SIXTH FRAMEWORK PROGRAMME
HORIZONTAL RESEARCH ACTIVITIES INVOLVING SME'S
CO-OPERATIVE RESEARCH

Project acronym:	ArraySBS
Project full title:	Array Based Sequencing by Synthesis
Project no.:	017623
Contract for:	CO-OPERATIVE RESEARCH

Final activity report M1-M30

Period covered:	30 August 2005 to 29 February 2008
Date of preparation:	04 April 2008
Start date of project:	30 August 2005
Duration:	2 years Extension of 6 months
Project coordinator name: Project coordinator organisation name:	Per Johan Ulfendahl/ Joachim Engels Quiatech AB/ Goethe University Frankfurt
Revision:	First Draft

Table of Contents

1. Publishable executive summary : ArraySBS *3*

2. Project objectives and major achievements during the whole project period *6*

2.1 List of participants **6**

2.2 Combined results of the respective periods **6**

2.2.1 Achievements of work package 1 7
2.2.2 Achievements of work package 2 10
2.2.3 Achievements of work package 3 12
2.2.4 Achievements of work package 4 12
2.2.5 Achievements of work package 5 14

3 Work package progress of the period *16*

3.1 Contractual deliverables **16**

3.2. Work package description forms **19**

4 Consortium management *36*

4.1 Organization of scientific collaboration **36**

4.1.1 Kick-off meeting in Frankfurt/Germany (31.08.2005 - 02.09.2005) 36
4.1.2 Second ArraySBS project meeting in Vilnius/Lithuania (01.03.2006 - 03.03.2006) 36
4.1.3 Third ArraySBS project meeting in Uppsala/Sweden (23.08.2006 - 25.08.2006) 37
4.1.4 Fourth ArraySBS meeting in Tartu/Estonia (08.03.2007 - 09.03.2007) 38
4.1.5 Fifth ArraySBS project meeting in Frankfurt/Germany (31.08.2007 - 01.09.2007) 41
4.1.6 Sixth ArraySBS project meeting in Copenhagen/Denmark (11.01.2008) 43
4.1.7 Visits and communication 44
4.1.8 Consortium Management – Problems 44
4.1.9 Major achievements concerning dissemination of knowledge 44

4.2 Use of person-month and resources **45**

5. Other issues *51*

5.1. Safety **51**

5.2. Gender issues **51**

5.2.1 First year 51
5.2.2 Second year 51

1. Publishable executive summary: ArraySBS

Goal of the project

The aim of the ArraySBS project is to provide proof-of-principle for the sequencing-by-synthesis (SBS) setting on an array of primer features, where the 3'-end of the primer is extended with a reversible terminator. This new system will permit cycles of incorporation of one single nucleotide and its immediate identification. Hence it is not based on gel electrophoresis and consequently will improve throughput and reduce costs substantially, while at the same time increase accuracy.

Workpartners

WP1: Oligovation (former Quiatech AB) (Sweden), Goethe University Frankfurt (Germany)
WP2: Fermentas UAB (Lithuania), Goethe University Frankfurt (Germany)
WP3: Silex AB (Sweden), Asper Biotech (Estonia), Oligovation (Sweden)
WP4: Asper Biotech (Estonia), University of Tartu (Estonia)
WP5: Asper Biotech (Estonia), Fermentas UAB (Lithuania), Goethe University Frankfurt (Germany), Oligovation (Sweden), Silex AB (Sweden), University of Tartu (Estonia)
WP6: Goethe University Frankfurt (Germany)

Work packages 1 and 2: Synthesis of dye-labelled 3'-modified reversible terminators (WP1) and identification and improvement of a DNA polymerase (WP2)

a) Objectives:
Initially for the ArraySBS project, the chemical and biotechnological work to create a reversible terminator system was of highest importance. The chemists part of the work (WP1) can be divided into the following tasks:

- Development of a 3'-*O*-protection group which is tolerated by the polymerase and quantitatively cleavable.
- Synthesis of three or optionally four 3'-end blocked dNTPs, testing of all 3'-*O*-modified nucleotides at Fermentas labs regarding their applicability as a reversible terminator.
- Development of a suitable cleavable linker for all reversible terminators that enables the dye-attachment and can be quantitatively cleaved afterwards.

For the enzymologist (WP2) the task was the identification, characterisation, cloning and purification of a DNA polymerase that tolerates the 3'-*O*-modification with high selectivity for the dye-labelled nucleotides A, C, G and T.

b) Work performed:
Until month 6, the 3'-*O*-MethylThioMethyl group (MTM) and the Dithio*Tert*butylMethyl group (DTM) had been used as a blocking group on the 3'-hydroxyl function of a nucleoside. This chemistry had been stopped after month 6 because of the refinancing plans at Quiatech.
The formacetal-based 2-CyanoEthoxyMethyl group (CEM) and the ether-based 2-CyanoEthyl group (CE) were then selected as alternatives to the thioacetal groups. Therefore 3'-*O*-CEM-thymidinyl-5'-*O*-triphosphate and 3'-*O*-CE-thymidinyl-5'-*O*-triphosphate had been synthesised. Fermentas UAB confirmed that the 3'-*O*-CE-thymidinyl-5'-*O*-triphosphate was better incorporated into the DNA template by numerous polymerases than the CEM derivative. The second dNTP bearing the CE-group, 3'-*O*-CE-adenosinyl-5'-*O*-triphosphate, was synthesised at Frankfurt University and successfully tested at Fermentas.
As a result of the first project period, it was found that the cyanoethyl (CE) moiety as a 3'-*O*-capping group for the reversible terminators was incorporated best by the DNA polymerase evolved at Fermentas' lab. This cyanoethyl group can be removed quantitatively within a few minutes using 1 M TBAF/THF solution at 60°C. Based on the TBAF (fluoride) cleavage chemistry, a fluoride cleavable

dye-linker system attached to the base moiety of each terminator was developed during the second project year. This multi-step synthesis led to the development of three complete reversible terminators based on the three nucleobases adenine (A), cytidine (C) and thymidine (T). All three terminators are incorporated by several mutants of the polymerase found and purified during the first project period.

Work package 3: Development of a microfluidic device

a) Objective:

- Design and manufacture of a prototype for a microfluidic device which is applicable for the SBS technology.

b) Work performed:
In cooperation with Asper, Silex AB developed and delivered the first prototype at month 3. As a suitable surface material, gold-coated chips were selected which enable immobilization 5'-thiol-modified oligonucleotide primers.
The second microfluidic prototype was delivered delayed at month 13 (foreseen month 12) due to the changes in chemistry. The chemistry based on the fluoride-cleavable 3'-*O*-cyanoethyl group strongly influenced the design of the second device. The time delay was caused by the discussion about a suitable surface material which should be resistant against the silicon-etching fluoride ions. The second microfluidic devices containing fluoride- resistant Teflon coating were delivered at month 13.
The third microfluidic device prototype was designed and manufactured at Silex. It was delivered at month 23 and is the final version of the microfluidic device that enables the re-sequencing.

Work package 4: Bioinformatics and selection of primers

a) Objective:

- Development of a software for primer design and of a software prototype for analysing sequencing-by-synthesis results.

b) Work performed:
Studies on how DNA primer properties influence the results from primer extension microarrays were made. The analysis was done on experimental data from genotyping microarrays, which is likely to behave very similarly with the developed sequencing-by-synthesis platform.
Advanced statistical methods were used to evaluate which primer properties are responsible for strong fluorescent signal and low failure rate. Based on these tests statistical models for prediction of call rate and signal intensities of extension primers on microarrays could be developed.
A microarray normalization method was developed which can reduce variation between microarrays and helps to design base-calling algorithms during the second year of the project. A web-based software prototype has been created for automatic design of re-sequencing primers with strong signal and low failure rates. After developing an initial software prototype the work was continued with adapting the prototype for re-sequencing. For this purpose several algorithms have been tested and it was found that the most appropriate algorithm for genotyping is the EM algorithm, which uses equal shape and volume of genotype clusters (type EEE and EII). Named algorithms were implemented into BsPred software and adapted for base calling of re-sequencing data.

Work package 5: Development of a re-sequencing assay (WP5)

a) Objective:

- Development of an assay for the p53 gene coding sequence (1218 bp) in a microfluidic reaction device and validation of its results by comparing the results with the ones produced by APEX and Sanger sequencing.

During the first project period Asper's commonly used APEX-based TP53 re-sequencing system glass microarrays with fluorescently labelled dideoxy terminators and ThermoSequenase DNA polymerase was used to evaluate the origin of possible variations in re-sequencing results.

Due to the fluoride-based cleavage chemistry used for the reversible terminators delivered by the contractors of WP1, an urgent need for a suitable chip surface material came up which led to time-demanding research on this task. The reason for this is that glass as the common chip surface material (Silex, Asper) is etched by the fluoride ions in the TBAF/THF-solution. Testing the fluoride resistance of several DNA chips, it was found that CodeLink™ Activated slides are suitable for this chemistry. Due to the low duplex stability in TBAF/THF for the performance of primer extension experiments on these slides, hairpin oligonucleotide probes with 12 or 17 nucleotide linker were designed by Tartu University and delivered to Asper. Both 12 and 17nt linker probes have been spotted on the slides for performing primer extension reactions. One result of the extension experiments was that hairpin oligonucleotide probes carrying a 17 nucleotide linker worked better for primer extension than probes with 12 nucleotide linker. In the second stage of the primer extension experiments, 64 oligonucleotide probes for testing up to three consecutive cycles of terminator incorporation/removal were designed by Tartu University. There is a 3 nucleotide motif after each hairpin. All oligonucleotides have been tested for proper secondary structure. These oligonucleotide probes were used to spot 100 primer microarrays to perform proof-of principle experiments of array SBS procedure.

