approach, Mg-withdrawal or polymerase & apyrase-mediated approach. Here removal of an aliquot is indicated as one option. An alternative is to use different primers for the initiation of the positioning extension for each bit, and where the primers are shifted in reference to the start position of the target-of-interest (compare with figure 2.2 above where k-1 region is used to shift the IIs recognition motifs in relation to the target-of-interest). Thereby all bit positions are exposed to the same number of iterative extensions, and all reactions can be carried out in one 'tube' or compartment. (Top right panel) Demonstration of a 10-nucleotide readout using Klenow and apyrase. (Low right panel)

Single base binary extension determines the mutational status of BRAF mutation (before, extended and mutation control is visualized)

The proof-of-principle demonstration has successfully been carried out on a glass solid support for the binary tag protocol and issues with background signal from previous work in solution has been addressed (Figure 3). The transfer of this protocol to on disc and microfluidic channel systems with lyophilized reagents is in the process to be determined, as well as identifying the molecular details associated with template capture.

As indicated previously, there has been significant progress in field of sequencing technology during the project period that relates to the WP2. Illumina has recently launched their NextSeq500 instrument that provides more rapid turnaround sequencing. This is achieved by binary sequencing principles, i.e. using only two labels to deduce the DNA sequence. Time is reduced by two times as only 2 instead of 4 labels are scanned and analysed. The cost is also reduced by using fewer reagents, however the instrument cost is > 200 000 USD, i.e. significantly larger than the DVD system developed through this initiative. In terms of design polymer technology, there has been some reported progress at Stratos Inc developing a competing technology with expandamers with the target market being nanopores. Recently Roche made a significant investment in Stratos (> 100 MUSD). Of particular interest is that LVs technology share many of the key features and LV has an earlier filing of the design polymer concepts.

Addressable binary conversion (WP3)

Overview of the indexing technology

In order to increase the power we have extended the sequence analysis by adding further functionality to the molecules to be sequenced. Two molecular tags (barcodes) are added; the first tag - the sample ID tag - defines the molecular origin (i.e., which nucleic acid fragment it describes), while the second tag - the time point tag - identifies the relative position of the fragment in the original molecule. Use of the tagging approach facilitates the in silico assembly of sequencing reads, and provides means for virtually increasing the read length. This approach also enables analysis of longer contigs (3-10 kb), copy number variations or genotyping over a long distance.

The tagging protocol, the Tiled End Sequencing (TES), can also be used outside the setting of the optical disk-based sequencing. For example, the protocol is valuable in all sequencing settings where short reads are generated or where longer reads are required. The protocol can be applied to both converted and non-converted material.

The principle of the protocol is based on three features; (1) indexing of the starting molecules so that each molecule has a unique barcode; (2) exonuclease treatment of DNA fragments without affecting the identification index so that fragments are degraded over time in only one direction and (3) indexing individual time points during degradation so that discrimination between time points can be achieved. Analyses demonstrate that long sequences of very high accuracy can be obtained.

In addition, to increase the throughput of the tiled end sequencing approach, a new technique that allows vast pre-barcoding of millions of samples, prior to TES, has been developed. The method is based on serial compartmentalization using emulsion technology and enables even higher multiplexing of sample introduced into a sequencing experiment.

Part of the activities of this work package also includes an ‘enzymology sub project’. This work evaluates the choice of different enzymes for the key molecular steps that are used in the other WPs.

Key aspects of work

For any short-read DNA sequencing strategy longer read length and positional information is desired. Through indexing, difficult regions of key interest can be obtained which is addressed here. In addition, molecule indexing is combined with temporal indexing in order to sort short fragments according to time or their relative position in the template molecule is of high interest. Indexing is also a key importance for companion diagnostics to identify unique molecular events instead of so called PCR duplicates. Furthermore we have investigated, from the point of binary conversion protocols, a

set of key DNA modifying enzymes. It is important that these enzymes, which include type IIs restriction endonucleases, ligase and polymerase, possess selected key characteristics such as high efficiency and specificity in processing DNA targets present both in solution and immobilized on a surface.

Results

Indexing of DNA molecules has been achieved at several levels. It can be fairly straightforward, using only a random set of adapters added onto the target by ligation. It is clear from Figure 5 and from the next section how the indexing can be efficiently used to extend the sequencing reads. This is of particular interest for short read platforms. Furthermore, we demonstrated the possibility to perform molecule indexing of several targets in a high throughput manner. The finalised protocols are detailed in an article published by Lundin et al (2013). Again, the protocols have been automated on a liquid handling platform.

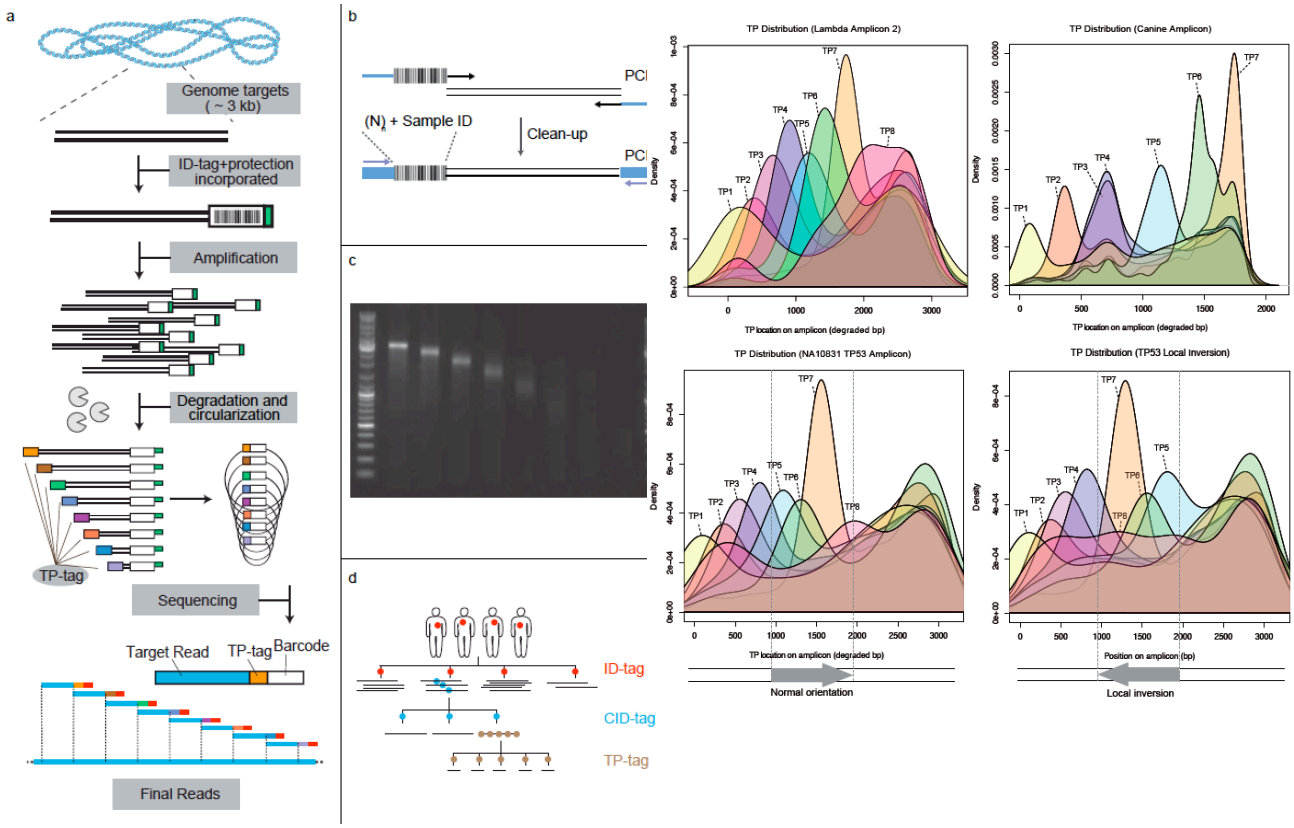


Figure 5 - Schematic description of the tiled-end sequencing protocol. Exonucleases are used to in a controlled manner to degrade prepared substrate molecules. The different degradation time points are indicated using a Position ID / TimePoint tag. Figure from Lundin et al, Scientific Reports, 2013. (Right panel) Distribution of Time Point tags for the different samples

We have achieved sample indexing and time point indexing (Figure 6) on three different target templates. Exonuclease degradation of the samples has enabled time point indexing while keeping the molecule indexing intact. There are many future applications of this technology as this can solve a major problem in DNA sequencing – obtaining high quality long reads - needed in all current and most of future DNA sequencing platforms.

We have shown that tags can be incorporated to in multiple layers of information to supply target origin, molecular origin and positional origin. By utilizing these indices, virtual read lengths equal to the lengths of the targets can be achieved, surpassing that of traditional Sanger sequencing as well as the capacity of spanning difficult region. We have showed PCR based tagging of nineteen 3000-bp targets representing the lambda genome. Viewed as an improvement of read length, this constitutes a 30-fold improvement to the Illumina system used, and illustrates the principle of converting short-read

technologies into long-range analysis. We also illustrated potential applications by including a variable target region of canine mtDNA as well as the major part of the TP53 gene for cancer cell lines that, in principle, could greatly improve either accuracy or dynamic range of detection. Illumina has recently launched a similar product, Moleculo, that use the same principles of assembling reads based on a tagging approach, although TES was the first published approach with significant public attention. The principle of TES is covered by LV patents.

In the serial compartmentalization approach using emulsion technology we demonstrate that the method in a high throughput manner enables unique barcoding, monoclonal amplification and phasing of amplicons from single DNA molecules in millions of discrete compartments. This technique complements the tiled end sequencing (TileSeq) approach by enabling barcoding of many different DNA fragments. In addition, with minor adjustments the number of phased products per genomic fragment could be increased by employing random amplification instead of target-specific primers. By performing complete phasing of 10-100 kb genomic fragments such a technology would substantially reduce the bioinformatical load of genome assembly (Borgström et al, 2014, manuscript).

We have also in this work developed bioinformatics tools to identify unique tags that do not match human and mouse sequences (Costea et al 2013). These can be used as more directed forms of indicies i.e. creating a library of index that can be used for labeling DNA or RNA targets. Another use of this resource, employed with in the project, was in the design of bit sequences suitable for the design polymers (WP3).

Finally, we performed for the first time a comprehensive in depth analysis of type IIs Restriction Enzymes demonstrating variable digestion patterns in a variable manner.

Type IIS RE	Distance	$\langle E_D \rangle$	$\langle E_L \rangle$	$\langle E_C \rangle$	n
SmuI	4/6	0.87	0.60	0.52	3
GsuI	16/14	0.73	0.40	0.30	5
BsmFI	10/14	0.85	0.56	0.47	4
BpmI	16/14	0.88	0.38	0.33	1
Eco57MI (A)	16/14	0.83	0.38	0.32	1
FauI	4/6	0.85	0.56	0.48	3
BbvI	8/12	0.89	0.33	0.29	1
BpuEI	16/14	0.87	0.30	0.26	1
BtgZI	10/14	0.88	0.23	0.21	5
BsgI	16/14	0.86	0.25	0.21	1
BseMII	10/8	0.81	0.20	0.16	1
BseRI	10/8	0.76	0.18	0.13	1
Eco57I	16/14	0.73	0.18	0.13	3
AcuI	16/14	0.81	0.26	0.21	4
FokI	9/13	0.83	0.13	0.11	1
Eco57MI (B)	16/14	0.81	0.11	0.09	1
BceAI	12/14	0.65	0.05	0.03	4
MmeI	20/18	0.30	0.03	0.01	1

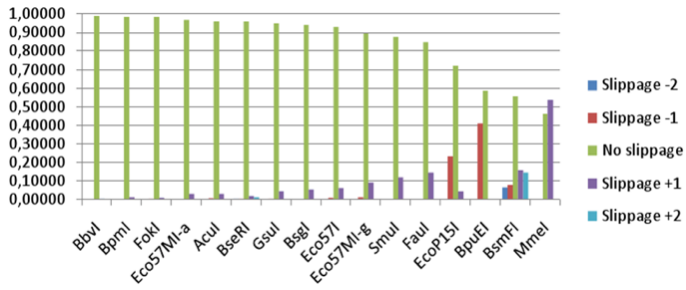


Table 1 – (Left panel) Combined digestion and ligation efficiencies for 17 type IIs restriction enzymes screened. "Distance" denotes the distance in nucleotides between the recognition site and the restriction site of each enzyme. "$\langle E_D \rangle$" denotes the average digestion efficiency, "$\langle E_L \rangle$" denotes the average ligation efficiency, "$\langle E_C \rangle$" denotes the average combined efficiency and "n" denotes the number of independent experiments used to calculate the average. (Right panel) The different enzymes are reported along the x-axis. Eco57MI has two different recognition sites, both of which were tested in this assay (Eco57MI-a = CTGAAG, Eco57MI-g = CTGGAG). Slippage is reported as +1, +2, -1 or -2, where +1 means that the observed restriction site is situated 1 nt downstream of the expected site and -1 means that the observed restriction site is situated 1 nt upstream of the expected site and similar for +2 and -2.

The design polymer protocol however requires at least two enzymes that digest over a significant distance. We also observe that the re-ligation efficiency of digested substrate molecules is variable, and reaches 60% for the enzymes with highest efficiency, indicating that there is a significant risk of the restriction enzymes remaining bound to the substrate after digestion. Combined efficiency of digestion of surface bound molecules and ligation reaches a maximum of ~50% for the enzymes with highest efficiencies. The publication describing patterns of slippage is under revision (Lundin et al, 2014).

Binary DNA read out on microscope slides (WP 4)

Overview of the read out technology

Activities in this workpackage aim at providing a glass-based platform that enables work on the integration of the binary conversion protocol with a surface for readout, during the period when the optical disk-based readout platform was in early development stage. Readout of the glass-based platform is achieved using existing instrumentation, for example by fluorescence microscopy. In practice, the glass-based platform provides a stepping-stone for the optical disk-based platform.

An aspect of the work is to provide the interface between both the template molecules and the surface of the disk (glass slide), or alternatively the externally amplified template and the disk (glass slide). This step is termed the target-to-disk (T2D) step and several technical implementations is envisioned depending on the amplification strategy, binary conversion protocol, sequencing mode (targeted or genome-wide), etc.

A key challenge for the project is the attachment of various detection bit oligonucleotides (DBO) in a controlled manner and position (both absolute positioning in reference to the disk, and relative positioning in reference to each other). For high-throughput applications the density of these is preferably in the sub-micrometer range, while for targeted re-sequencing applications (companion diagnostics) a lower density is sufficient. A repertoire of concepts for sub-micron positioning can be envisioned; a partial listing is available in Annex I of the project. It is important to emphasis that approaches based on x-y stages cannot deliver the required density, at least not in high-throughput manner. Therefore, a focused activity for the project has been to identify means for accurate DBO positioning in a manner that is both economically feasible, lends itself to automation and has the potential to in the future to be developed into a manufacturing process. A natural choice for this system has been to use the addressing functionality of the optical disk drive. Based on the initial good progress on the laser-induced transfer (LIT) work in WP5, the work on high-density deposition was down prioritised in this WP and work focused more on the actual attachment chemistry, which could then be transferred to WP5 and the disk setting.

The choice of the amplification method has a major impact on the user friendliness of the technology and depending on the sequencing applications, several alternatives can be envisioned. For large-scale sequencing (genome-wide applications), approaches based on either solid-phase amplification, emulsion PCR or on-disk compartmentalized amplification reactions can be used. Solid-phase amplification and emulsion PCR are established amplification technologies and described in the scientific literature, hence is not addressed in this project. These methods are also in commercial use; the Illumina sequencing technology utilizes solid-phase amplification, while the 454 sequencing (Roche) and Ion Torrent technologies are based on emulsion PCR.

Compartmentalized amplifications in micrometer-sized wells on disk (glass slide) offers the benefit of enabling efficient amplifications in solution, while at the same time providing means for immobilisation of amplification product onto surface through sequence complementarity (compare with emulsion PCR). In this work package, we have evaluated the use of 1-20 µm-sized compartments casted in PDMS and use of PCR, RCA and/or helicase-dependent amplification (HDA) methods to amplify a template in a manner that generates amplicons hybridised to surface-attached primers.

To get access to compartments in the size range of 1-20 µm, custom-manufactured structures need to be used. A candidate method is through generation of silicon-based master and subsequent stamper, followed by casting of the microstructures in PDMS. Use of PDMS as substrate for amplification purposes is however challenging due to adsorption related issues. As part of the work in this WP, we have established a method for permanent modification of the PDMS surface to avoid these unspecific interactions.

Another key activity of this work package is the demonstration DNA conversion resulting in detectable binary features on the surface, and the initial evaluation of labelling strategies that are also compatible with the optical disk-based readout. Labelling strategies that can be envisioned include use of beads as physical marker (light scattering), or enhancement of manometer-sized gold particles to larger structures using chemical deposition of silver ions onto the initiating gold particle to create either a reflective surface or larger scattering particle.

Key aspects of work

In this work developing a DBO attachment chemistry to surfaces relevant in optical disk manufacturing was done. The attachment method was based on coupling of a terminal amine-group of the DBO to a surface functionalised with either carboxyl- or epoxy groups. The substrate surfaces were either (i) commercial glass slides (Nexterion E) or (ii) surfaces coated in house with GOPTS. The initial use of glass slide substrates facilitated work on the biochemical reactions in parallel to the DBO surface attachment work on-disk, which was considered a more complex task. A number of complementary strategies for attaching DBO to disk surfaces have also been tested. Because of the lateral precision requirements (on the micrometer scale), these alternative methods are presented in WP5, which focuses on the optical disk as substrate.

In addition we developed a method for patterning a surface with detection bit oligonucleotides (DBO) with accurate sub-micron resolution. The primary challenge relates to the high DBO density and the associated requirements on both absolute and relative positioning. A secondary challenge relates to identification of an attachment chemistry that can be used on both glass slides and can with minor modifications be transferred to the optical disk setting.

Results

Detection bit oligonucleotides have been successfully coupled to microscope slides, both commercial Nexterion slides and home coated slides. The molecular functionality after surface coupling has also been demonstrated (Figure 6 - 7).

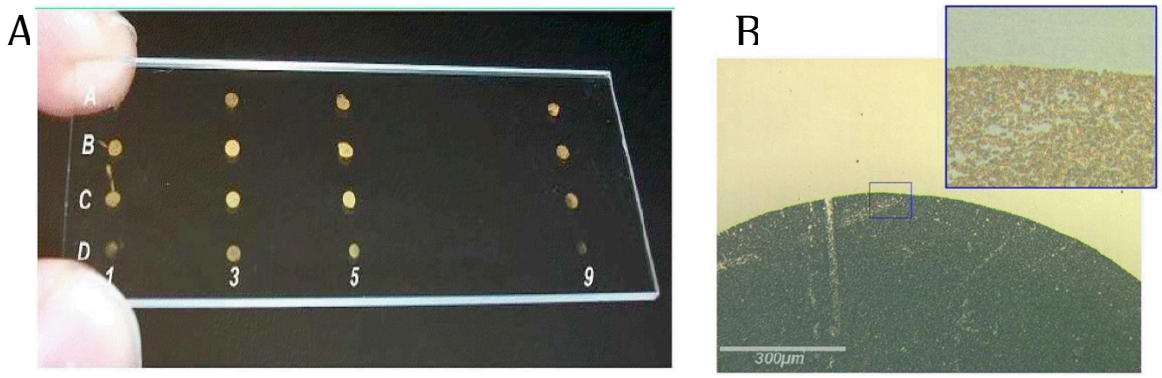


Figure 6 – A) Photograph of slide with spots of surface-attached oligonucleotide after labelling with MyOne beads and washing by submerging the slide in 70°C water with agitation. B) Close-up of a spot (spot 5c in A) showing surface attached oligonucleotide subsequently labelled with MyOne beads. (Ref: TR460)

On the optical disk, the DBO must be attached at predefined positions with sub micrometer precision. Hence, the optical disk system (optical disks and optical drive) was utilised in the continued DBO attachment work in WP5.

We also evaluated alternative attachment chemistries for the disks (reported in detail in WP5) and include: (i) Thiol-terminated DBO to gold surface (ii) Amino-terminated DBO to silane-epoxy coated SiO2 or SiON (iii) Amino-terminated DBO to activated spin-coated polycarbonate. In addition we investigated methods for creating sub-micrometer DBO spots with high lateral precision, (also reported in WP5): (i) Contact printing (ii) LIT (laser-induced transfer).

The second route of investigation - compartmentalized in micrometer-sized wells on disk (glass slide) offers several benefits of integration and significant progress was achieved. On the way to perform amplification in microcompartments, currently the most appropriate amplification system is based on helicase-dependent amplification, the hybridisation assay for the multiplex detection DBOs on the microarray glasses has been established, and the in situ amplification protocol in macro scale has been optimized. The problem with the PDMS porosity has been detected and solved, as confirmed by the

hybridisation assay in the microcompartments. The final goal, amplification in the microcompartments, has been reached as planned.

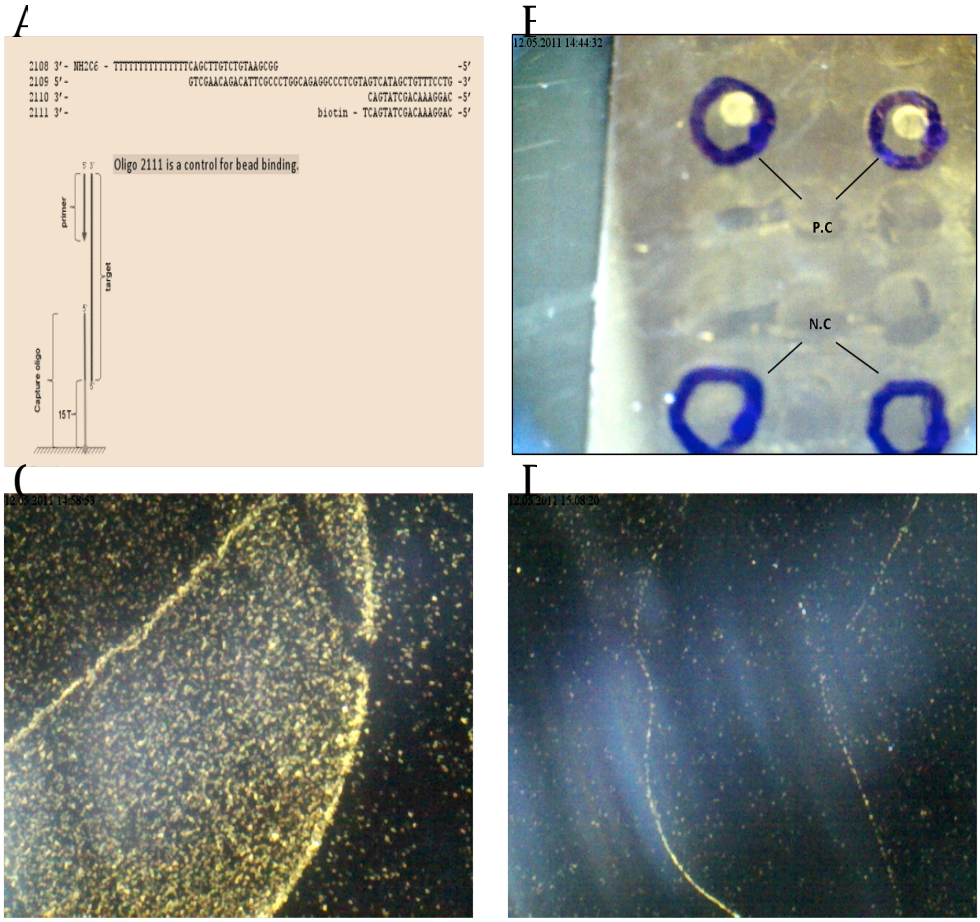


Figure 7 – Molecular functionality demonstrated using a hybridisation and extension-based assay. A) Schematic principle of assay. B) Comparison of positive and negative cases, where negative are polymerase-negative controls. C) Magnification of a positive and D) negative feature.

On the optical disk, the DBO must be attached at predefined positions with sub micrometer precision. Hence, the optical disk system (optical disks and optical drive) was utilised in the continued DBO attachment work in WP5.

We also evaluated alternative attachment chemistries for the disks (reported in detail in WP5) and include: (i) Thiol-terminated DBO to gold surface (ii) Amino-terminated DBO to silane-epoxy coated SiO2 or SiON (iii) Amino-terminated DBO to activated spin-coated polycarbonate. In addition we investigated methods for creating sub-micrometer DBO spots with high lateral precision, (also reported in WP5): (i) Contact printing (ii) LIT (laser-induced transfer).

The second route of investigation - compartmentalized in micrometer-sized wells on disk (glass slide) offers several benefits of integration and significant progress was achieved. On the way to perform amplification in microcompartments, currently the most appropriate amplification system is based on HDA, the hybridisation assay for the multiplex detection DBOs on the microarray glasses has been established, and the in situ amplification protocol in macro scale has been optimized. The problem with the PDMS porosity has been detected and solved, as confirmed by the hybridisation assay in the microcompartments. The final goal, amplification in the microcompartments, has been reached as planned.

The results of microcompartment amplification presented in Fig 8 - 9. Here, we demonstrate the production of high-density array of amplicons representing genetic material in digital format with high amount of clonal DNA in the spots. Due to the high yield of clonal DNA and high flexibility of amplicon design, our protocol can easily be adapted as sample preparation step for next generation sequencing in general, and in particular for the DVA platform envisioned in Digital Sequencing. The results obtained in this work has been accepted for presentation in the highly competitive Conference on Miniaturized Systems for Chemistry and Life Sciences international conference (rejection rate of 60%) that will be held in San Antonio, Texas, USA, October 26 - 30, 2014[1].