Main achievements and results of the project

- Based on the very good incorporating and terminating properties of the 3'-*O*-cyanoethyl(CE) group, three 3'-*O*-CE dNTPs (A,C and T) labelled with a fluorophore have been synthesised and tested. The polymerase did also tolerate these nucleotides, which also contain a new invented, fluoride cleavable linker system. The synthesis of these three terminators was optimized and finished successfully.
- Due to the cyanoethyl-group bearing terminators, several mutants of the DNA polymerases that tolerate the 3'-*O*-cyanoethyl(CE)-group as well as the dye-linker system on the base-moiety of the terminator have been evolved, purified and characterized. At the end of the project Fermentas' polymerase accepting the three reversible terminators was optimized. This polymerase can be used for commercial purpose after the project end.
- The activities regarding WP1 and WP2 led to a provisional US patent 60/824453 covering the chemistry, its application for sequencing-by-synthesis, and a kit using this chemistry, enzymes and composition, the incorporation of these by a polymerase and the conversion of the 3'-*O*-protection to a free 3'-hydroxyl group.
- A suitable microfluidic device prototype enabling the SBS procedure could be drafted and manufactured. Also suitable, fluoride stable slides could be found and stable hairpin oligonucleotides were designed. A setup for the 100 primer microarray proof-of-principle was developed.
- A well functioning software tool for array based sequencing-by-synthesis and primer design was developed.

2. Project objectives and major achievements during the whole project period

2.1 List of participants

Part. Role*	Part. Type	Part. no.	Participant name	Participant short name	Country	Date enter project	Date exit project**
CO	SMEP	1	Oligovation (former Quiatech AB)	Oligovation	Sweden	Month 1	Month 30
CR	SMEP	2	Asper Biotech Ltd	Asper	Estonia	Month 1	Month 30
CR	SMEP	3	Fermentas UAB	Fermentas	Lithuania	Month 1	Month 30
CR NCO	RTD	4	Johann Wolfgang Goethe University	JWG U	Germany	Month 1	Month 30
CR	RTD	5	Tartu University	Tartu	Estonia	Month 1	Month 30
CR	RTD	6	Silex Microsystems	Silex	Sweden	Month 1	Month 30

*NCO = New Coordinator since month 13
CR = Contractor
** = The new exit date was set due to the project extension of 6 months

2.2 Combined results of the respective periods

Main project objective

The main objective is to develop methods and components enabling the development of a cost effective and fast DNA array based sequencing-by-synthesis system allowing point-of-care personalised medicine.

Operational objectives

- To develop and synthesise three (optionally four) dNTPs (A, C, G, T) with blocked 3'-end and labelled with a fluorescent dye, i. e. three complete reversible terminators. **(WP 1)**

- To identify and develop a DNA polymerase that accepts the dye labelled reversible terminators. The best DNA polymerase that accepts fluorescent dye labelled reversible

terminators will be identified and cloned not later than month 14, this polymerase has also to be improved and purified during the second project period. **(WP 2)**

- To develop a microfluidic device prototype that allows array SBS reactions by integrating incubation steps at different temperatures with controlled agitation of reaction components over the entire area of a microarray, including washing and regeneration procedures. **(WP 3)**

- To develop bioinformatic tools for primer design and for analysis of sequencing-by-synthesis generated data. The software prototype for handling the image data from SBS experiments should be done until month 14. Besides this objective, the base-caller should be modified and suitable primers for the p53-resequencing should be identified. As an outcome and milestone of this, software for primer prediction and sequence data handling should be delivered at the end of the project. **(WP 4)**

- To validate the technology through the re-sequencing assay of a coding sequence of a gene. After 18 months the array SBS procedure should have reached a proof-of-principle stage with up to 100 primer features on a microarray. During month 30 the p53 gene (the coding region) should be sequenced with the array SBS procedure. **(WP 5)**

2.2.1 Achievements of work package 1

Tasks and achievements:

- *task1.1: Synthesis and delivery of the first 3'-O-modified reversible terminator (D3 month 1)*

  Achievements: The first main project objective for the chemists at Quiatech as well as at the JWG University was the development of a suitable 3'-*O*-protection group. Therefore Quiatech as one of the contractors from WP1, which had already experience in thioacetal chemistry (see M. Kwiatkowski, U.S patent US2004/0175726: COMPOUNDS FOR PROTECTING HYDROXYLS AND METHODS FOR THEIR USE and M. Kwiatkowski European patent EP 0808320: NOVEL CHAIN TERMINATORS, THE USE THEREOF FOR NUCLEIC ACID SEQUENCING AND SYNTHESIS AND A METHOD OF THEIR PREPARATION), recommended to introduce the MethylThioMethyl blocking group (MTM) and the DithioTertbutylMethyl blocking group (DTM) at the 3'-hydroxyl function of a nucleoside.
  According to that idea, the first potential reversible terminators 3'-*O*-MTM-thymidinyl-5'-*O*-triphosphate was synthesised. The triphosphate synthesis appeared time-demanding and laborious. Two purification steps have to be carried out (ion exchange FPLC and RP-HPLC), because of the high purity that is needed for the polymerase acceptance tests.
  The first 3'-*O*-modified reversible terminator, could be delivered until month 1.

- *task1.2: Synthesis and delivery of a biotinylated 3'-O-modified reversible terminator (D5 month 2)*

  Achievements: The biotinyl-labelled 3'-*O*-MTM-thymidine was synthesised at Quiatech. For the attachment of the biotin group a propargylamine-functionality was introduced at the base-moiety of the nucleoside. Subsequent conversion into the 5'-*O*-triphosphate and two-step purification (FPLC, RP-HPLC) led to the 3'-*O*-MTM-5-propargylamine-uridinyl-5'-*O*-triphosphate as the first complete reversible terminator and was delivered in month 2.

- *task1.3: Synthesis of the second 3'-O-modified reversible terminator (D8 month 9)*

  Achievements: The second 3'-*O*-modified reversible terminator, 3'-*O*-DTM-thymidinyl-5'-*O*-triphosphate was synthesised until month 4 at Quiatech and provided to Fermentas.

- *Results of the polymerase acceptance and cleavage tests of the first and second reversible terminator*

  Achievements: Although the delivered triphosphates were purified two times by the chemist, Fermentas had to accomplish an additional enzymatic purification step (Mop-Up) to remove last traces of 3'-*O*-unmodified nucleotides. For the incorporation tests a radioactively labelled DNA template with a poly-A overhang was used to coevally check the termination properties of the incorporated 3'-*O*-blocked nucleotide.
  Cleavage tests of the 3'-*O*-blocking groups were accomplished at Quiatech by incorporating the modified nucleoside via its amidite into a primer sequence.
  The 3'-*O*-MTM-thymidinyl-5'-*O*-triphosphate was accepted and incorporated into the DNA template by several polymerases, but the MTM-function could not be removed quantitatively.
  The DTM-moiety of 3'-*O*-DTM-thymidinyl-5'-*O*-triphosphate showed instability under the conditions of the polymerase incorporation step. In fact the acceptance by the tested polymerases was good but the removal of this group proceeded too easily.
  The biotinylated 3'-*O*-MTM-thymidinyl-5'-*O*-triphosphate was first synthesised in only small amounts and had to be synthesised on larger scale to have enough material for the polymerase acceptance tests.
  Before the synthesis of the biotinylated nucleotide could be repeated and before polymerase acceptance and stability of the thioacetal-based 3'-*O*-blocking groups (MTM and DTM) potentially could have been improved, other problems occurred in the first 8 months of the SBS project and the researchers could not continue working on the base of thioacetal chemistry. The reasons for that were refinancing plans at Quiatech during springtime 2006. The company, which is the owner of the thioacetal chemistry containing patents as mentioned above, sold these patents to a third party.

- *task1.4: Synthesis of the third 3'-O-modified reversible terminator (D8 month 9)*

  Achievements: Quiatech found a new 3'-*O*-blocking group, the novel fluoride cleavable CyanoEthoxyMethyl (CEM) formacetal group and used it for synthesising the third 3'-*O*-modified reversible terminator. In month 6, 3'-*O*-CEM-thymidinyl-5'-*O*-triphosphate had been synthesised both at JWG University and Quiatech and was sent to Fermentas immediately.

- *Optional task1.5: Synthesis of the fourth 3'-O-modified reversible terminator (D8 month 9)*

  Achievements: Between month 6 and 9, the chemists from Quiatech and JWG University came up with the synthesis of another potential reversible terminator, which was based on the CyanoEthyl (CE) ether group. The 3'-*O*-CE modification could be easily installed and the nucleoside was subsequently converted into the 5'-*O*-triphosphate and sent to Fermentas at month 8.

- *Results of the polymerase acceptance and cleavage tests of the third and fourth reversible terminator*

  Achievements: The 3'-*O*-CEM-thymidinyl-5'-*O*-triphosphate was found to be accepted by several polymerases.
  Cleavage tests at Quiatech also provided promising results. The cleavage experiments were carried out with 1 M TetraButylAmmoniumFluoride (TBAF) in dry THF and the reversion to the 3'-hydroxyl group proceeded quantitatively in 2 minutes at 37°C.
  Fermentas confirmed that the substrate bearing the 3'-*O*-CE group was even better incorporated into the template DNA than the 3'-*O*-CEM derivative.
  It was found by Quiatech that the cleavage of the CE group took place under the same conditions (fluoride ions) that are used for the CEM group.