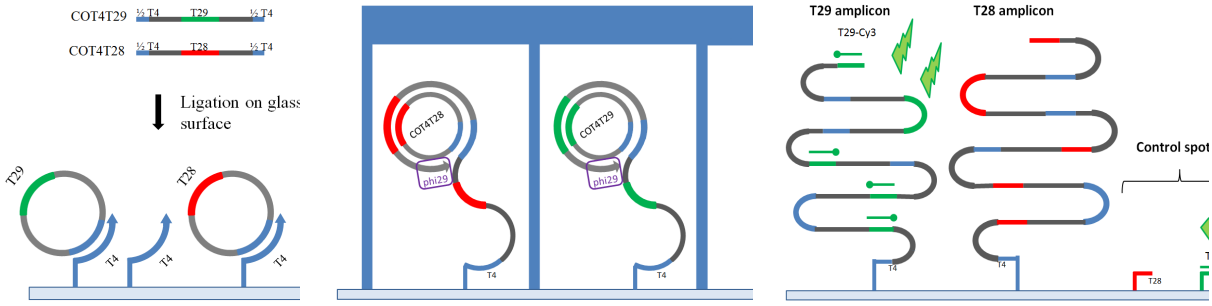


Figure 8. Microcompartment amplification (Left)) The COT4T28 and COT4T29 padlocks were hybridized on the T4 oligonucleotides (blue) printed on glass surface and circularized by ligase T4. (Middle) Initial phase of microcompartment amplification. Glass was covered with PDMS stamper carrying RCA mixture in the microwells. Phi29 amplifies the COT4T28 and COT4T29 padlocks using T4 oligonucleotide as primer. Branching (not shown here) of the primary chain was used to convert linear amplification to exponential. (Right) T29-Cy3 and T28-Cy3 was used respectively as probes for detection of the T29 and T28 amplicons on T4 spots. T28 and T29 spots (no amplification) were used as hybridization assay controls.

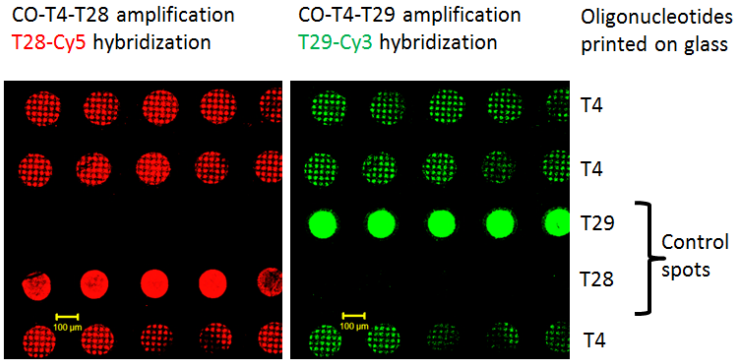


Figure 9. Fluorescent microscope images of the 10 µm spots with RCA amplicon on glass surface. Confocal fluorescent microscope was used to detect the fluorescent signals on the glass surface after hybridization. The image with red signals corresponds to T28Cy5 hybridization with T28 amplicons (spots with pattern) or T28 control (spots with continuous signal). The image with green signals represents T29Cy3 specific hybridization in the same way, amplicons with pattern and controls with continuous signals. Only the spots containing T4 capturing oligonucleotides accumulate amplicons in the microcompartments and have pattern with the spot size of 10 µm.

The hybridisation was selected as a quick assay to test the microcompartment quality. The HDA protocol should be used as the final proof for the microcompartment functionality. One of the primary challenges for this objective is the optimisation of the HDA protocol and to enable clonal DNA amplification in a single microcompartment. Also, for the binary conversion strategy that takes places in two parallel microfluidic reaction channels and utilises apyrase and Klenow polymerase, a simplified model system for the microfluidic compartments has been developed.

On the modification of PDMS surface, it is possible to conclude that the plasma treatment creates different surface functionalities (silanol, amino, PEG groups). Two substances APTES and PEG are important for the positive treatment result. Formation of thin film of aminosiloxane stabilizes the surface against the rearrangement and PEG after grafting by plasma co-polymerization increases the surface wetting properties. However, PDMS as material is not suitable for industrial fabrication, and the material absorbes moleculas. Within Digital Sequencing, KTH has therefore developed a novel OSTE polymer based microcompartment microwell array and is currently adapting the system for direct arraying on the DVD substrate (Figure 10 -11).

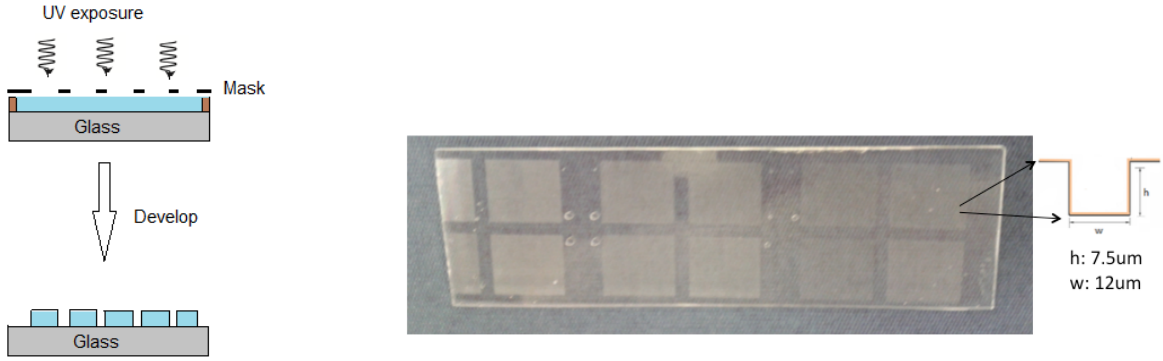


Figure 10. (Left) Fabrication process of the microarray wells on glass slide. The OSTE polymer is exposed to UV through a mask, and following a developing step, the array are permanently deposited on the glass slide. (Right) OSTE based microfabricated microwell array. The wells are 12x12 mm square with a height of 7.5 mm. The array allows for spotting DNA direct into the glass slide, and hence significantly reduce the amount of steps involved in microcompartment based DNA amplification.

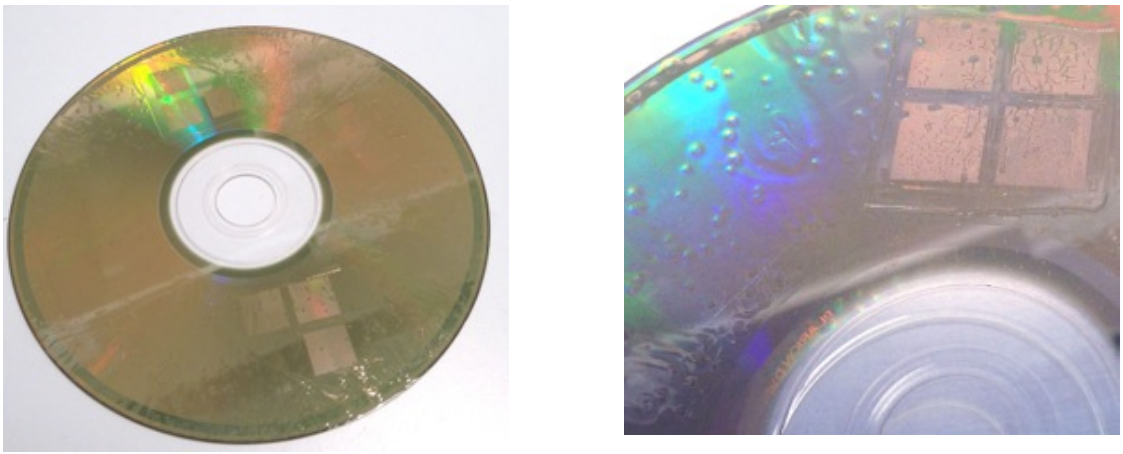


Figure 11. OSTE based microfabricated microwell array on DVD substrate. The wells are 12x12 mm square with a height of 7.5 mm. The microwell array are clearly visible.

A indicated previously, an bioinformatics pipeline was established that can generate large number of unique index DNA sequences that can be used in various workpackages including both of the binary codes proposals, DBOs and a set of random indexes that may be used in the tiled end sequencing protocol (WP2). The pipeline is implemented using C++ and parallellised using Open MPI platform. Input sequences can be created either using random sequences or by parsing an archea genome. The pipeline generated >100 000 unique index sequences aimed to have no or minimal cross interaction. We believe that this pipeline can be used for any given application that requires indexing of sequences of any number and it is implemented in a flexible way so that it can be adapted for various lengths, sequence content as well as their interference with experimental sequences (Costea el at 2013).

Experimental validation of the hybridization events in real time with binary DNA demonstrate a sensitive and specific hybridization pattern without any cross talk between binary bits.

Binary DNA read out on optical disks (WP5)

Overview of the on disc technology

The aim is to demonstrate that the core biochemical elements of the digital sequencing process can be realised on an optical disc surface and the binary value detected using an optical drive. Further, according to the stepping stone strategy, at the final level it should be possible to carry out the on-disc chemistry at a density suitable for the equivalent of an entire genome sequencing or a more targeted strategy suitable for diagnostic purposes, relaxing the specifications.

Central to the low cost requirement of the DNA digital sequencer, we will use a commercially available optical drive (e.g., DVD drive). Such drives typically operate at single laser wavelengths of 658 nm and it is preferred to use this as our optical detector. A common method of molecular labelling uses fluorescence markers. These are unsuitable for our requirement. Wavelength of 658 nm is not ideal for exciting fluorescent markers and indeed, extra drive features would be needed in order that the fluorescence be detected; adding unwanted extra cost.

The integration and detection of the biochemical conversion steps on the optical disk surface can be achieved using either reflection or transmission-based approaches. For reflection based approaches access to the DVD drive's firmware or physical capture of a signal is required, making the approach challenging. For transmission-based approaches a second detector can be added to the drive, and this is used to record the amount of light transferred through the disk and the products of the biochemical conversion process placed on top of the disk.

Key aspects of work

Here we demonstrate controlled immobilisation of detection bit oligonucleotides on metallic or dielectric film surfaces. Its achievement is an essential step toward the goal of performing DNA sequencing on an optical disc surface. In addition to being able to attach the oligonucleotides to the surface it must be bonded strongly (e.g. covalently) and most importantly retain molecular functionality in order that further biochemical steps such as hybridisation, extension and ligation can be carried out.

For monitoring of chemical events we chose to investigate the transmission-based approaches to detected locally altered transmission through the disk. Altering of the transmission through the disk was achieved by having particles bind specifically to the positions where a positive reaction had been taken place. The particles were of two types: (i) streptavidin coated nanogold that were grown into micrometer size particles by a silver enhancement process, or (ii) opaque (iron oxide) streptavidin coated 1-3 μm large paramagnetic microbeads.

This silver enhancement process involves wet chemistry but has the advantage that it can optically distinguish sub-micron features accurately. First, the disc surface is exposed to a solution containing streptavidin conjugated nanogold particles. These bind strongly to any surface attached molecules containing a biotin group. Secondly, the nanogold particles are exposed to a silver enhancement solution which preferentially deposits silver on the nanogold, building up a reflective surface over time. The silver enhancement is rapid but easily controllable and within a comfortable time span (e.g. minutes). Amplification of the gold signal by 10-100X is readily (and routinely) achieved. The reaction is insensitive to light, is simply stopped by washing in water, and needs no fixing. The silver enhancement process builds up a layer of reflective silver over the target molecules which have been previously selectively conjugated to streptavidin nanogold particles via a biotin linker. The reflective silver layer is ideal for detection in an optical drive.

Also streptavidin-conjugated microbeads are suitable for the optical detection of surface attached DBO molecules. The benefit of this approach is the simple, direct detection of the label without enhancement reactions of any kind. Due to their size, the beads are not suitable for labelling sub-

micron features, but are useful where the feature size is of the order of 10's of microns. The microbeads are of sufficient size and density to diffract the incident laser light of the DVD drive, thus altering the reflectivity and transmission signal in the region where they are attached, which can be detected. Further, this technique has some important advantages over silver enhancement in that it involves one fewer experimental step, and importantly, the density of the beads attached to the surface affects the size of the electrical signal detected by the drive. Consequently, this method of labelling can provide quantitative as well as qualitative information about the reaction being studied.

A variety of investigated methods for attaching detection bit oligonucleotides to thin film surfaces on optical discs was evaluated. Each approach is described in detail in the Periodic Reports.

Submersion: The thiol - gold submersion method (Figure 12) is important for two reasons; i) it provides a controlled system that facilitates investigations and understanding the mechanisms involved in formation of thiol-based self-assembled monolayers (SAM) and ii) it is needed for the backfilling step to block the remaining gold surface after laser-induced transfer (LIT) deposition of the DBO in the SAM approach. The different aspect that were evaluated were the non-specific binding of proteins and DNA, capability to block unwanted silver enhancement (SE) and possibility to attach DBO to the surface in a well-presented manner.

PDMS printing: The next phase was to refine and miniaturise the SAM formation process by printing the thiols onto the substrate using micro-patterned polydimethylsiloxane (PDMS) stamps. DBO spot sizes in the region of 10's of microns are achievable using this technique. SAM analysis methods used in this work included imaging ellipsometry and surface plasmon resonance (SPR).

Laser Induced Transfer: Following on from PDMS printing, the LIT process allows transfer of spots of DBO's at sizes an order of magnitude smaller (in the sub-micron range). LIT utilises the focussing ability of a commercially available low-cost DVD laser diode to deposit DBO's in tiny spots onto addressable substrates at high speed. Figure 13 below shows one implementation of LIT. The combination of DVD technology and precise oligo attachment allow this method to potentially achieve full genome DNA sequencing capacity on a single DVD size optical disc.