- *task: Synthesis of biotinylated 3'-O-CE modified reversible terminator*

Achievements: In order to evolve polymerases that can recognize the reversible terminators even more efficiently, Fermentas needed a dye-labelled potential reversible terminator. Therefore the chemists at JWG University synthesised the 3'-*O*-CE-5-propargylamine-uridinyl-5'-*O*-triphosphate. At Fermentas amino modified nucleotide was labelled with biotin as a dye-prototype (month 12).

- *task: Synthesis of second 3'-O-CE modified reversible terminator*

  The second dNTP, which was synthesised at JWG University and sent to Fermentas, was the 3'-*O*-CE-adenosinyl-5'-*O*-triphosphate (month 12).

- *task1.6: Synthesis and delivery of the first dye-labelled 3'-O-modified reversible terminator (D12 month 12)*

  Achievements: In order to attach a fluorescent dye to the base the chemists are working on the modification of the above mentioned propargylamine-modified nucleotides with different fluorophores (dyes) bound to a suitable cleavable linker. This task and its deliverable is postponed to month 15 as described in detail in section 3.2 work package 1, "Main achievements" and "Problems".

- *task1.7: Synthesis of the remaining 3'-O-modified reversible terminators with dyes (D21 month 24)*

  Achievements: For a proof-of-principle of the re-sequencing, it was decided to prepare three complete reversible terminators, that means the dNTPs of A, C and T bearing the 3'-*O*-blocking group and a dye-linker system with a specific dye for each nucleobase. Based on the results of the first project year, it was found that the cyanoethyl-function as a 3'-*O*-blocking group was best accepted by several polymerases. Besides that, our idea for a cleavable linker at the base moiety of the nucleotide was to use the cyanoethyl-motif in the linker. With this cyanoethyl-group it should be possible to cleave the 3'-*O*-capping and the dye-linker system under the same conditions. During the second period, the chemists strongly focussed on the development of a glycol-type linker as well as finding the optimal cleavage conditions for both 3'-*O*-capping and linker. We found that both moieties could be quantitatively removed using a standard solution of 1 M TBAF/THF (40 equivalents TBAF to 1 equivalent nucleotide) at 60°C. The cleavage time under those conditions were 6 minutes for the 3'-*O*-cyanoethyl-group and less than 1 minute for the cyanoethyl-containing cleavable linker. With this cleavage optimization and with the establishment of a suitable synthetic strategy for the three reversible terminators, we could achieve the deliverable D21 at the end of the second project period.

- *Optional task1.8: Synthesis of fourth 3'-O-modified reversible terminators with dye (D21 month 24)*

  Achievements: Besides the synthesis of the three reversible terminators, the 3'-*O*-cyanoethyl-dNTPs with the fluoride-cleavable linker and dye (A = Texas Red; T = Fluoresceine; C = Cy3), the fourth reversible terminator (i. e. G containing the Cy5-labelled fluoride-cleavable linker and the 3'-*O*-CE group) could not be synthesised within the timeframe of the project.

- *Results of the polymerase acceptance and cleavage tests of the three complete reversible terminators bearing a cleavable linker, a specific dye and the 3'-O-blocking group*

  Achievements: At the end of the second project year, three complete reversible terminators were available for the polymerase acceptance tests: 3'-*O*-cyanoethyl-dNTPs A, C and T with the fluoride-cleavable linker and the base-specific dye. All three terminators are well incorporated into the DNA template, the linker and the 3'-*O*-cyanoethyl group can be cleaved at 60°C with an excess of 1 M TBAF/THF solution.

- *task1.9: Set up large scale and purification methods of the 3'-O-modified reversible terminators (D21 month 24)*

  Achievements: As a result of the syntheses of all three complete reversible terminators, the purification methods of these labelled nucleotides were optimized during the preparation of these in several batches. It was found that the nucleotides showed very good purity after two purification steps (RP-FPLC and RP-HPLC with fluorescent detection). The optimized labelling reaction led to sufficient amounts of Fluoresceine-labelled dTTP, TexasRed-labelled dATP and Cy3-labelled dCTP for structural analysis (^{1}H and ^{31}P-nuclear magnetic resonance spectroscopy) and biological tests.

2.2.2 Achievements of work package 2

The project objectives of WP2 for this period were the identification, characterization, cloning and purification of a DNA polymerase that incorporates reversible terminators with minimal discrimination between the dye-labelled nucleotides A, C, G, and T. This part of the project was mostly covered by the work of the workpartner Fermentas in collaboration with Quiatech and JWG University.

Tasks and achievements:

- *task2.1: Development of an assay protocol describing the technique that allows detection of reversible terminator incorporation by purified DNA polymerases (D4 month 2)*

  Achievements: A protocol describing the "Development of technique for detection of reversible terminator incorporation by purified polymerases" as well as a protocol describing "Development of technique for detection of reversible terminator incorporation by crude or partially purified cell extracts" was created until month 4.

- *task2.2: Identification of a DNA polymerase that incorporates reversible terminators and its purification (D9 month 10)*

  Achievements: To test the ability of polymerases to incorporate modified dNTPs into DNA, a set of enzymes (more than 30 members) representing all families of DNA polymerases was collected. Model compounds, including dideoxy-, 3'-*O*-acetate- and 3'-*O*-methyl-dTTP were used for fine-tuning of screening assay. Fermentas found that polymerases exhibiting prominent nucleolytic activity are poorly compatible with the assay, unless specific measures are taken. Full range screening program was executed using four modified dTTPs featuring 3'-*O*-Dithio*Tert*butylMethyl (DTM), 3'-*O*-MethylThioMethyl (MTM), 3'-*O*-CyanoEthoxyMethyl (CEM) and 3'-*O*-CyanoEthoxy (CE) protective groups. DTM protective group has been found to be unstable during enzymatic incorporation into DNA; thus, under the agreement reached by project members it has been dropped from the list of possible terminators. Screening performed using 3'-*O*-MTM-dTTP was highly successful; however, efforts taken to remove the protecting group (keeping DNA structure intact) failed. It should be noted, however, that this compound has a potential to be used in those applications where only one step of modified dNTP incorporation is required (i.e. investigation of single nucleotide polymorphism). CEM and CE protecting groups proved to be stable both in free (nucleotide) and incorporated (into DNA) form and to be incorporated into DNA by numerous polymerases. Fermentas found two polymerases which are able to incorporate 3'-*O*-CEM-dTTP and six polymerases which are able to incorporate 3'-*O*-CE-dTTP.

- *task2.3: Purification and cloning of best polymerase (D14 month 14)*

  Achievements: The genes of the selected polymerase for improvement were available and owned by Fermentas. Therefore no work was necessary for the cloning and purification and the improvement (task2.4) could start immediately.

- *task2.4: Improvement of the DNA polymerase by gene modification (D19 month 24)*
  Achievements: Fermentas decided to design an *in vitro* evolution system, which could be used not only to improve the incorporation of modified dTTP by any polymerase, but also explored for evolution of many other enzymatic properties of polymerases. The system is based on the well-known ribosome display selection widely used for protein affinity selection. In order to make ribosome display selection system suitable for polymerase evolution, it is necessary to modify the traditional scheme by attaching covalently the DNA substrate to the 5'-end of mRNA before the translation step. Then mRNA with the attached DNA substrate will be *in vitro* translated, purified and incubated with modified dTTP which contains biotin attached to the base. Molecules of polymerase able to incorporate the biotinylated modified nucleotide will be fished out using streptavidin beads and their co-isolated mRNA molecules amplified by RT-PCR. There are 4 major achievements in this selection approach: (1) *in vitro* translation of active polymerase under investigation, (2) attachment of DNA substrate to mRNA, (3) general optimization of experimental setup, and (4) construction of mutant library. At the moment two initial objectives are fulfilled. Final optimisation of the whole experimental setup of polymerase selection and construction of its mutagenic library is on the way. Based on rationally designed mutants targeting surroundings of the polymerase's active side and on crystal structures of those mutants, it was possible to improve the terminator accepting polymerase genetically. Besides that, mutations allowing enhanced flexibility of non-structured loops within this trajectory were introduced. Potential candidates were partially purified and tested for affinity for 3'-CE modified nucleotides and primer extension velocity, using these nucleotides. As an alternative to this polymerase improvement technique, a colony-based screening assay suitable for detection of mutants with higher incorporation rate of modified nucleotides was developed. The screening is based on incorporation of 3'-CE modified nucleotides (dye labeled on the base moiety) into the immobilized double-stranded DNA substrate. This step is followed by immune enzymatic detection of the dye. The approach proved itself as both highly selective and high-throughput, allowing screening through up to 30-50 thousand colonies on a single plate.

- *task2.5: Improving the DNA polymerase by gene modification and task2.6: Purification of the improved polymerase (both D19 month 24)*

  Achievements: Altogether, a number of mutants were obtained by rational design and by colony-based screening. Ten mutants were purified and tested. The best candidates featured up to 20% increase in efficiency of incorporation of 3'-CE-modified nucleotides bearing fluorescent label, based on the results obtained from experiments executed in microchip mode

Main achievements from work packages 1 and 2

The chemical and biotechnological work, done by the contractors of WP1 and WP2 during the first period of 12 months, lead to a provisional US patent 60/824453 covering the chemistry, its application for sequencing-by-synthesis, and a kit using this chemistry, enzymes and composition, the incorporation of these by a polymerase and the conversion of the 3'-*O*-protection to the free 3'-hydroxyl group.
The intensive collaboration among the researchers of Fermentas and JWG University led to three complete reversible terminators, which are all incorporated by several mutants of the DNA polymerase. The 3'-*O*-cyanoethyl function of these terminators is fluoride-cleavable, and also the dye-linker system can be removed under the same conditions. This is a proof of the demanded reversibility for these terminators. The enzymologists developed a terminator accepting polymerase

on the one hand, but they also made big efforts in the technique development of improving potential reversible terminator- accepting polymerases.