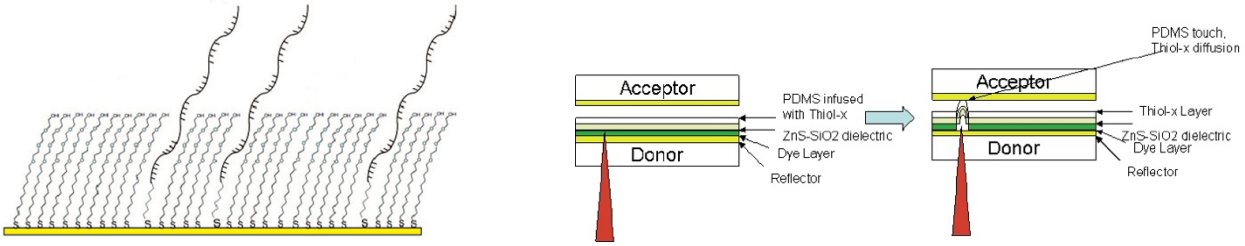


Figure 12 – (Left panel) Mixed self-assembled monolayer containing oligonucleotides and polyethylene glycol functionalised with thiols. (Right panel) Schematic showing laser-induced transfer of a non-absorbing oligo using DVD 658 nm laser diode.

Alternative option: Inkjet printing and attachment by incubation in humid atmosphere of amine conjugated DBO's on dielectric or polycarbonate surfaces. The deposition of sub-micron regions of 40 specific DBO in regular arrays via the LIT approach requires many new processes and is a high risk task. An alternative Digital Sequencing approach has been identified which removes the need to attach unique DBO for each position in the 20 base sequencing path, but instead generates linear paths consisting of identical DBO in conjunction with known pre-deposited reagents. Advantages of this route include an easier manufacturing since there are far fewer DBO variants to deposit/attach. Further, the alternative route achieves DNA sequencing density via both temporal as well as spatial resolution which relaxes the need for sub-micron features. This allows use of larger spots, meaning that inkjet printing could be used to apply the DBO. It also allows use of conjugated microbeads (in size range few microns which effectively scatter incident laser light to produce a detectable change in reflectivity/transmission) in place of silver enhancement for specific optical detection of DBO attachment / hybridisation.

Results

SAM formation by submersion: Thiol-functionalised molecules in a solution will spontaneously adsorb to any gold surface present, forming a self-assembled monolayer (SAM). Generally, the quality of SAMs increases with incubation time. However, when dealing with mixed monolayers they tend to change their composition over time (while incubating). In the first stage of the adsorption process, small molecules reach the surface quicker than large ones and therefore have an advantage. Later on in the assembly process, the length of the alkyl chain is important due to the intermolecular lateral vdW bonds. In thiol SAM formation, the optimal number of carbons in the alkyl chain is in the range 12-16 because it is associated with the lowest energy for the system. Bulky groups, such as oligonucleotides, in the molecule are energetically unfavourable and such molecules will normally be replaced by less bulky counterparts with time. The time needed for assembly of high quality SAMs is dependent on the concentration of the thiol molecules, but normally varies between 16 and 40 h. It should be noted that in this section, the percentages given for the molecules participating in the SAMs are those in the incubation solution.

Before initiating work on thiol-functionalised oligonucleotides, biotinylated thiols (EG_6Bt) were used. In this way, the system could be investigated without involving the challenges of oligonucleotide presence in the SAM and hybridisations, extensions etc. To obtain suitable surface concentration of biotins, the biotinylated molecules were mixed with similar molecules but without the biotin moiety (EG_xOH, where x = 3, 4 or 6). The device seen in Figure 13A was constructed to be able to subject a single disk substrate to many different solutions at the same time without risking cross contamination. In Figure 13B, where the resistance of SAM-coated gold to silver enhancement detection (SE) is investigated, the resulting 5 mm diameter spots are seen. It was found that all the ethylene glycol-containing thiols provided the gold surface with sufficiently protective coating to prevent unwanted SE.

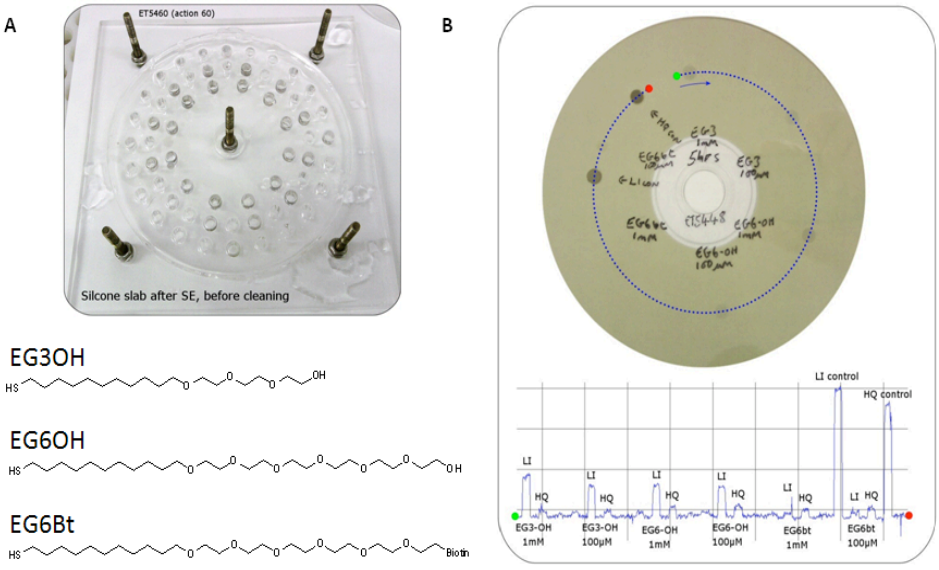


Figure 13 - A) A moulded polydimethylsiloxane (PDMS) rubber slab was used to form circular reaction compartments on the substrates. To prevent leakage, pressure was applied to the substrate and PDMS slab between two PMMA plates held together by screws. B) The resistance to non-specific SE was evaluated for two polyethylene glycols with different length, EG_3OH and EG_6OH, as well as for a similar but biotinylated polyethylene glycol. The vertical axis corresponds to reflection and any silver enhanced areas gives rise to an increase in signal. The different SE kits, LI and HQ, were used and they display slightly different behaviour. The controls were made at the gold surface of the substrate (i.e. no thiol monolayer deposited). Interestingly, EG_6Bt shows best SE preventing properties, possible due to the hydrophobic nature of the biotin moiety that is likely to reduce the ability of any aqueous solutions to contact the substrate surface.

The functionality on disk substrates was also investigated and as seen in Figure 14. The resistance to unwanted SE (pink circles) on three types of SAMs (blue, brown and orange polygons) is very good. It is also seen that the non-specific binding of SA-Au is the SAMs lacking biotin is very low (red circles

in blue and brown polygons). The SAMs that do have biotinylated DBOs in them display nearly as strong response as the positive controls (yellow circles), which was SE done on unmodified gold (i.e. maximum response).

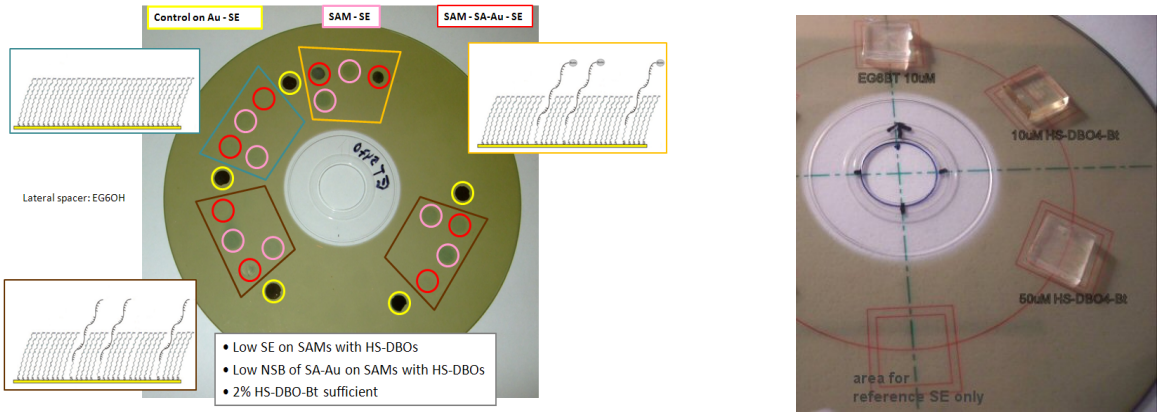


Figure 14 – (Left panel) Three different SAMs on a gold coated disk substrate. Blue polygon: 100% EG6OH, brown polygon: DBO:EG6OH 1:49 and orange polygon: DBOBt:EG6OH 1:49. As expected, small responses to SE and SA-Au + SE (pink and red circles, respectively) are seen for the non-biotin SAMs (blue and brown polygons), while clear responses are seen for SAMs containing biotinylated DBOs and the positive controls (SE straight on the gold coating). (Right panel) Arrangement of printing on a substrate: clockwise EG_6GBt 10 µM, HS-DBO_4-Bt 10 µM, and HS-DBO_4-Bt - 50 µM. The 6 o'clock area was left blank to silver enhance as a reference at the end of the experiment

Summary of SAM formation by submersion

(i) Thiolated alkyl chains containing biotinylated ethylene glycol moieties (e.g., HS-C11-EG6Bt) are detectable optically after silver enhancement at concentrations of just 1% in SAM's formed by submersion. This is good news for the LIT process, meaning that only a small fraction of the LIT'd molecules need be present and presented in the correct orientation in the SAM after LIT and backfill is complete.

(ii) Attempts to produce SAM's containing longer chain DBO molecules (e.g., HS-C6-DBOBt where the DBO is typically 20 bases long) were less sensitive. Detectable silver enhancement was observed when HS-C6-DBOBt concentration of 2% or above was present.

(iii) The use of long chain DBO molecules caused two problems - i) IRAS measurements showed they were significantly more prone to desorption during submersion in backfill solution, ii) SAM films containing DBO's were more prone to damage during subsequent on-disc chemical steps such as hybridisation.

Ways forward centre on using DBO's which have the same number of carbon atoms in the alkyl chains as the PEG thiols used for backfill. (Typically, the DBO molecules contain 6 C atoms in the alkyl chain compared to 11 for the PEG thiols rendering them energetically less stable). It is expected that this would reduce preferential desorption of the DBO and improve the quality of the SAM film such that it is better able to resist hybridisation steps.

Relating to the LIT process, the macro-scale submersion experiments showed that if just 2% of the LIT'd DBO molecules are attached and correctly oriented, then it should be possible to detect them optically on disc via silver enhancement.

SAM formation by PDMS stamp (microprinting): The LIT process for DBO's is in essence a very fast printing method on the sub-micron scale (Figure 13). To mimic the LIT, but on slightly larger scale, several experiments were made to print SAM's of molecules of interest onto gold films on optical disc surfaces using PDMS stamps. It was attempted to provide answers to questions like: "What quality could be achieved? What conditions were needed?".

The microprinting was carried out on an ungrooved substrate freshly sputtered with a 12 nm thick film of Au. Several PDMS stamps inked with biotinylated molecules at 10 and 50 µM concentrations were

carefully placed on the gold film surface (Figure 15). Microprinting contact was maintained for 30 minutes with 10 and 50 µM solutions of the molecules. After the stamps were removed, the remaining bare gold surface between the printed spots was back-filled with EG_4OH.

The back-filled spots were then treated with nanogold streptavidin SA-Au, and silver enhanced with LI silver for 40 minutes. Optimised conditions were obtained under the following conditions: (i) printing from higher 50 µM concentration of the biotinylated oligos. (ii) printing from a buffered solution was preferred to water. In addition it was found unnecessary to incubate the printed oligos in humid atmosphere to arrange into a better quality SAM. The SAM quality was found to be sufficient immediately after the contact printing step.

To confirm that we can determine the presence of the printed oligonucleotides in an optical drive, the disc with the silver enhanced printed thiols was analysed on Pulstec ODU1000 658 nm DVD tester. A measure of the disc reflectivity trace around the circumference of the disc through the silver enhanced printed spots was made (Figure 14, left panel).

It is easy to detect macroscopic regions of high reflectivity due to silver enhancement in the printed regions. It is expected that the fully silver enhanced control will show the greatest increase in reflectivity and, as the film is not patterned, the waveform in this area will be smooth. In the areas where pattern of the SE thiols developed we expect reflectivity increase from the enhanced spots alternating with lower reflectivity in the blocked background. The increase of reflectivity in the spots is dependent on the availability of the nanogold sites attached to the biotinylated oligos or thiols.