2.2.3 Achievements of work package 3

Silex as the main contractor of WP3 had to develop at least three prototypes of a reaction device that allows array SBS reactions by integrating incubation steps at different temperatures with controlled agitation of reaction components over the entire area of a microarray, including washing and regeneration procedures. The microfluidic reaction device should also be able to perform the re-sequencing during 10 cycles. The design and manufacturing of the devices was strongly influenced by the deprotection chemistry applied for the terminators and also by the demanded template immobilization conditions.

Tasks and achievements:

- *task3.1 and 3.2: Design and manufacture of the first microfluidic device prototype (D7 month 3)*

  Achievements: Until month 6, Silex designed in cooperation with Asper (one of the contractors from WP4) the first microfluidic device prototype. The device consists of a microfluidic platform (reaction chamber) etched on the frontside, with backside holes microfluidic interconnections. Furthermore a recess in the silicon is made around the reaction chamber. This recess is used to place an O-ring that maintains permeability while placing the lock on the reaction chamber. This microfluidic device was then manufactured and delivered to Asper until month 12. In parallel, Silex delivered diced slides with different coatings for evaluation of which coating is best suited for future prototypes (bio-compatibility).

- *task3.3 and 3.4: Design and manufacture of the second microfluidic device prototype (D11 month 12)*

  Achievements: Because of the change from thioacetal to acetal/ether chemistry (regarding results and achievements from WP1), the second microfluidic prototype could not be delivered until month 12. This delivery date is postponed to month 14/15. The reason for that delay is the discussion about a suitable surface material for the second prototype. This material should be resistant against the silicon-etching fluoride ions that are used during the cyanoether cleavage-procedure. At the 3rd SBS meeting in Uppsala/Sweden in month 12, it was decided to use a gold surface for immobilization of the DNA template in the reaction chamber. The second microfluidic device was delivered in month 15.

- *task3.5 and 3.6: Design and manufacture of the third microfluidic device prototype (D18 month 23)*

  Achievements: A third microfluidic device prototype design was designed and manufactured at Silex. The third microfluidic device was delivered at month 23. It was decided (meeting in Copenhagen, 11th January 2008) that it was not necessary to go on and include extra integrated features such as heaters to the device since the point to test it would not be reached within the timeframe of the project.

2.2.4 Achievements of work package 4

The development of bioinformatic tools for identification of primers for chip sequencing as well as software for analysing sequencing-by-synthesis results was mainly the project objective of Tartu University. In order to apply the invented software for the Array SBS re- sequencing procedure, Tartu University had a strong collaboration with Asper for developing software for primer prediction and sequence data handling for enabling the p53-re-sequencing.

Tasks and achievements:

- *task4.1 and 4.2: Development and delivery of a method and a first version of software for primer design of SBS microarrays (D10 month 11)*

  Achievements: Maido Remm and his researcher group at Tartu University first started statistical modelling of primer properties on a similar microarray platform. Initial analysis was based on data from APEX microarrays. APEX is a four-channel microarray genotyping platform that is using primer extension (http://www.asperbio.com/process.htm). In the first stage of the project different normalisation methods were investigated. Signals from each microarray and each channel were normalised by different methods to achieve comparable signal values in each channel of each microarray. The effect of normalisation was tested by measuring array-specific variation component in MANOVA. The best normalisation was achieved by the following formula:
  norm_signal = log2 (raw_signal / 90%_quantile of all signals in given channel)
  With this normalisation method, it was possible to reduce microarray specific variation from initial 23% to nearly 0%. This method and software prototype (1st version) for primer design had been delivered in month 11.

- *task4.3: Improvement of the parameter settings of the method and the 1st software prototype (D15 month 14)*

  Achievements: In the second stage of their work, the researchers from Tartu University tried to find the causes of primer-specific variations of signals. By identifying such factors it would be possible to predict and prevent primers that generate too low signal intensities or too low call rates. This was a necessary step before reliable primer design software for any given microarray could be created. To correlate primer properties with their signal intensities, 36 sequence-related factors for each primer were calculated. Generalised Linear Models were used then to find out which of these sequence-related factors influences signal intensity and call rate of a given primer. The results indicated that the most important factors were the deltaG of the last 16 nucleotides from the primer 3'-end and the strength of secondary structure within the primer sequence. However, these factors explained only approximately 20% of primer-specific variations; the other part of variations still remains unexplained. Now the group at Tartu University started additional searches for causes of variation using combinations of nucleotides within and around the primer 3'-end.

- *task4.4: Modify Base Caller for ArraySBS application; task4.5: Identify primers for p53 re-sequencing and task4.6: Analyse data from instrument; all three tasks belong to deliverable D22 (month 24)*

  Achievements in task4.4: High-throughput genotype analysis needs automatic tools to manage image data and to make correct predictions of genotypes. The idea of base calling is to convert microarray signal intensities into possible genotypes. Base calling procedure includes several steps such as recording signal intensities of experimental data, normalization of signal intensities between channels and arrays, training of genotype models based on training data and predicting genotypes given new experimental data. The base calling method should be highly accurate and reliable. In this project a particular need was to develop a procedure for re-sequencing base calling. The development of a base calling software and its optimization was achieved in four steps:
  1. Research on experimental data for testing and a method for comparing algorithms;
  2. Examination of the best SNP genotyping algorithm;
  3. Improvement of the best SNP genotyping algorithm;
  4. Adaption of the best solution for re-sequencing;

As an outcome of this work, the optimal base-calling algorithm was implemented into software called BsPred. The web interface for BsPred was implemented with BsPred_common model and is available at http://bioinfo.ut.ee/bspred/.
Achievements in task4.5 and task4.6: The principles of primer design software were described in the first periodic activity report (D10). Within the second project period, the formula for prediction of call rates and signal intensities of primers was finalized. As a final part of this study, re-sequencing primer design software called "SBS Designer" was created for prediction of call rate and signal strength of candidate primers. This program is able to produce a list of primers that can be used to re-sequence given DNA region, accompanied by statistics, such as predicted call rate that is calculated using the prediction model, melting temperature (Tm), length, genomic positions etc. The invented software is applicable for a sequencer since bioinformatics analyses were based on the results of APEX ARGACE and ABCR genotyping microarray experiments conducted previously by Asper Biotech Ltd.

2.2.5 Achievements of work package 5

The main project objective for the contractors of WP5, Asper Biotech, is the development of an assay for the p53 gene coding sequence (1218 bp) in a microfluidic reaction device and the array's validation by comparing the results with the ones produced by APEX and Sanger sequencing.
In this first period of the project Asper Biotech, who is a contractor of WP4 and WP5 has been mainly doing research for the deliverables to be completed in the second period and working together with Silex and Tartu University. Asper was involved in designing the first prototype of microfluidic device in collaboration with Silex. From different coating materials of microfluidic chamber offered by Silex, Asper first selected Teflon as a suitable coating based on its chemical resistance.
The research for microarray carrier material became complicated because of the changes of the 3'-*O*-blocking group and thus of the cleavage conditions. The use of fluoride posted significant challenges to Asper to find surfaces and technologies working with the enzymes and new reversible terminator technology developed by the other partners. From silicon chips coated by Silex with the following protective materials: thermal oxide, nitride, C_4F_8, aluminium and gold, the gold-coated chips were then selected for further tests in regard to 5'-thiol group modified oligonucleotide primer binding properties. The primer binding stability resistance tests against chemicals used for terminator group reversal are currently under way.

Tasks and achievements:

- *task5.1 and 5.2: Development of SBS method on APEX chips and the implementation of the first SBS settings (D16 month 18)*

  Achievements: Asper's current APEX-based TP53 re-sequencing system glass microarrays with fluorescently-labelled dideoxy terminators and ThermoSequenase DNA polymerase was used to evaluate the origin of possible variations in re-sequencing results. The DNA fragments used for APEX were amplified from one individual DNA three times in order to create sufficient amount of material for three independent analyses. The amplifications, purifications and pre-APEX treatments were carried out on different days. The APEX reactions were performed using Asper's standard conditions; the analyses of results could successfully be completed at Tartu University Bioinformatics department by the selection of best possible re-sequencing primer candidates. Asper also ran tests to understand the factors influencing the design of the ArraySBS system microchips by applying automated hybridization station HS-400 from Tecan AB.

- *task5.3: Proof-of-principle of the array SBS procedure with a 100 primer microarray (D16 month 18)*

Achievements: During the second project period, a series of experiments were performed in order to find out a suitable DNA chip material resistant to fluoride-based deprotection chemistry. For this purpose different DNA chip carrier materials were tested. It was found that CodeLink™ Activated Slides were the most appropriate DNA chips regarding the fluoride-caused surface etching and the reliable template immobilization.

Several primer extension reactions with hairpin oligonucleotides designed by Tartu University and branched hairpin oligonucleotides delivered by Oligovation were carried out. It was found that normal oligonucleotides carrying a 17 nucleotide linker worked best for primer extension.

Further 64 oligonucleotides for testing up to three consecutive cycles of terminator incorporation/removal were designed by Tartu University partner. There is a 3 nucleotide motif after each hairpin. All oligonucleotides have been tested for proper secondary structure. These oligonucleotide probes were used to spot 100 primer microarrays; these experiments are still in progress and could not be finished within the time frame of the project.

2 Work package progress of the period

3.1 Contractual deliverables

The table below presents the status of the contractual deliverables as defined in the Description of Work. The status of internal deliverables is given in each work package in sections 2.2. and 2.3.