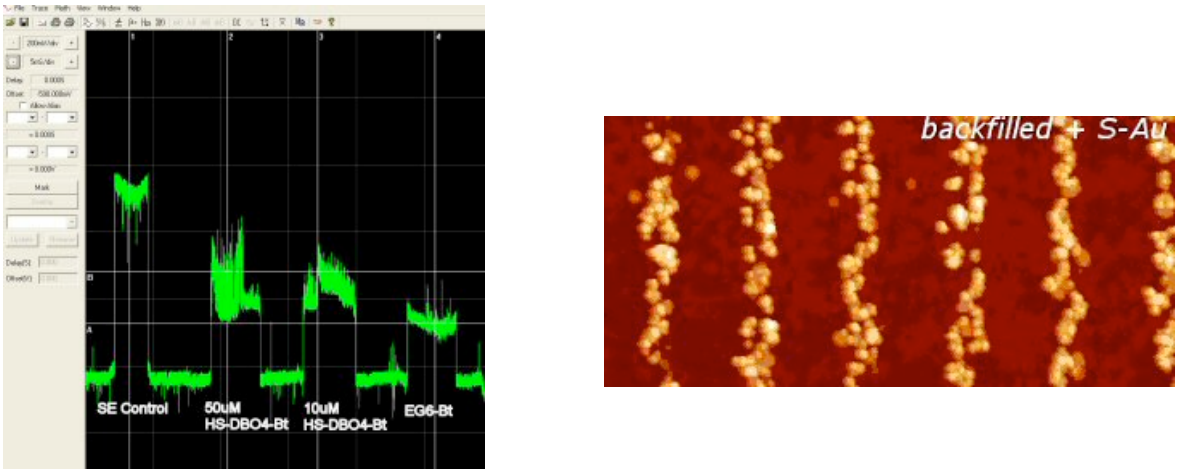


Figure 15 – (Left panel) Pulstec waveform read at 5.19 ms^{-1} velocity (1.5X) with 1 mW read power. The baseline corresponds to the reflectivity of the untreated gold. The peaks show the reflectivity of the silver enhanced spots of the printed biotinylated SAMs: From left to right for: Silver Enhanced control Au surface; 50 µM SH-DBO_4-Bt; 10 µM SH-DBO_4-Bt and 10 µM EG_6Bt . (Right panel) AFM image, 5 µM wide, of an Au coated acceptor disc after lamination to a grooved donor disc topped with a PDMS thin film inked with EG_6Bt, backfill with EG_6OH, conjugation of SA-Au and silver enhancement. Silver enhancement is only present in the regions where the tops of the PDMS coated grooves came into contact with the gold film. Groove lines from the donor disc have been imprinted onto the ungrooved acceptor disc. What remains are vertical lines of silver enhancement corresponding to the successful transfer of EG_6Bt where the PDMS film was in contact with the surface.

Printing of SAMs onto gold coated DVD surfaces from PDMS thin films with the LIT process requires the printing of HS-DBO-Bt molecules in sub-micron scale from a donor DVD disc coated with a thin dielectric film. On top of the dielectric it is proposed to coat a thin film of PDMS and ink this surface with the required DBO. The effectiveness of microcontact printing of biotinylated thiol molecules from a thin film was evaluated (Figure 15, Right panel) and we can conclude that experiments clearly shows that where the inked PDMS surface comes into contact with the tops of the grooves of the acceptor disc, there is transfer of biotinylated material which subsequently silver enhances. Importantly, this experiment also allows us to see how thickly the silver grows on the printed features. We find that up to 30 nm of silver will grow in the contact regions. This is sufficient to generate a significant optical change on the disc surface which is detectable in an optical drive.

Summary of the microprinting experiments have shown

(i) bulk printing

(ii) accurate patterned printing of both EG_6-Bt and SH-DBO_4-Bt. on gold in 10's of microns scale. At concentrations between 10-50 µM, printed regions were easily optically detectable using 658 nm laser after backfill, streptavidin nanogold (SA-Au) attachment and silver enhancement.

(iii) in order to achieve repeatable print spot size, the stamp application pressure, contact time and ink concentration must be controlled.

(iv) printing from thin films

(v) accurate printing at nano scale is possible from a PDMS thin film coated on DVD grooves (0.74 µm).

SAM formation by LIT (nanoprinting): SAM formation via Laser Induced Transfer (LIT) of the SAM material to an acceptor substrate was a completely new concept, unreported in literature. Several preliminary experiments were made, to establish viability, in the lead up to the LIT of a SAM material. Figure 16 (Left panel) shows a AFM micrograph of top disc shows clear rectangular spots of material which has been transferred (LIT'd) to the mirror disc surface in a pattern determined by the laser writing pulse length and frequency For addressability and readback, it was necessary to LIT material onto an addressable substrate which could be subsequently read back in a DVD drive (Figure 16, Right panel). Using this technique, it is possible to position SAM materials in precise locations such that multiple different materials (for example, 40 different oligos) could be placed into an known, addressable arrangement.

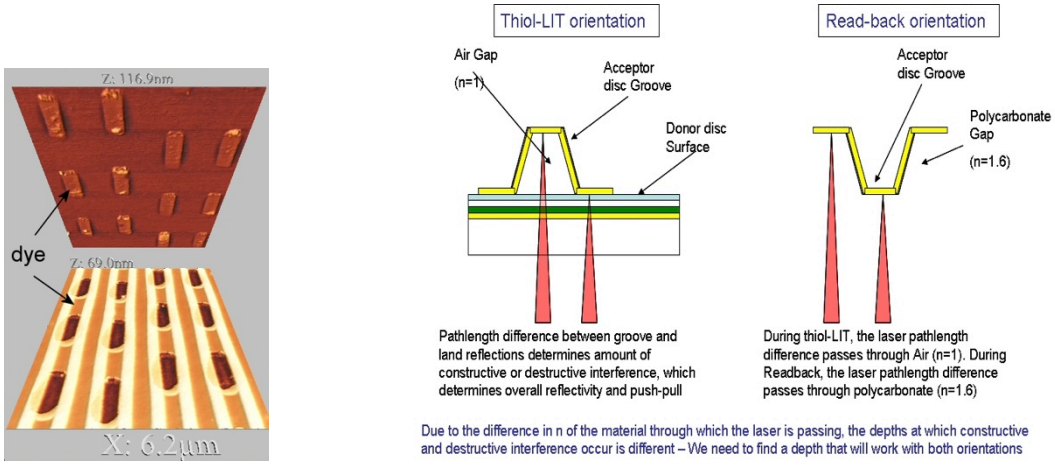


Figure 16 – (Left panel) Atomic Force Microscope images showing organic dye (Ultragreen) Laser Induced Transfer from a grooved polycarbonate DVD-R donor disc (bottom) to a grooveless polycarbonate DVD 'mirror' acceptor disc (top (Right panel) Schematic of LIT and readback

During experiments, it was serendipitously discovered that the presence of an air gap between the donor and the acceptor led to a significant improvement in the performance. The air gap served not only to improve transfer but also to separate the land regions of the acceptor disc from the donor disc altogether, thus preventing unwanted touch transfer of material to the lands of the acceptor disc.

In order to determine if the silver enhanced LIT marks were really due to silver enhancement of a biotin molecule, a LIT comparison of biotinylated (HS-EG6Bt) vs. non biotinylated (HS-EG4OH) molecules was made side by side. It was expected that the LIT of molecules which do not contain Bt would not bind streptavidin gold and would not therefore silver enhance. The result in Figure 17 below supports this.

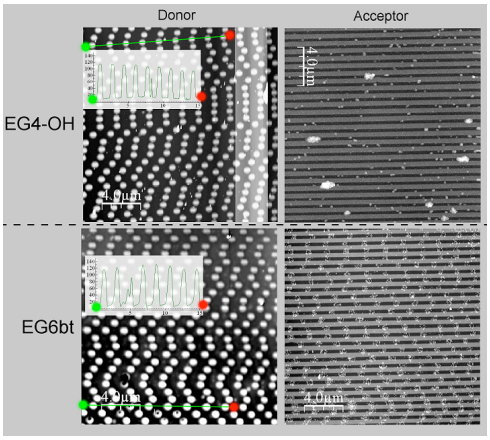


Figure 17 - Atomic Force Microscope topography images. Direct comparison of laser induced transfer of non-biotinylated HS-EG4-OH thiol against biotinylated HS-EG6-Bt. The silver enhanced spots in the bottom right hand image are due to gold labelled streptavidin binding to the biotin which is attached already to the HS-EG6Bt molecules. This result gives great confidence that HS-EG6Bt is still active after transfer.

Summary SAM formation by LIT (nanoprinting):

(i) The status reached with this work showed that biological molecules can be addressably arranged into sub-micron spots using LIT.

(ii) A patent has been filed to this effect.

(iii) Considering the viability of LIT for manufacturing settings, we have encountered some repeatability problems when performing LIT with longer chain oligos (DBO). They are not as conducive to forming neat SAM arrangements. Often the biotinylated end of the oligo was not presented for attachment to streptavidin gold and silver enhancement, leading to variable response.

Two alternative chemical methods for attaching and optically detecting oligonucleotides on the disc surface has also been demonstrated (i) DBO attachment to silanised SiO_2 dielectric surface on optical disc has been successfully achieved and selectively detected using microbeads. This method is thought to be suitable for DNA Digital Sequencing on disc and is regarded a lower risk route than LIT (ii) Attachment of amine conjugated DBO to silane treated SiO_2 surface via incubation in humid environment. But the density of the immobilised oligonucleotides is lower than density of oligonucleotides immobilised on the silanised SiO2 surface.

Two alternative methods have been used to label surface attached DBO molecules to enable their detection in optical disk drives in an non-fluorescent manner. And we have successfully demonstrated detection using streptavidin-coated nanogold and silver enhancement (Figure 18) and we have demonstrated the biochemical compatibility of the detection route utilising 1-3 um sized opaque beads by carrying out a probe attachment, hybridisation of unlabelled target, extension of probe using protruding target sequence using biotinylated nucleotides, and – finally – detection of extension using streptavidin-coated paramagnetic beads (Figure 19).

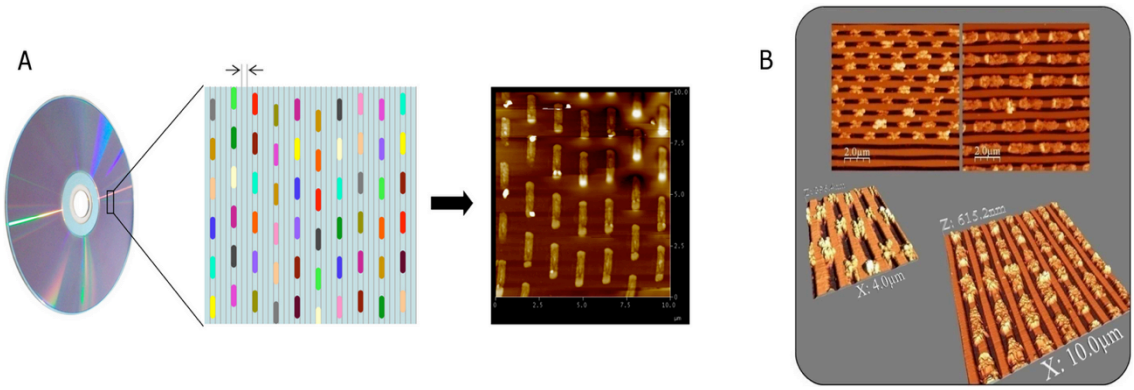


Figure 18 - The track pitch of DVDs is 740 nm and here spots have been deposited on every fourth track using a formatted donor substrate. The feature size is 1 µm x 350 nm and was imaged using atomic force microscopy. B: Functionalised alkyl thiols transferred to a formatted, gold-coated acceptor substrate and subsequently silver enhanced into reflection sub micron sized spots for ease of optical detection. These images were also acquired with atomic force microscopy.

In the efforts towards integration of all the techniques on optical disc surface we have obtained several key results. The work relating to the integration of the biochemical conversion with microfluidics-based reaction control has been primarily been done by KTH following the transfer of disc prototyping capability from Plarion and using the two installed Discipher instruments. To prepare microcompartment structures for target amplification, we casted structures in PDMS. The poly(dimethylsiloxane) (PDMS) rubber is a candidate material for fabrication of microwell part of optical disc (see further details in WP6 on the functionality of this).

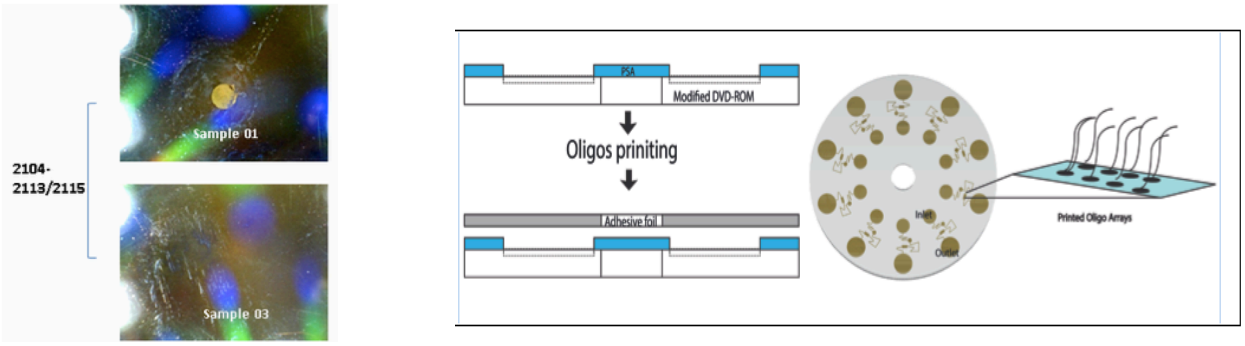


Figure 19 – (Left panel) Results showing that a polymerase-mediated extension with a biotinylated nucleotide can be used to generate a biotinylated binding site for the streptavidin-coated paramagnetic beads. (Right panel) Multi-layer Disc Fabrication for DNA microarray - fluidic structures on DVD. Laser cut PSA creates the fluidic structure on DVD ROM. The DNA is then spotted on the semi-transparent DVD ROM. The pre-spotted DVD substrate is finally bonded to top pressure-sensitive adhesive layer containing the fluidic channels.