Del. no.	Deliverable name	WP no.	Lead participant	Estimated person-months	Nature	Dissemination level	Delivery date (proj. month)
D1	A functioning public web page established.	6	Quiatech	0,25	O	PU	1
D2	A functioning internal web page established.	6	Quiatech	0,25	O	CO	1
D3	First 3'-end modified reversible terminator delivered.	1	JWG U Quiatech	2	P	CO	1
D4	A protocol describing the technique that allows detection of reversible terminator incorporation by crude or partially purified cell extracts and by purified DNA polymerases.	2	Fermentas	4	R	CO	2
D5	First 3'-end modified reversible terminator with the attached biotin group delivered.	1	JWG U Quiatech	4	P	CO	2
D6	A project leaflet	6	Quiatech	0,25	O	PU	cancelled[1]
D7	A first microfluidic device prototype adapted to the existing Asper detection platform.	3	Silex	3	P	CO	3
D8	Second and third (optionally fourth) reversible terminator (A,C,G,T) delivered.	1	JWG U Quiatech	14	P	CO	9
D9	Identification of a DNA polymerase that incorporates reversible terminators and its purification.	2	Fermentas	25	D	CO	10

[1] The project was announced with the press release 24th August 2005 at the Quiatech website and by Tartu University. As a result of this, two third parties contacted Quiatech and began to discuss business opportunities. In order to protect confidential material (details in chemistry etc.) no project leaflet had been written.

D10	Methodology and software for primer design of SBS microarrays.	4	Tartu U.	14	P	CO	11
D11	A second microfluidic device prorotype with, for example, channels, heaters, mixer and valves adapted to the existing Asper detection platform.	3	Silex	5	P	CO	12 postponed 14-15[2]
D12	First reversible terminators (A,C,G,T) with a dye delivered.	1	JWG U	4	P	CO	12, postponed 21[3]
D13	An intermediate plan for using and disseminating knowledge	6	Quiatech	0,25	R	CO	12
D14	Purification of the best identified DNA polymerase and cloning of the respective gene.	2	Fermentas	13	D	CO	cancelled[4]
D15	Software prototype for handling image data from SBS experiments.	4	Tartu U.	12	P	CO	14
D16	Proof-of-principle of array SBS procedure with 100 primers	5	Asper	24	R	CO	14, postponed 30[5]

[2] The start of chip manufacturing task had a time delay of approximately two months, depending on the late design input within the project. Therefore this deliverable was postponed to the second period.

[3] Initial changes in chemistry and the development of the best 3'-*O*-blocking group led to an estimated time delay of 3 months (month 15) for this deliverable. Afterwards further problems with labelling the triphosphate and purification of the final reversible terminator led to an additional delay of 6 months (month 21).

[4] As the genes of the polymerase of choice were available at Fermentas no cloning experiments had to be carried out.

[5] Because of significant changes from initial plans regarding the chemistry used for terminator and fluorescence group removal the time delay for this deliverable was first estimated to month 18. Further stability problems of slides, immobilization and duplexes caused an additional delay to month 30.

D17	Non-confidential project summary.	6	New coordinator	1	R	PU	14
D18	A third microfluidic device prototype with integrated features adapted to the existing Asper or alternative detection platform and optionally other detector platforms.	3	Silex	3	P	CO	23
D19	Isolation of the best improved version of DNA polymerase and its purification.	2	Fermentas	74	D	CO	24, postponed 29
D20	The whole p53 coding sequence microarray tested with array SBS procedure.	5	Asper	18	D	CO	Not completed within the timeframe of the project[6]
D21	The best reversible terminator-dNTP (A, C, G, T) with one different dye attached to each in "large" amount delivered.	1	JWG U	12	P	CO	24, postponed 28-29
D22	A software for primer design and software for handling image data from SBS experiments	4	Tartu U.	16	P	CO	24
D23	Non-confidential project summary	6	New coordinator	1	R	PU	31
D24	Plan for using and disseminating knowledge	6	New coordinator	1	R	CO	31

[6] Significant changes from initial plans in regard to chemistry used for terminator and dye-linker group removal led to a big time delay which retarded the re-sequencing of the p53 gene within the timeframe of the project.

3.2. Work package description forms

SYNTHESIS OF DYE LABELLED 3'-*O*-MODIFIED REVERSIBLE TERMINATORS

Work package number	**WP 1**	**Start date or starting event:**				0	
Participant Short Name	JWG U	Quiatech					
Person-months per participant	24	12					

Objectives and tasks

Development and synthesis of at least three of the four 3'-end blocked dNTPs (A/C/G/T) with at least one labelled with a fluorescent dye, where the 3'-end blocking group can be reverted to a 3'-hydroxyl group under mild conditions.

task1.1: Synthesis of first 3'-*O* modified reversible terminator (3'-*O*-MRT) (D3)

task1.2: Synthesis of biotinylated 3'MRT (D5)

task1.3: Synthesis of second 3'MRT (D8)

task1.4: Synthesis of third 3'MRT (D8)

task1.5: Optionally synthesis of fourth 3'MRT (D8)

task1.6: Synthesis of 3'MRT with dye (D12)

task1.7: Synthesis of the remaining 3'-*O*-modified reversible terminators with dyes. (D21)

task1.8: Synthesis of (optionally) fourth 3' -*O*-modified reversible terminators with dyes. (D21)

task1.9: Set up large-scale and purification methods of the 3'-*O*-modified reversible terminator. (D21)

Deliverable D21: All four reversible terminators with one different dye attached (D21)

Achievements on tasks

task1.1 and task1.3:

The first main project objective for the contractors of WP1 was to prepare 3'-*O*-modified nucleotide triphosphates labelled with a fluorescent dye located either at the base or the terminating group itself (see deliverables D3 and D5).

Because of Quiatech's experience in thioacetal chemistry (see M. Kwiatkowski, US patent US2004/0175726 and European patent EP0808320), it was decided to use both the MethylThioMethyl (MTM) as well as the Dithio*Tert*butylMethyl (DTM) group as a 3'-*O*-blocking group. As a model compound for the four nucleobases A, C, G and T, thymidine was chosen. After protecting the free 5'-hydroxyl function with a suitable protecting group like the MonoMethoxyTrityl (MMTr) group, the MTM-group was introduced *via* a radical reaction with dimethylsulfide and benzoyl peroxide to the 3'-position of thymidine. Regeneration of the 5'-OH-function with 10% *Para*ToluoylSulfonic Acid (PTSA) and conversion of 3'-*O*-MTM-thymidine into its corresponding 5'-*O*-triphosphate by using triazole activated phosphorous oxychloride and bis(*n*-butylammonium)pyrophosphate delivered the first 3'-end modified reversible terminator in month 1 (Deliverable D3).

The triphosphate synthesis appeared time-demanding and laborious. Two purification steps have to be carried out (ion exchange FPLC and RP-HPLC), because of the high purity that is needed for the polymerase acceptance tests.

Besides the MTM-group for capping the 3'-hydroxylfunction, the chemists at Quiatech and JWG University also synthesised 3'-*O*-DTM-thymidinyl-5'-*O*-triphosphate using the following procedure:

As a starting compound, again thymidine was chosen. 3'-*O*-MTM-5'-*O*-MMTr-thymidine was then prepared. The 3'-*O*-MTM-function of this compound was then activated with sulfuryl chloride to react subsequently with potassium thiotosylate and *tert*butylmercaptane to give 3'-*O*-DTM-5'-*O*-MMTr-thymidine. Removal of the MMTr-group and subsequent triphosphate synthesis and purification (two steps) led to 3'-*O*-DTM-thymidinyl-5'-*O*-triphosphate as a second potential

reversible terminator (month 4).
Both compounds were sent to Fermentas then for starting the polymerase acceptance tests (see in section "Main achievements" WP2).
Further cleavage experiments of the 3'-*O*-MTM group were accomplished at Quiatech by incorporating the modified nucleoside *via* its amidite into a primer sequence.
The group was expected to be cleaved under slightly acidic conditions. During the time we could work with the thioacetal chemistry it was not possible to find suitable conditions for a quantitative deprotection without affecting the rest of the DNA strand (compare problems).

task1.2:
After the first tests Fermentas asked for a biotinyl-labelled 3'-*O*-MTM-thymidinyl-5'-*O*-triphosphate as a model compound for a dye-labelled reversible terminator. For the synthesis of that model compound 5-Iododeoxyuridine was used and the first step was the introduction of the 3'-*O*-MTM group as described above.
To create a site for the attachment of the biotin group a propargylamine functionality had to be established at the base-moiety (thymidine-moiety) of the nucleoside by using the palladium catalysed Sonogashira coupling. Subsequent conversion into the 5'-*O*-triphosphate and two-step purification (FPLC, RP-HPLC) led to the 3'-*O*-MTM-5-propargylamine-thymidinyl-5'-*O*-triphosphate. Attaching the biotin to the propargylamine moiety led to the first complete reversible terminator.
So the first 3'-end modified reversible terminator with the attached biotin group was delivered in month 2 (Deliverable D5) but the investigations of this compound were also stopped by the changes at Quiatech (see problems).