The results showed that oxygen plasma treatment is most effective method which changes gradually the topmost surface of the PDMS making it highly hydrophilic. It makes possible to graft a uniform thin layer of aminosiloxane. The FTIR and XPS show that this layer is stable to surface rearrangement during 4 months of exposure in the lab air. Thus at the PDMS surface the amino and the silanol functionalities were successfully created. These functionalities can be used for further DBO immobilization. This protocol can be used for preparation of the microwell substrate.

We developed a robust surface modification protocol and adapted it to the DVD substrate. Following this, DNA arrays were printed on the silane-epoxy modified bottom DVD ROM and the semi-transparent DVD substrate was bonded to the top substrate consisting of pressure-sensitive adhesive with fluidic structures (Figure 19, right panel). Next, we performed DNA hybridization followed by either silver-gold enhancement chemistry and or bead based for visualization of the microarray spots. High quality, low-background, “images” of microarray spots were achieved inside the channels of DVD surface, available for automated analysis. To optimize the flow condition, we used fluid plugs in tubing and connected to external pump system connecting the fluidic channels. Figure 20 shows the plug-flow based integrated liquid handling to perform the following integrated steps: hybridization, washing, labeling and silver-gold based signal enhancement.

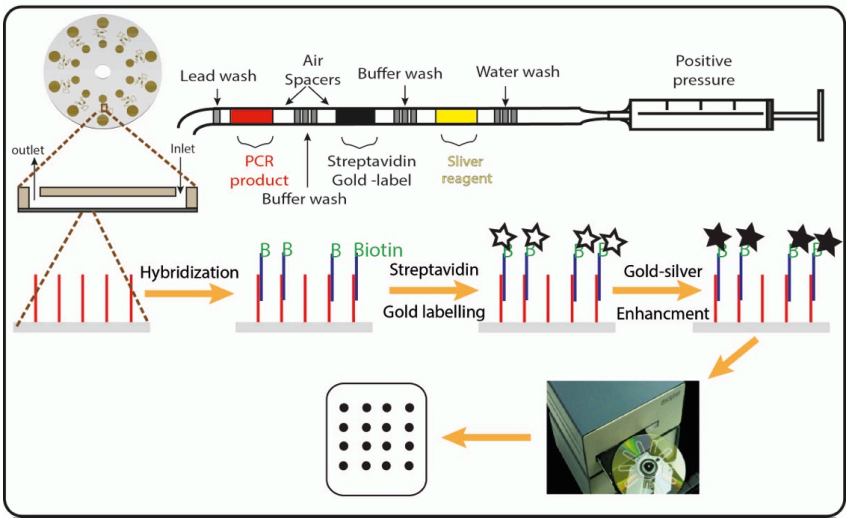


Figure 20. Microfluidic plug-flow system to control the liquid movement inside microchannels on the DVD substrates. The tube is pre-filled with the desired solutions separated by an air gap. A syringe pump is then used to precisely control the liquid movement through the microfluidic channels on the DVD substrate to execute the following consecutive assay steps: hybrirization (red color), labelling with streptavidine coated gold nanoparticles (black color), and silver enhancement reaction (yellow color). The microarray spot is the read inside the DVD reader for detection.

As a proof-of-concept for integrated microfluidics, the Discipher system was applied to the detection of genetically modified organisms (GMOs) in maize and soy food samples. DNA probes were printed in a microarray format on the polycarbonate layer of DVDs, and integrated controls guaranteed the absence of false-negatives and false-positives. After microarray spotting, a fluidic layer consisting of pressure-sensitive adhesive with microfluidic structures was bonded manually. The hybridization assay, including the washing protocols and development reaction, was performed by dispensation of samples and reagents through the inlet, with the centrifugal pumping and hydrophobic valves controlling the fluidic movement (Figure 21). After removing the fluidic layer, the disc was inserted into the DVD player and microarray images were captured. Hence, the only required materials were standard store-bought DVDs, plastic chambers, tips, pipettes, oven, and a standard DVD drive. Excellent correlation was achieved between the optical density registered by DVD pick-up and the GMO content (Figure 21).

Summary, the remaining challenges are to (i) To demonstrate all functional assays on the rotating disc inside a microchannel and (ii) Development of a fluidic system allowing for the biochemical reactions to take place in a controlled manner; here use of external controlling mechanisms such as syringes is allowed.

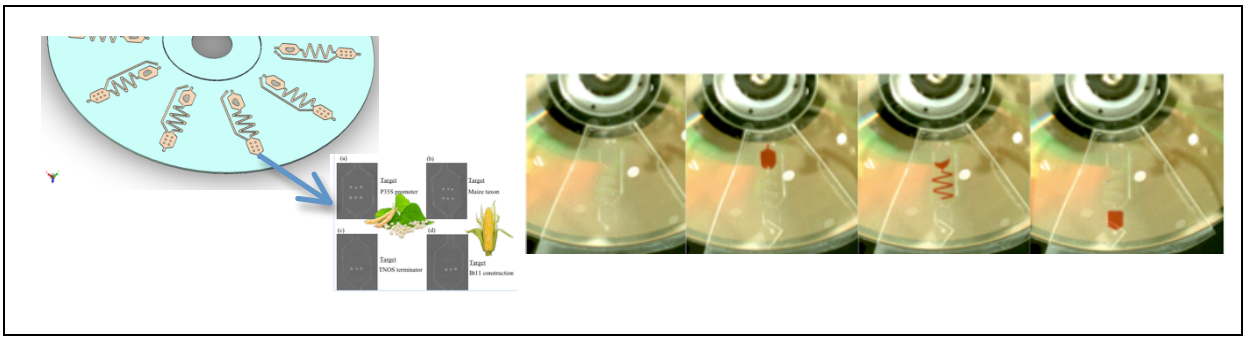


Figure 22. (Left) Schematic image showing DVD substrate for integrated hybridization and washing step. (Right) Example images captured using an external spin stand (is described in WP6) used to capture of images in fluidic design development and bioassay optimization steps. (Insert) Examples of microarray DVD image results obtained from four certified reference samples, showing the assay sensitivity and specificity

Development and validation of an optical disk based DNA sequencing system (WP6)

Overview of the optical disk based DNA sequencing system technology

The main aim is to develop the instrument and disc hardware to encompass and measure the on-disc chemical processes optimised in the previous work. Consistent with the Project Mission Statement, the hardware design is to be based around a consumer optical disc drive in order to leverage the technical benefits and low costs coming from 20 years successful industry development of this technology.

Key aspects of work

Running in parallel with Binary DNA read out on optical disks WP during the first 36 months of this project, we have in this work focused on developing instruments to work with: (i) LIT arrays of sub-micron DBO's (accessing spatial domain only) (ii) larger spots of printed DBO's distributed in linear paths (accessing spatial and temporal domains).

During the Project Period II, focus has been on establishing disc production capability and development and assembly of a prototype drive. In total six different production approaches have been evaluated. We have made the strategic choice to focus on the development of the system based on larger printed DBO's distributed in linear paths. Hence, during PPII, the development work towards the LIT compatible drive has been discontinued. The prosecution of the LIT patent application has been continued. The present prototype instrument is based on the work of the Project Period I, and during Period II we have focused on providing a stand-alone instrument with all required functionalities and during period III prototype instruments have been placed at consortium members.

Results

The two lead methods for on-disc DBO attachment require different disc structures for on-disc DNA sequencing. Prototype disc designs and production considerations have been addressed for both the lead DBO attachment methods.

LIT - Where DBOs are deposited in at addressable positions into sub-micron arrays of greater than 40 different spots. Central to the sequencing process is that the DNA sample is sequenced in short stretches of approximately 20 nt per read. The binary DNA concept represents each base by 2 bits. One DBO spot on the Biorecord surface represents 1 bit, consequently 40 DBOs are required to sequence each 20mer fragment. Sequencing is carried out in parallel in millions of microcompartments, each representing one 20mer DNA fragment. 50 million wells on a DVD can therefore decode 1 Gbp of sequence information (Figure 23). To produce each Biorecord disc, the production process consists of writing 40 DBO's on the disc in series. The Biorecord disc would be in turn vacuum laminated (no glue) to 40 separate donor discs, each carrying one of the 40 DBO's necessary for printing into its correct location in the overall array. After each lamination step, the disc is placed in a modified DVD drive and the DBO's are LITd into their correct locations. Once this process is complete, the disc is ejected from the drive, spliced apart and the next donor disc laminated etc. Figure 23 (middle) shows this process. Using an arbitrary waveform generator we fed precise write pulse trains into a Pulstec DVD ODU1000 tester which burned marks in specific locations on a DVDR disc (Figure 23, right). We first burned a train of single 3T pulses separated by much longer 69T gaps – (top). We then adjusted the trigger timing and burned another train of single pulses adjacent to the first set of marks. Continuing on we successfully burned 10 trains of marks in accurate positions demonstrating that the triggering accuracy was sufficient to attempt LIT. This result confirms that the proposed production route for the Biorecord disc is feasible.

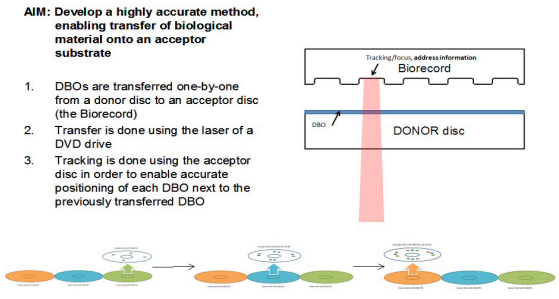


Figure 23 – (Left) Close up view of disc - How best to divide up the disc into individual regions of 40 different DBO's with minimal loss of capacity (Middle) Shows schematically how 40 different DBO's can be addressably LIT'd to the acceptor disc (which when completed becomes the Biorecord on which DNA sequencing is carried out). (Right) Oscilloscope waveforms showing the HF signal while tracking on a DVDR disc. From top to bottom the traces show how individual written marks can be accurately placed next to each other when written 'blindly', based only on the accuracy of a disc trigger mark and timing clock

Though confirmed as feasible, many aspects of the disc construction and sequencing process involved novel steps involving high technical risk. The process was reviewed by all the beneficiaries. An alternative sequencing route was devised which retained consistency with the project aims and which we believe involves lower risk.

Inkjet printing - Where DBOs are printed in multiple larger spots (10's of microns) forming linear paths that act as the sequencing route. One key element of the alternative sequencing design is to make use of temporal as well as spatial dimension to provide sequencing capacity. Adding a temporal element allows for use of reduced numbers of larger DBO spots of size range which can be inkjet printed rather than LIT'd. A second simplification is to use significantly fewer DBOs (Figure 24, left). This has an impact on reagent cost and reduces risk of misprinting of the wrong material in the wrong place. It also simplifies the future manufacturing of the disks. Thirdly, the reduced density and use of a linear path of DBOs allows for the sequencing process to be made in microchannels rather than individual wells. This eliminates the need to conduct parallel hybridisation in 50 million wells at the same time and all the risks of leakage that that method entails. Finally, the use of microchannels and larger spots allows for microbeads to be used as the labelling method. This is a more simple, potentially quicker and potentially reversible route that is preferred to silver enhancement. The drawback of the approach includes a reduced capacity per disc, but provides means for a simplified end user processing.

By placing a photodiode in the DVD drive situated above the disc, it is possible to collect the transmitted light and interrogate reactions taking place in the microchannels. Specifically in the case of digital sequencing, it is possible to interrogate and count microbead labels. The presence of and quantity of microbead labels imaged directly relates to the sequence of the target of interest By placing a photodiode in the DVD drive situated above the disc, it is possible to collect the transmitted light and interrogate reactions taking place in the microchannels. Specifically in the case of digital sequencing, it is possible to interrogate and count microbead labels. The presence of and quantity of microbead labels imaged directly relates to the sequence of the target of interest.

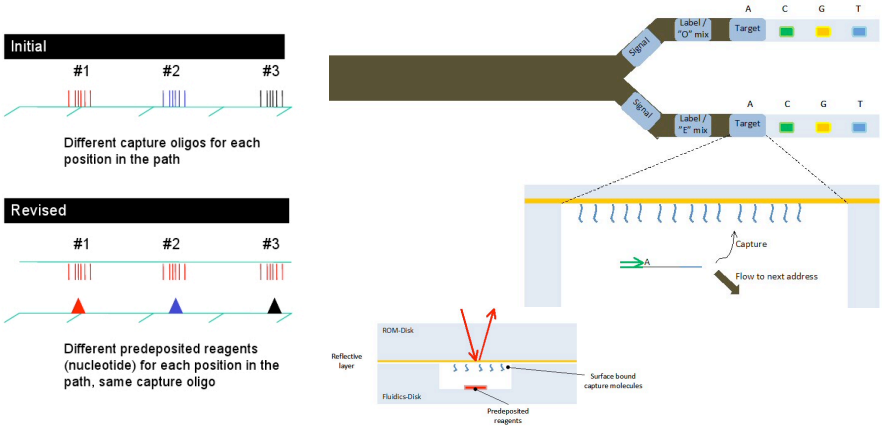


Figure 24 – (Left) By depositing the same oligo in linear paths but with different predeposited nucleotides (dNTP, symoblised by triangles) we significantly reduce the need to print lots of different DBOs. (Right) Cross sectional view of the disc showing the DVDROM (top), semireflective layer (yellow) and microchannel sections (bottom). The disc contains standard DVD ROM data on one half, i.e., looks exactly like a DVD which means it can be read by a standard DVD drive. The difference is to bond a microchannel substrate to the top of the DVD half (where a normal DVD disc would contain a blank dummy substrate). (Middle) Schematic of sequencing process in microchannels. Essentially, the amplified target of interest to be sequenced flows along the channel, is partially captured at each DBO position in the path. At each DBO a biochemical detection reaction takes place, consisting of the steps 1) positioning, 2) binary labelling, and 3) binary detection. The three steps may be separated in time.

Development and manufacturing of working prototype disc. Starting from the bottom, the discs used in the project utilise a DVD-ROM disc. This 0.6 mm disc contains all the operational information required to read the disc in a standard DVD drive, including the spiral groove that has a 0.74 µm track pitch. These discs are typically injection moulded PC or PMMA, and also have the important function of focusing the incident light such that it focuses at the reflector. There are several producers of such DVD-ROM discs, including Eximpo/Northern Star (Czech Republic) and Axxicon (The Netherlands).

On top of the DVD-ROM a set of thin layers are sputtered. Collectively, these few nm thick layers are termed the stack of the disc. We have in previous deliverable reports and quarterly scientific progress reports discussed the development and optimisation of this stack. Turning to the key functionalities of the layers in the stack, some are there to provide strong adhesion to the adjacent layer, some (typically metals) function as a mirror reflecting the incoming light of the optical pickup unit of the DVD drive back to the detector, while some (typically at the top of the stack) provide a surface that can be activated to bind biological molecules. In some instances an intermediate activation step is carried out after sputtering; for example, if the top sputtered layer is SiO2, then this could be silanated with an epoxysilane. The epoxy groups can then be used to bind aminated biomolecules.

Two disks were glued to perform microfluidic devises. The main task was to develop the method to control adhesion and adhesion stability in the bond. PosiTest pull off adhesion setup was applied to bonded disks with different adhesives. After pull off test the failure mode was inspected. The failure takes place across PC substrate-Au interface which is most weak boundary. It was found relatively low adhesion force between the PC disks with gold coatings comparing with the reference PC – adhesive-PC bond. It was pointed out that improving the adhesion across Au-PC interface is needed for producing the next devices.

We have demonstrated two disc manufacturing routes providing discs for both rapid development work and for future production like settings. These are based on either PSA cut structures, or bonding of injection moulded discs halves using screen printing technology.

Prototype reader. During the Project Period II a stand-alone prototype instrument has been assembled. This instrument is used to both control the biochemical steps during the sequencing, as well as for readout of the results. A key property of the prototype instrument is the strategy and use of components enabling high-volume, low-cost production of the instruments.

This means that although low cost DVD drive hardware can be used as the basis of the instrument for performing the DNA sequence data read back, it needs significant modification to be able to decode the data. The schematic in Figure 25 below shows what is necessary. An extra link to the RF amplifier is needed to extract the data information on the disc and pass it through bespoke read hardware and on to the PC for analysis. In addition, address data from the DVD+R disc is read and passed to the PC at the same time. This combination of binary data coupled to an address position allows the software to decode the data it receives.

In terms of productionisation of such units, initial cost estimates were in the range of a few thousand euros for the writer / reader units (which would be used at the manufacturing site) and a few hundred euros for the reader only units (which would be used at the customer site).

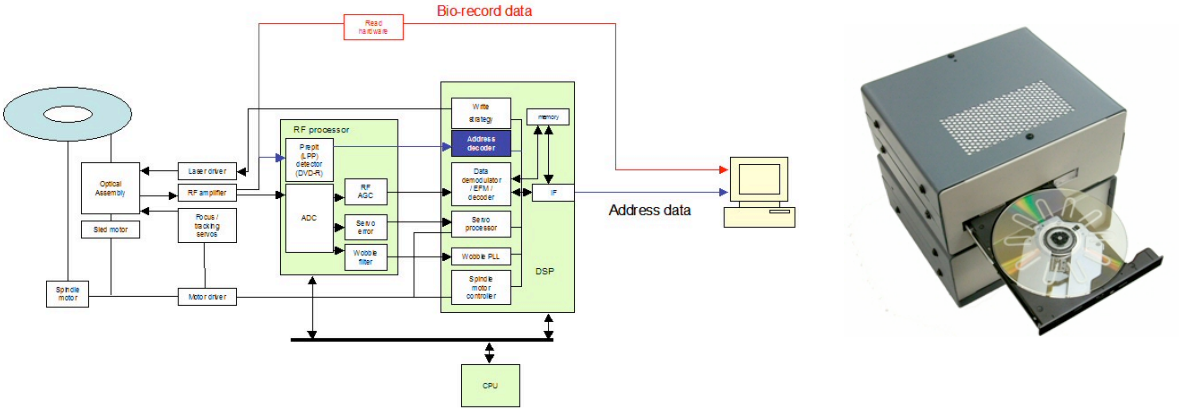


Figure 25 – (Left) Schematic of drive modifications required in order to read and decode DNA sequencing information from the disc surface. (Right) Photograph of prototype DNA Sequencing Instrument.

An early prototype version of the instrument with key functionality established had been assembled during the Project Period I, but the instrument was controlled by external devices such as an Lakeshore PID controller for temperature, an oscilloscope for capturing the signals from the photodiodes, and an external ADC to convert the analogue signal from the oscilloscope into a digital format. Figure 25 shows a photograph of the unit. During Project Period II we have designed, developed and manufactured a PCB that carries out all of these steps. The current prototype instrument is a stand-alone instrument enabling temperature control of the reactions, control of liquid movement using rotational microfluidics, and imaging of results using transmission-based strategies. The digitalised data is sent further to a PC for processing using custom prepared software.

A standard low cost (~$20-40) DVD drive is used as the core of the system. The key modifications and functionalities that have been added to the system include: (i) a series of photodiodes and a bespoke controlling board to collect the transmitted light, (ii) temperature detectors, heaters and fans to control the temperature according to user instructions, (iii) an extra inbuilt motor to provide a fine level of control of the rotation, (iv) a PCB and associated software that integrate and control all functionalities of the drive, and (v) a custom software to process and analyse the data, including a GUI for user presentation

The key difference compared to a standard DVD drive is the use of a second photodiode (D2) that is distal to the laser source in reference to the disc. This approach enables monitoring of reactions both in the microfluidic channel structure placed above the reflector (distal, blue square), and of objects placed within the focal plane (not shown). The fundamental principle of data generation using the prototype system is based on end user instructing (via the GUI) the DVD drive to read from a known

start position and following the spiral track (groove) of the disc, while the D2 captures the amount of light transmitted through the disc. SW based processing of captured data generates a 2-dimensional image that is subsequently analysed. The prototype system includes software at various levels. We have implemented a software development strategy that ensures adequate control of stability, but that also provides different users with the required flexibility. The key software function of the software is to capture, process and analyse the images generated during the readout of the sequencing reactions. The labelling approach that has been chosen is based on use of micrometer sized paramagnetic (opaque) beads generating a reduction in transmission. The signal captured by the PD is generated into a 2D image, such as the one shown in Figure 26. A key feature of the software is the ability to quantify the signal by counting the number of beads captured. In the long term this may be done by custom developed software or by using existing open source software. The current approach in the Project utilises imageJ, an sopen-source software, in this step. Figure 26 (Right) shows a comparison between the bead count and the concentration of a target molecule in model system, demonstrating a good starting point for further optimization. An update of the Discipher software has enabled image capture with sustained heating. Finally, extensive evaluation of the two instruments at KTH shows the robustness of the developed Discipher instrument (Figure 27). Future development is required to enable automated quantification.

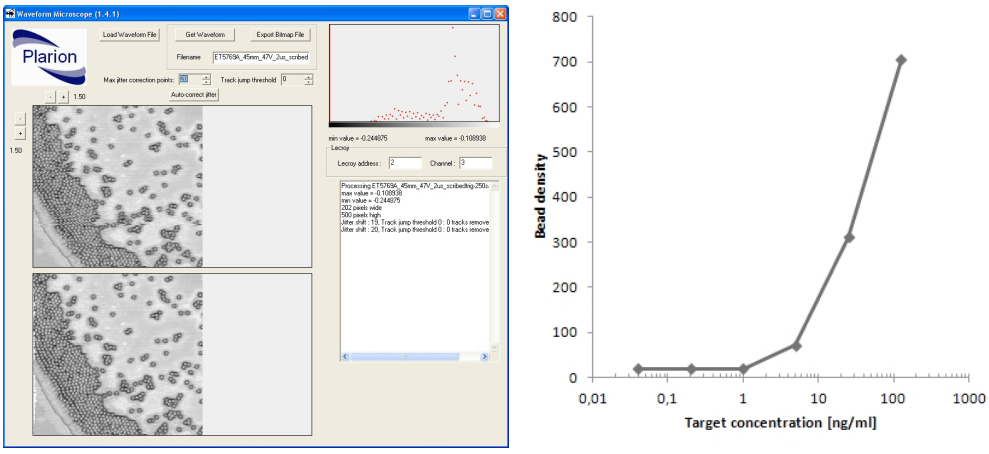


Figure 26 – (Left) The figure shows a slightly fuzzy bead image, and below it, the same image with the jitter corrected. The beads are much more crisply defined. (Right) Observed bead density vs. concentration of a target molecule

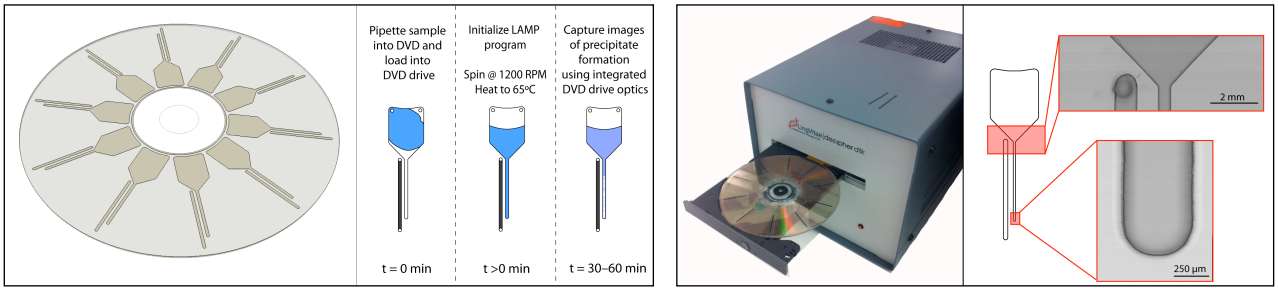


Figure 27 (A) Schematic of the microfluidic DVD. (Left) Each device is arrayed 10 times for use in multiplexing or on-DVD control samples. (Right) Overview of LAMP assay process steps. Once the assay reagents are pipetted into the DVD, the chamber is sealed with tape, and placed into the DVD drive. All following steps including heating and scattering-based detection are automated within the system. (B) The modified commercial DVD drive used in this work. A heater and secondary photodetector module was added to enable the temperature control and image capture steps for an automated LAMP assay. (Right) Images captured using the internal DVD laser system show detailed sections of the microfluidic device. The system is capable of capturing images down to 1 micron resolution.

Summary

The prototype instrument for controlling the biochemical reactions and for readout of the sequencing results has been designed, developed and manufactured

Ten prototype instruments have been manufactured and have been delivered to project beneficiaries and selected third party developers (see Exploitation activities in the Management report for further details).

The use of the prototype reader in adjacent application areas has been carried out and published in Ramachandraiah H, Amasia M, Cole J, Sheard P, Pickhaver S, Walker C, Wirta V, Lexow P, Lione R, Russom A. Lab-on-a-DVD: standard DVD drives as a novel Laser Scanning Microscope for image based point of care diagnostics. Lab on a Chip, 2013 Apr 21;13(8):1578-85

The potential impact (including the socio-economic impact and the wider societal implications of the project so far) and the main dissemination activities and exploitation of results

A successful launch of a high-throughput low-cost DNA sequencing instrument will represent a major catalytic event within the life sciences, resulting in both increased number of users and overall level of usage. Hence, the first company to introduce such an instrument will be having an unparalleled competitive advantage and vast market potentials arising from many sectors.

The current end user market for DNA analysis consists of scientists engaged in research areas such as drug discovery, diagnostics, health care, agriculture, and forensic medicine to name a few. The next generation sequencing (NSG) market is rapidly evolving with a large number of developments taking place to increase accuracy and speed, and reduce costs of sequencing. It is the fastest-growing and most lucrative segment in the genomics space with an estimated growth of 16.3%. The global NGS market was valued at $1.3 billion in 2012 and is poised to reach $2.7 billion by 2017. Despite the significant market size already, it is predicted to continue to grow 20-30% per year once the cost of DNA sequencing is driven down to more affordable levels, i.e. once a new generation sequencing platform facilitates a replacement of current generation platforms. Consequently, the business potential for companies involved is extremely high, and the impact on the health factor will be substantial.

The field of DNA sequencing has evolved rapidly over the most recent years with introduction of massive high throughput instruments such as Illumina HiSeqXTen targeting population studies, rare diseases and cancer diagnostics. Large population studies are currently in the process of being established such as Denmark, England, China and recently the Middle East. Rare diseases are the focus of the Genomics England effort and well as the Swedish genome project. While whole genome cancer studies are more related to fundamental understanding of tumor progression and treatment. All of these applications deal with whole genome sequencing to uncover novel variants in the genome and the HiSeqXTen system, albeit with high investment costs of 10 MUSD, provide a human genome for 1000 USD at 30-fold coverage. Currently there are no competitors that can compete with the Illumina platform in terms of cost per genome and accuracy. It is obvious that low cost and high accurate human genome sequencing will have major impact to these areas of research.