task1.4 and task1.5:
After the usage of the thioacetal based groups (MTM and DTM) was not possible anymore we developed a new 3'-*O*-blocking group using the novel fluoride cleavable CyanoEthoxyMethyl (CEM) formacetal group. This process was based on activation of the 3'-*O*-MTM derivative with sulfuryl chloride and reaction of the formed chloromethyl function with cyanoethanol. We further adopted the Ludwig-Eckstein procedure for the synthesis of the nucleoside 5'-*O*-triphosphate and delivered the third reversible terminator in month 6.
The 3'-*O*-CEM-thymidinyl-5'-*O*-triphosphate was found by Fermentas to be accepted by several polymerases. Cleavage testes at Quiatech also provide promising results. The cleavage experiments were carried out with 1 M **T**etra**B**utyl**A**mmonium**F**luoride (TBAF) in dry THF and the reversion to the 3'-hydroxyl group proceeded quantitatively in 2 minutes at 37°C.
Next we synthesised a reversible terminator based on the 3'-*O*-CyanoEthyl (CE) protection group. Here the direct reaction of 5'-*O*-benzoyl protected thymidine with acrylonitrile and caesium carbonate as base yielded a Michael adduct, which was after 5'-*O*-deprotection with sodium hydroxide converted subsequently into the appropriate triphosphate by using the Ludwig-Eckstein procedure. So the fourth reversible terminator could be delivered in month 8.
Fermentas confirmed that this substrate was even better incorporated into the DNA template than the CEM derivative. In order to evolve polymerases that can recognize our reversible terminators even more efficiently, Fermentas needed a biotin modified thymidine reversible terminator. Therefore we synthesised the 3'-*O*-CE-5-iodo-deoxyuridine and substituted the iodo-function *via* Sonogashira-coupling with propargylamine. After conversion into the 5'-*O*-triphosphate this amino modified nucleotide was labelled with biotin at Fermentas as a dye-prototype (month 12).
The second dNTP bearing the 3'-*O*-CE group, which was synthesised at JWG University, was the 3'-*O*-CE-adenosine. After conversion into its corresponding 5'-*O*-triphosphate it had been sent to Fermentas as well. So the second reversible terminator using the CE-protecting group was delivered in month 12.
We applied for a provisional US patent 60/824453 covering our chemistry, its application for sequencing-by-synthesis, and a kit using this chemistry, enzymes and composition.

task1.6:
In order to attach a fluorescent dye to the base we also used the above described propargylamine-modified nucleoside triphosphate. What is also needed is a suitable cleavable linker so that the dye

can be cleaved from the base moiety. The first reversible terminator with a dye attached was Fluoresceine labelled 3'-CE-dTTP, delivered at month 21.

task1.7 and task 1.8:
During the first project period it was found that the cyanoethyl(CE)-moiety used as 3'-*O*-blocking group is accepted and incorporated best by several polymerases from Fermentas. The chemists investigated many cleavage conditions for the removal of this 3'-*O*-capping, as this could be achieved with a 1 molar **TetraButylAmmmoniumFluoride** (TBAF) solution in THF at 60°C.
Based on the fluoride-induced cleavage, the chemists developed and synthesised a dye-linker system which is also cleavable by the use of TBAF. This was achieved by the development of a glycol-based, cyanoethyl-moiety containing linker attached to the base motif of the nucleotide.
This linker was also tested regarding its stability and it could be revealed that it is cleaved even much faster under the same conditions like the 3'-*O*-cyanoethyl group. With these results in hands the chemists synthesised the three dNTPs with the fluoride-cleavable linker and 3'-*O*-capping as well as with three different dyes (A = TexasRed, C = Cy3, T = Fluoresceine). The terminators could be delivered to Fermentas in months 27/28. Within the timeframe of the project it was not possible to synthesise the fourth complete reversible terminator, although precursors for it could be synthesised. The synthesis of this terminator will be finished after the end of the project in order to publish the chemistry and its enzymatic application.

task 1.9:
As the synthesis of the terminators without the dye-linker system was known since the first project period, it could be optimized and the three propargylamine-bearing 3'-*O*-CE dNTPs (N = A,C,T) were prepared and purified in large amount. The final steps in synthesis of the complete reversible terminators are the dye-linker attachment and the triphosphate synthesis. Both steps could also be optimized (regarding A,C and T) in the last stage of the project, which enabled the delivery of the three complete terminators in sufficient amount to Fermentas' lab.

Deliverables done

D3 (**2 person-months allocated, 2 person-months spent**). Month 1. First 3'-end modified reversible terminator delivered. **Done month 1.**

D5 (**4 person-months allocated, 4 person-months spent**). Month 2. First 3'-end modified reversible terminator with the attached biotin group delivered. **Done month 2.**

D8 (**14 person-months allocated, 22 person-months spent**). Month 9. Second and third reversible terminators (A, C, T) delivered. **Done month 9.**

D12 (**4 person-months allocated, 20 person-months spent**). Month 12. First reversible terminators (A, C or T) with a dye delivered. **Done month 21.**

D21 **(12 person-months allocated, 52.5 person-months spent)** Month 24.The best reversible terminator-dNTP (A, C, G, T) with one different dye attached to each in "large" amount delivered. **Done month 28-29.**

Milestones
During the first project period, the cyanoethyl-(CE) function as a 3'-*O*-blocking group was found to be best incorporated. Therefore both 3'-*O*-CE-adenosinyl-5'-*O*-triphosphate and 3'-*O*-CE-thymidinyl-5'-*O*-triphosphate were selected as reversible terminators that are accepted by six identified DNA polymerases.
As a result of the first two terminators, furthermore three complete reversible terminators bearing the 3'-*O*-blocking group and the cleavable linker-dye system each could be synthesised,
Both 3'-*O*-CE blocking group of the terminators as well as the cyanoethyl-moiety containing cleavable linker were successfully cleaved using 1 M TBAF in dry THF at 60°C.

Problems

A big time delay was caused by the investigation in the preparation of initially chosen thioacetal based 3'-*O*-blocking groups for the reversible terminators. On the one hand, the cleavage of MTM was not quantitative, on the other hand the DTM- blocking group was instable under the polymerase reaction conditions. Besides the chemical obstacles, Quiatech got into refinancing plans during springtime 2006 and changed its' name into "Oligovation". The company, which is the owner of the thioacetal chemistry containing patents sold these to a third party then. The consequence was that the chemists had to try another synthetic approach for establishing suitable reversible terminators which contain the 3'-*O*- cyanoethyl- group. At the end of the project, the fourth reversible terminator (based on 7-deaza-7-iodo-2'-deoxyguanosine) could not be synthesised completely and therefore could not be tested in polymerase acceptance tests. With the limited timeframe of the project only precursors of this compound exist. The synthesis will be completed after the project end.

IDENTIFICATION AND IMPROVEMENT OF DNA POLYMERASE

Work package number	**WP 2**		**Start date or starting event:**		0		
Participant Short Name	Fermentas	JWG U					
Person-months per participant	110	6					

Objectives on tasks

Identification, characterization, cloning and purification of a DNA polymerase that incorporates reversible terminators with minimal discrimination between dye-labelled A, G, T and C.
task2.1: Assay development (D4)
task2.2: Identification of a DNA polymerase that incorporates reversible terminators and its purification (D9)
task2.3: purification and cloning of best polymerase (D14)
task2.4: Technique development for improvement (D19)
task2.5: Improving the DNA polymerase by gene modification (D19)
task2.6: Purification of the improved polymerase (D19)

Achievements on tasks

task2.1:

I. Development of technique for detection of reversible terminator incorporation by purified polymerases:

We started development of the technique with the assessment of quality of modified nucleotides provided by our partners. Unexpectedly, modified nucleotides were found to be substantially contaminated by natural counterparts which are incorporated much faster compared to modified ones, and thus may lead to misinterpretation of results. The solution of this issue requested additional human resources. Nevertheless, we solved the problem of contamination by enzymatic depletion of natural nucleotides from HPLC-purified stocks of modified nucleotides, and made the step of enzymatic exhausting obligatory for use of every 3'-*O*-modified dNTPs. Next, several synthetic DNA oligonucleotide duplexes labelled either with radioactive or fluorescent label at their 5'-terminus were designed and used to monitor the incorporation of nucleotides under investigation. Given that project started with modified nucleotides derived from dTTP, the template DNA strand was designed to contain several consecutive A nucleotides immediately downstream the 3'-end of primer. Such structure of DNA substrate allows distinguishing between single and multiple incorporation of nucleotides. Then, we identified experimental conditions which are suitable for detection of incorporation of known terminators like dideoxy-, 3'-*O*-acetate- and 3'-*O*-methyl-dTTP. Finally, we elaborated conditions of subsequent step, in which natural dTTP is added to the reaction mixture. This step is very important since it provides opportunity to recognize if incorporated nucleotide terminates polymerization. The developed technique was used to screen modified nucleotides, provided by our partners, for incorporation into DNA primer using preparations of different polymerases.

II. Development of technique for detection of reversible terminator incorporation by crude or partially purified cell extracts:

The use of cell extracts for assessing polymerase activity using modified nucleotides is highly challenging. To avoid exonucleolytic degradation of DNA template, inevitably resulting due to the complexity of reaction mix, locked nucleic acid (LNA)- and phosphorothioate (PTO)-based substrates (six in total) with modified residues located at or near the 5'- and/or 3'-terminus of DNA substrate were used.

Crude cell extracts from *Escherichia coli* and *Bacillus subtilis* were used to screen for incorporation of modified nucleotides following assay and conditions described above. We found that modified substrates are substantially resistant to nucleolytic degradation as expected. However, we observed that the presence of minor amounts of natural nucleotides within crude cell extracts results in efficient primer extension. This precluded possibility to detect incorporation of modified nucleotides due to

significantly lower affinity of polymerases for latters. Dialysis of crude cell extracts was found to be insufficient to remove natural dNTPs to the acceptable level, while single step of chromatography removed dNTPs efficiently thus allowing detection of polymerase activity and incorporation of dideoxyTTP which was used in that experiment. Such an approach has a potential to be used for screening of polymerases accepting 3'-*O*-modified nucleotide by using crude cell extracts as a source. However, chromatographic preparation of individual samples is poorly compatible with required high-throughput mode of screening. Considering both the very low throughput of screening among the wild type microorganisms and hopeful results of incorporation of modified nucleotides by commercially available polymerases further screening efforts using crude extracts have been discontinued by decision made on meeting of Management Committee in Vilnius.

task2.2:

Two main sources of polymerases were considered: (I) those available from commercial sources, and (II) Fermentas UAB proprietary thermostable DNA polymerases.