Another trend is the focus on targeted DNA sequencing of a panel of genes or single genes. Several instrument providers are focusing on these applications with a key aim to develop an integrated and rapid systems with minimal hands on time. In addition to the use in basic research the overaching goal is to introduce these systems in a diagnostic setting with enormous market potential that will impact many fields in the health sector. Today you can design a panel of approximately 500 genes to address all genes couple to a drug/treatment in cancer. The status of a gene and its potential mutations can therefore guide treatment as well as be used to monitor residual disease. The term companion diagnostics has been introduced to represent the specific analysis of a gene coupled with a specific drug treatment. For example Roche has developed both a diagnostic kit and a treatment alternative for melanoma cancer directed towards the BRAF gene. Currently, Life Technologies, Pacific Biosystems and Illumina have instruments directed towards these applications. Still most of these instruments require significant investments, but it is clear that the societal impact will increase substantially during the next decade.

Ancillary technology should also be considered in the context of developing next generation sequencing systems. These include sample preparation systems for all types of applications such as: whole genome sequencing, targeted re-sequencing, de novo sequencing, RNA-Seq, ChIP-Seq and

methyl-Seq. Another area that has been somewhat neglected is bioinformatics - the raw sequencing data generated by today's sequencers does not reflect on any meaningful information in itself. Also, open source tools for analyzing NGS data require adequate bioinformatics training. These factors open up huge opportunities for commercial NGS bioinformatics software, workbenches, and services, whose market is currently fragmented and has a relatively smaller size.

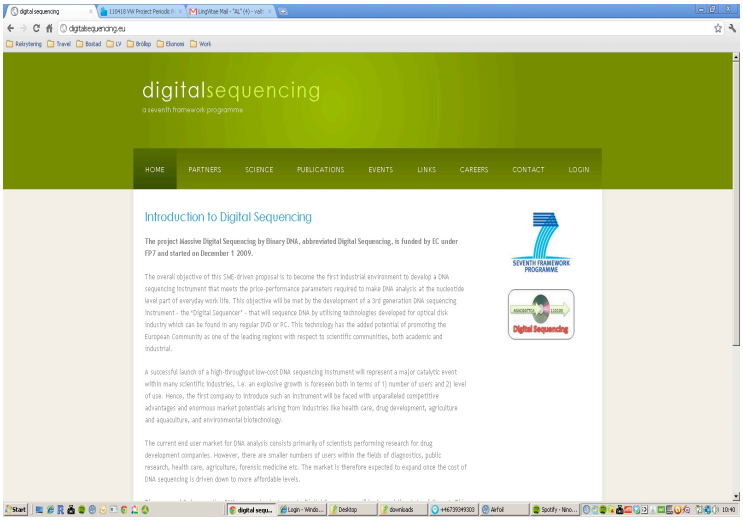
The digital sequencing project has provided stepping stones towards a low cost DNA sequencing instruments using the low cost DVD platform with significant potential impact in the field. The work has included ancillary technologies such as sample preparation with barcoding, automated solutions and software development. The main dissemination activities within Project include scientific publications and scientific presentations. In particular, the developed Tile-Seq development has received attention especially as this was published just before the launch of Moleculo, a similar technology to in silico assemble long reads from short reads. Another key demonstration was the use of the DVD platform as a diagnostic tool with accompanying national and international publicity. In addition Ling Vitae AS, the former coordinator, has had several commercial activities and discussions with leading commercial entities.

The DVD platform for sequencing was not fully completed within the scope of the project yet important achievements were obtained and demonstrated. To address the high capacity instrument we demonstrated that we can accurate position genetic information onto a DVD disc using the LIT technology (patent pending) with the potential to obtain 1 Gbp of information per disc. However, during the later phase of the project we focused on a less high capacity DVD solution addressing the needs in targeted DNA sequencing such as companion diagnostics. A market niche that remains unexplored for novel low cost instruments.

The concept of binary sequencing has become more accepted in the scientific community demonstrated by the launch of Illuminas NextSeq500 using only two labels to deduce the 4-digits of the genetic code. In addition binary sequencing is also the key technology for Stratos that recently made a strategic alliance with Roche (15 MUSD investment), most probably as a part in the puzzle to provide robust and accurate nanopore sequencing. Roche has also recently purchased Genia Technologies for 350 MUSD to achieve such a synergy.

The address of the project public website

http://www.digitalsequencing.eu/ (administrated by the Ling Vitae AS, the former coordinator)


digitalsequencing
a seventh framework programme
HOME
PARTNERS
SCIENCE
PUBLICATIONS
EVENTS
LINKS
CAREERS
CONTACT
LOGIN
Introduction to Digital Sequencing
The project Massive Digital Sequencing by Binary DNA, abbreviated Digital Sequencing, is funded by EC under FP7 and started on December 1 2009.
The overall objective of this SME-driven proposal is to become the first industrial environment to develop a DNA sequencing instrument that meets the price-performance parameters required to make DNA analysis at the nucleotide level part of everyday work life. This objective will be met by the development of a 3rd generation DNA sequencing instrument - the 'Digital Sequencer' - that will sequence DNA by utilising technologies developed for optical disk industry which can be found in any regular DVD or PC. This technology has the added potential of promoting the European Community as one of the leading regions with respect to scientific communities, both academic and industrial.
A successful launch of a high-throughput low-cost DNA sequencing instrument will represent a major catalytic event within many scientific industries, i.e. an explosive growth is foreseen both in terms of 1) number of users and 2) level of use. Hence, the first company to introduce such an instrument will be faced with unparalleled competitive advantages and enormous market potentials arising from industries like health care, drug development, agriculture and aquaculture, and environmental biotechnology.
The current end user market for DNA analysis consists primarily of scientists performing research for drug development companies. However, there are smaller numbers of users within the fields of diagnostics, public research, health care, agriculture, forensic medicine etc. The market is therefore expected to expand once the cost of DNA sequencing is driven down to more affordable levels.
SEVENTH FRAMEWORK PROGRAMME
Digital Sequencing

Use and dissemination of foreground

This report contains a summary description of the dissemination activities work of the Project along with a summary of the used resources. This enables the reader to get an overview of the dissemination in the context of respective work packages.

This section (A) has been structured as follows:

- Each work package is described independently
- The first section contains brief description summary of resources.
- The second section contains brief description of dissemination
- The third section

WP 2 Automated and manual binary conversion

Summary of the use of resources

Work described above towards the listed objectives has been carried out according to the description provided in Annex I. Deviations from the set plan have included the following:

An improved design polymer protocol has been established that enables conversion using fewer cycles (switch from 2- to 5-nt conversion). Pilot experiments with a nanopore have also been carried out using a design polymer. The positioning step of the binary tag-based approach has been re-evaluated, and the binary detection based on 2-nt (and 5-nt) overhang ligations described in Annex I has been replaced by a new candidate method utilising polymerase extensions of labelled (biotinylated) nucleotides. Pilot experiments using the binary extension principle has been demonstrated to be successful.

In general, the work in WP2 has been more challenging than initially expected and a delay in the deliverables could not be avoided. Also, the most recent work has been focused on binary conversion approach that is directly relevant for the aimed integrated sequencing strategy.

During second phase of the Project we have strengthened the team working on the binary conversion process by including a new group to the beneficiary KTH. Dr Russom has joined the Project and has previous experience on miniaturisation of pyrosequencing reactions. Dr Russom has primarily been active WP4-6 but in close interaction with activities in WP2-3. WP resources have merged in the final period with binary polymerase extension sequencing being analysed on glass and DVD surfaces (incl re-iterations, barcoding of samples) using both fluorescens and bead detection

Summary of the dissemination

The results generated in this work will be used both for filing of new IP (binary tag-based protocol; assessment of feasibility initiated) and for scientific publications. To date, the following dissemination activities have been identified:

- Automated purification method based on paramagnetic beads has been developed and published suitable for any next generation sequencing platform.
- Discussion with commercial entities working on DNA nanopores regarding the use of the automated design polymer conversion protocol, with the aim of securing a future licensing deal
- Variable digestion length for the type IIs restriction enzymes. These enzymes are used in various protocols, but the amount of data in the public domain on the ‘slippage’ is very limited. (in revision PLOS One. See WP3)
- File a patent application on the binary tag sequencing approach (assessment of feasibility on-going)
- Selected elements of the binary tag-based conversion protocol was presented at the AGBT meeting in Florida in Feb 2013.
- Scientific presentations were made at several occasions presenting binary sequencing

Publications

PLoS One. 2011 Apr 27;6(4):e19119. doi: 10.1371/journal.pone.0019119.

Large scale library generation for high throughput sequencing.
Borgström E, Lundin S, Lundeberg J.

PLoS One. 2011;6(7):e21910. doi: 10.1371/journal.pone.0021910. Epub 2011 Jul 7.
Scalable transcriptome preparation for massive parallel sequencing.
Stranneheim H, Werne B, Sherwood E, Lundeberg J.

PLoS One. 2014, in revision
Endonuclease specificity and sequence dependence of Type IIS restriction enzymes
Lundin S, Jemt A, Terje-Hegge F., Foam N., Pettersson E., Käller M., Wirta V, Lexow P, Lundeberg J,

WP 3 Addressable binary conversion

Summary of the use of resources
Work described above towards the listed objectives has been carried out according to the description provided in Annex I. Deviations from the set plan have been minor, but include the following:

• Work on target selection and indexing has been included as part of work carried out in this WP. This is required for the binary tag-based sequencing protocol, where extended ('positioned') amplicons are captured to the disk surface using adapter molecules.

• Work on high throughput DNA barcoding (aka as indexing) has been included as part of work done in WP2. This work is important to avoid individual amplification reactions in and thereby increase the capacity of the method which is required in TileSeq assay and to met the enormous capacity of NGS platforms.

Summary of the dissemination
The results generated in this WP has been used at least in the following manner:
• The indexing protocol has been described in various posters on scientific conferences, including the AGBT sequencing conference (major event for DNA sequencing related work)
• The serial compartmentalization assay has been presented in a workshop (Knut and Alice Wallenberg workshop) and will be presented in the next AGBT conference
• A publication on the index tagging and automated protocol has been prepared and accepted
• Commercial exploitation of the TileSeq technology will be assessed
• Article highlight on TileSeq on Genome Web w. interview
• Scientific presentations were made at several occasions presenting TileSeq

Publications

Sci Rep. 2013;3:1186. doi: 10.1038/srep01186.
Hierarchical molecular tagging to resolve long continuous sequences by massively parallel sequencing. Lundin S, Gruselius J, Nystedt B, Lexow P, Käller M, Lundeberg J.

PLoS One. 2013;8(3):e57521. doi: 10.1371/journal.pone.0057521. Epub 2013 Mar 4.
TagGD: fast and accurate software for DNA Tag generation and demultiplexing.
Costea P, Lundeberg J, Akan P.

WP 4 Binary DNA read out on microscope slides

Summary of the use of resources
Work described above towards the listed objectives has been carried out according to the description provided in Annex I. Deviations from the set plan have been minor, but include the following:

Both Plarion (beneficiary 3) and IC (beneficiary 5) have contributed to the results obtained in this WP, although not initially indicated in Annex I. Plarion's work has been overlapping with the

WP5 and hours reported here as the WP4 contribution has been small calculated in hours. IC contribution to this WP is approximately 3 man months and has addressed the modification of the PDMS stamper.

Work on sub-micron precision in the attachment of DBO was transferred to the WP5 where suitable instrumentation, the optical disk drive, is available. Work on the objectives in WP4 therefore focused on development of an attachment chemistry that could be transferred to the optical disk setting.

Summary of the dissemination

The results obtained in this WP are currently planned to be used as follows:

Poster presentation on microcompartment amplification and PDMS modification has been carried out

The pipeline for design of index tags and binary units has been presented on the Beyond the Genome conference in a poster format.

The conversion of DNA into a detectable has been transferred to WP5 and DVD drive based readout

WP 5 Binary DNA read out on optical disks

Summary of the dissemination

The use of results obtained in this WP:

- A patent application has been filed on the LIT technology
- The LIT technique has been orally presented on the Technology World conference in London
- The project web site contains a brief information letter describing the principle of the LIT technology, and that encourages interested parties to contact LingVitae for further discussions. The method is available for licensing.
- The attachment chemistry developed can also be used for attaching cells proteins to the disc surface; this is of particular interest in immunoassays where antibodies can be attached to the disc surface. This is currently being utilised in parallel exploitation activities where the prototype discs and drives are evaluated for use in decentralised diagnostics.

WP 6 - Development and validation of an optical disk based DNA sequencing system

Summary of the dissemination

The prototype drive that has been manufactured enables further development of the sequencing protocol.

It can also be used in parallel to explore other adjacent application areas, for example in low-cost diagnostic settings. These exploitation activities have been initiated.

Lab Chip. 2013 Apr 21;13(8):1578-85. doi: 10.1039/c3lc41360h
Lab-on-DVD: standard DVD drives as a novel laser scanning microscope for image based point of care diagnostics.
Ramachandraiah H[1], Amasia M, Cole J, Sheard P, Pickhaver S, Walker C, Wirta V, Lexow P, Lione R, Russom A.