To test the ability of polymerases to incorporate modified dNTPs into DNA, an exhaustive set of enzymes (more than 30 members) representing all families of DNA polymerases was collected. Model compounds, including dideoxy-, 3'-*O*-acetate- and 3'-*O*-methyl-dTTP were used for fine-tuning of screening assay. We found that polymerases exhibiting prominent nucleolytic activity are poorly compatible with assay, unless specific measures are taken (see task2.1, part II for details).

For proprietary Fermentas polymerases, five mutants with reduced nucleolytic activity were prepared and tested. Even then, some polymerases had to be dropped from list narrowing it to 23 representatives with lowered or absent nucleolytic activity, both of mesophylic and thermophylic origin. Full range screening program was executed using four modified dTTPs featuring 3'-*O*-Dithio*Tert*butylMethyl (DTM), 3'-*O*-MethylThioMethyl (MTM), 3'-*O*-CyanoEthoxyMethyl (CEM) and 3'-*O*-CyanoEthoxy (CE) protective groups.

DTM protective group has been found to be unstable during enzymatic incorporation into DNA; thus, under the agreement reached by project members it has been dropped from the list of possible terminators. Screening performed using 3'-*O*-MTM-dTTP was highly successful; however, efforts taken for uncovering method to remove protecting group (keeping DNA structure intact) failed, making this compound useless for the project. It should be noted, however, that this compound has a potential to be used in those applications where only one step of modified dNTP incorporation is required (i.e. investigation of single nucleotide polymorphism).

CEM and CE protecting groups proved to be stable both in free (nucleotide) and incorporated (into DNA) form and are being incorporated into DNA by number of polymerases (see below). Work on deprotection conditions for these terminators is in progress to meet demands of cyclic nature of sequencing-by-synthesis protocol.

In summary, two polymerases able to incorporate 3'-*O*-CEM-dTTP and six polymerases incorporating 3'-*O*-CE-dTTP were discovered. Of importance, incorporation of modified nucleotides mentioned above ensures efficient termination of DNA synthesis.

task2.3:

The cloning of gene for DNA polymerase was planned in conjunction with identification of novel polymerase in cellular extracts. After initial experiments, efforts in this direction have been discontinued (see task2.1, part II). Moreover, it was found that polymerases capable for incorporation of 3'-*O*-modified nucleotides are available in purified form from commercial vendors, including ourselves. All of discovered enzymes are recombinant thus making need for cloning elusive. Most important, gene of the polymerase of choice for improvement (see task2.4) is available in-house at Fermentas for further experiments. Thus, no cloning experiments were carried out, and human resources devoted for that task were switched to experiments aimed to develop the evolution *in vitro* scheme (see below).

task2.4,task2.5 and task2.6:

Several different approaches were employed to address this task which is related to deliverable D19. First, number of rationally designed mutants targeting surroundings of the polymerase active site was

generated. Several available crystal structures were used to build a model for ternary substrate DNA-polymerase-nucleotide complex and key residues identified, critical for nucleotide entering into the active site. Then, the key residues were substituted by residues ensuring better accessibility for nucleotides towards active site; also, mutations allowing enhanced flexibility of non-structured loops within this trajectory were introduced. Potential candidates were partially purified and tested for affinity for 3'-*O*- CE modified nucleotides and primer extension velocity, using these nucleotides that have been provided by the chemists from JWG University.

In parallel, an alternative, colony-based screening assay suitable for detection of mutants with higher incorporation rate of modified nucleotides was developed. The screening is based on incorporation of 3'-*O*-CE modified nucleotide (which has a fluorophore attached to the base) into the immobilized dsDNA substrate. This step is followed by immune enzymatic detection of the fluorophore. The approach proved itself as both highly selective and high-throughput, allowing screening through up to 30-50 thousand colonies on a single plate.

Finally, we have designed the original and versatile *in vitro* evolution system addressing the incorporation of modified nucleotides by any polymerase of choice and expandable for improvement of other enzymatic properties as well.

The system combines two *in vitro* evolution approaches which were never combined before - ribosome display and water-in-oil compartmentalization. In order to make it suitable for isolation of polymerase mutants which incorporate modified nucleotides efficiently, prior to *in vitro* translation the mRNA molecules must be modified in such a way to provide the double-stranded DNA substrate. The latter serves for incorporation of 3'-*O*- CE modified nucleotide (which contains the biotin group attached to the base) during selection.

Molecules of polymerase able to incorporate the biotinylated nucleotide are separated using streptavidin beads and their co-isolated mRNA molecules are further amplified by RT-PCR. There are 4 major milestones in this selection approach:

(1) *in vitro* translation of active polymerase;
(2) attachment of DNA substrate to mRNA;
(3) general optimization of experimental setup;
(4) isolation of improved polymerase.

Both milestones are fulfilled. Final optimization of the whole experimental setup of polymerase selection and identification of improved polymerase is expected to be completed in the nearest future.

Altogether, a number of mutants have been obtained by rational design and by colony-based screening. Ten mutants were purified and tested. The best candidates featured up to 20% increase in efficiency of incorporation of 3'-*O*-CE modified nucleotides bearing fluorescent label, basing on the results obtained from experiments executed in microchip mode. Our results allow the suggestion that several rounds of random mutagenesis followed by screening of most promising mutants are required to increase the efficiency of incorporation substantially.

Deliverables done

D4 (**4 person-months allocated, 4 person-months spent**). Month 2. A protocol describing the technique that allows detection of reversible terminator incorporation by crude or partially purified cell extracts and by purified DNA polymerases. **Done month 2.**

D9 (**25 person-months allocated, 25 person-months spent**). Month 10. Identification of a DNA polymerase that incorporates reversible terminators and its purification. **Done month 10.**

D14 (**13 person-months allocated, 13 person-months shifted to D19**). Month 14. Purification of the best identified DNA polymerase and cloning of the respective gene. **Required genes are available.**

D19 (**74 person-months plus 13 person-months from D14 allocated, 87 person-months spent**). Month 24. Isolation of the best improved version of DNA polymerase and its purification. **Done month 29.**

Milestones

At month 2, a protocol describing the most suitable technique for evolution of polymerases had been delivered. At month 10, several DNA Polymerases that incorporate reversible terminators had been identified. At month 29, the best improved version of DNA polymerase was evolved and purified (see D19). This polymerase accepts and incorporates the fluoride-cleavable, reversible terminators provided by the contractors of workpackage 1.

Problems

1. Use of crude cellular extracts has been found to be incompatible with required high-throughput mode of screening. At the same time commercially available polymerases which incorporate modified nucleotides were identified. Therefore Management Committee decided to stop search of suitable polymerases in crude extracts.
2. DTM terminating group is highly unstable, whereas MTM was found to be impossible to deprotect keeping DNA intact. Therefore, CEM and CE were selected for further work on polymerase improvement.

DEVELOPMENT OF A MICROFLUIDIC DEVICE

Work package number	**WP 3**	**Start date or starting event:**		0			
Participant Short Name	Silex	Asper	Quiatech				
Person-months per participant	8	2	1				

Objectives and tasks

The goal of the microfluidics development will be to deliver a reaction device that allows array SBS reactions by integrating incubation steps at different temperatures with controlled agitation of reaction components over the entire area of a microarray, including washing and regeneration procedures.

task3.1: Draw first prototype (D7)

task3.2: Make first prototype (D7)

task3.3: Draw second prototype (D11)

task3.4: Make second prototype (D11)

task3.5: Draw final prototype (D18)

task3.6: Make final prototype (D18)

Milestone: Delivery of the third prototype (D18)

Achievements on tasks

task3.1 and task3.2

Silex has together with Asper designed the first microfluidic device prototype. Silex manufactured and delivered this prototype to Asper. In parallel, Silex delivered diced slides with different coatings for evaluation of which coating should be best suited for future prototypes (bio-compatibility).

Justification of ressources:

- Made design of the first microfluidic device.
- Made process design for manufacturing the first microfluidic device.
- Made processing for 6 wafers coated with 6 different coatings and diced them into single devices of the same dimentions as the real microfluidic devices, for biocompatibility testing.
- Made CAD design and ordered masks for the first microfluidic device fabrication.
- Manufactured and delivered the first microfluidic device. The device consists of a microfluidic platform (reaction chamber) etched on the frontside, with backside holes microfluidic interconnections. In addition a recess in the silicon is made around the reaction chamber. This recess is used to place an O-ring that maintain permeability while placing the lock on the reaction chamber.
- Purchase of material for device production (silicon wafers and mask set).
- Purchase of O-rings for testing.

task3.3 and task3.4

Because of the changes in chemistry regarding the nature of the 3'-*O*-blocking group of the reversible terminator as well as its cleavage conditions, the suitable coating (gold) for the surface of the reaction chamber had been chosen in month 12. The design and manufacture of the second prototype for the microfluidic device is therefore delayed (see D11).

task3.5 and 3.6:

The first two versions of microfluidic device prototypes were delivered during the first year of the project. The data from the work regarding the chemistry (WP 1), the enzymology (WP 2) and the sequencing technique (WP 5) during the second year strongly influenced the design and production of the last prototype (see fig. 1 and fig. 2).

Both tasks are fulfilled by performing the following working steps:

- Made design of the different microfluidic devices.
- Made process design for manufacturing the microfluidic devices.

- Made processing for 6 wafers coated with 6 different coatings and diced them into single devices of the same dimentions as the real microfluidic devices, for biocompatibility testing.
- Made CAD design and ordered masks for the different microfluidic devices fabrication.
- Manufactured and delivered the different microfluidic device. The device consists of a microfluidic platform (reaction chamber) etched on the frontside, with backside holes microfluidic interconnections. In addition a recess in the silicon is made around the reaction chamber. This recess is used to place an O-ring that maintains permeability while placing the lock on the reaction chamber.
- Purchase of material for device production (silicon wafers and mask set).
- Purchase of O-rings for testing.

Deliverables done

D7 **(3 person-months allocated, 3 person-month spent)**. Month 3. A first microfluidic device prototype adapted to the existing Asper detection platform. **Done Month 3.**

D11 **(5 person-months allocated, 4.4 person-month spent)**. Month 12. A second microfluidic device prototype, with for example, channels, heaters, mixer and valves adapted to the existing Asper detection platform. **Done Month 15.**

D18 **(3 person-months allocated, 4 person-month spent)**. Month 23. A third microfluidic device prototype with integrated features adapted to the existing Asper detection platform and optionally other detector platforms. **Done Month 23.**

Milestones

A final (third) microfluidic device prototype design was designed and manufactured at Silex. The third microfluidic device was delivered month 23. It was decided (meeting in Copenhagen, 11th January 2008) that it was not necessary to go on and include extra integrated features such as heaters to the device since the point to test it would not be reached within the timeframe of the project.

Problems

The start of chip manufacturing task was a bit delayed, approximately of two months. Mainly due to late design input within project. Therefore deliverable D11 was postponed to the second period (to month 14-15).

BIOINFORMATICS AND SELECTION OF PRIMERS

Work package number	**WP 4**		**Start date or starting event:**			0	
Participant Short Name	Tartu	Asper					
Person-months per participant	36	6					

Objectives and tasks

Develop bioinformatics tools for identification of primers for chip sequencing and software for analysing sequencing-by-synthesis results.

task4.1: Method development step1 (D10)

task4.2: Implement primer design software (D10)

task4.3: Improvement of the parameter settings of the method and the 1st software prototype (D15)
task4.4: Modify Base Caller for ArraySBS application. (D22)
task4.5: Identify primers for p53 re-sequencing (D22)
task4.6: Analyse data from instrument (D22)

Milestone: A software for primer design and software for handling image data from SBS experiments (D22)

Achievements on tasks

task4.1 and task4.2:

As the ArraySBS technology is still in development, we started statistical modelling of primer properties on similar microarray platform. Initial analysis is based on data from APEX microarrays. APEX is a four-channel microarray genotyping platform that is using primer extension (http://www.asperbio.com/process.htm). We used data from 243 microarrays with 889 genotyping primers each.
In the first stage of the project different normalisation methods were investigated. Signals from each microarray and each channel were normalised by different methods to achieve comparable signal values in each channel of each microarray. The effect of normalisation was tested by measuring array-specific variation component in MANOVA. The best normalisation was achieved by the following formula: norm_signal = log2 (raw_signal / 90%_quantile of all signals in given channel).
With this normalisation method we were able to reduce microarray specific variation from initial 23% to nearly 0%.

task4.3 and task4.4:

The idea of base calling is to convert microarray signal intensities into 10 possible genotypes: AA, AC, AG, AT, CC, CG, CT, GG, GT, and TT. Base calling procedure includes several steps such as recording signal intensities of experimental data, normalization of signal intensities between channels and arrays, training of genotype models based on training data and predicting genotypes given new experimental data. The base calling method should be highly accurate and reliable. In this project a particular need was to develop procedure for re-sequencing base calling. The main difference between SNP genotyping and re-sequencing is in the training sets. Whereas in SNP genotyping one can expect 3 genotypes and it can also be expected that all of them are present in training data (Figure 1A). In the re-sequencing process, the assumption is made that any of 10 genotypes is possible even if only one genotype is seen in the training dataset. Therefore the re-sequencing base calling algorithm needs slightly different approach than SNP genotype base calling.
The co-workers at Tartu University decided to work their common way to required base calling software gradually, in several steps:
1. Find suitable experimental data for testing and a method for comparing algorithms
2. Find the best SNP genotyping algorithm
3. Try to improve the best SNP genotyping algorithm

4. Adapt the best solution for re-sequencing.

1. Finding the dataset and method for comparing base calling algorithms

To decide which base calling method is the best, appropriate characteristics are needed. The researchers used the ROC (receiver operating characteristic) curve to compare performances of different algorithms. The ROC curve describes performance as a trade off between selectivity and sensitivity.

The analyses performed were based on the results of APEX ARGACE and ABCR genotyping microarray experiments conducted previously by Asper Biotech Ltd. In order to get reliable results the data were pre-filtered to remove low-quality microarrays, and primers from our dataset. The final dataset contained 18883 base calls from 175 different primers.

2. Choosing the best algorithm for base calling

At first, several principally different clustering algorithms on SNP genotyping were tested to see which algorithm should be chosen for further development of re-sequencing software. Comparative analyses were conducted using two algorithms used by Asper Ltd. current genotyping software; using PAM-method (a variant of K-means clustering), LDA (linear discriminant analysis) and genotyping software Sniper which is based on Gaussian clustering and uses EM (expectation maximization). The most promising results were achieved by the **LDA algorithm** and **EM-algorithm.** The researchers assumed that EM algorithm can be more flexible for re-sequencing and therefore it was implemented in new base calling software BsPred (Base Predictor).

3. Improving the algorithm - should the complexity of genotype model be increased?

To improve their EM-algorithm performance, they tested different possibilities of cluster variance modelling. They carried out training-testing analyses with ten different cluster variance models as described for comparing both algorithms. The best results were achieved with the **EII** (currently used in BsPred) **and EEE models**. More complex models are probably too flexible and can produce over-fitting effects too easily.

4. Adapting software for re-sequencing

After developing initial software prototype it was continued with adapting the prototype for re-sequencing. The EM-algorithm needs a training set for estimating cluster parameters such as mean intensities and variances in different channels. All different genotypes must be represented in the training set. It is not possible to call genotypes that are missing in the training set. However, in re-sequencing data it is needed to call also rare genotype variants that are likely to be missing in training data. One approach for this problem is the prediction of the location of genotype clusters that are missing in training set. There have been tried four different genotype modelling methods for this purpose.

a) BsPred_a. This method predicts missing clusters by transposing the training cluster into appropriate channel (orthogonal model). This relationship can be expressed by general formula $X_{AA} = Y_{CC}$; $Y_{AA} = X_{CC}$; etc. for all possible genotype pairs. X and Y correspond to normalized signal intensities of A channel and C channel respectively.

b) BsPred_b. This method predicts missing clusters by using linear correlation coefficients which represent the pair wise relationship between all different genotype clusters. This relationship uses only intercepts and can be expressed by general formula $X_{AA} = X_{CC} + b_1$; $Y_{AA} = Y_{CC} + b_2$; etc. for all possible genotype pairs) where $b_1..b_n$. are parameters estimated from other primers in training set.

c) BsPred_c. This method predicts missing clusters by using linear correlation coefficients which represent the pair wise relationship between all different genotype clusters. This relationship uses slopes and intercepts and can be expressed by general formula $X_{AA} = a_1 * X_{CC} + b_1$; $Y_{AA} = a_2 * Y_{CC} + b_2$; etc. for all possible genotype pairs) where $a_1..a_n$ and $b_1..b_n$. are parameters estimated from other primers in training set.

d) BsPred_common. This method uses one common model for all primers. Thus genotype clusters need not to be predicted (other primers would cover missing genotypes in training set). However, because this is not primer-specific model, its variance and thus error rate in prediction could be higher.

Method a) is completely primer-specific (different models for each primer), methods b) and c) are partly primer-specific and method d) is not primer specific at all.
It was found that method (c) works best out of 4 tested methods, although the differences are not large. Linear discriminant analysis gave poorer results compared to EM-algorithm used in BsPred software.

5. Developing software and graphical user interface for base calling
The best base-calling algorithm was implemented into software called BsPred. The web interface for BsPred was implemented with BsPred_common model and is available at http://bioinfo.ut.ee/bspred/.

As a conclusion on achievements on both tasks, several algorithms have been tested and it was found that the most appropriate algorithm for genotyping is EM algorithm, which uses equal shape and volume of genotype clusters (type EEE and EII). Named algorithms were implemented into BsPred software and adapted for base calling of re-sequencing data.

task4.5 and task4.6:
The principles of primer design software were described in previous periodic activity report (D10). Within the current project activity period, the formula for prediction of call rates and signal intensities of primers was finalized. The final formula has six factors and can predict 28% of variation in signal intensities or 5% of variation in call rates.
As a final part of this study, re-sequencing primer design software called SBS Designer was created for prediction of call rate and signal strength of candidate primers. SBS Designer can either evaluate the quality of existing primers or design new probes for a DNA sequence. The sequence can be provided in three different ways: entered manually into web text-fields; loaded from a text file (in FASTA format); or searched for by using chromosome coordinates or gene ID. As a result, the program will produce a list of primers that can be used to re-sequence given DNA region, accompanied by statistics, such as predicted call rate that is calculated using the prediction model, melting temperature (Tm), length, genomic positions etc. The primer design algorithm works by first designing primer for all nucleotide over the input DNA region. The probes can either be of fixed length (minimum is 15) or have uniform melting temperatures, in which case such primer length will be automatically chosen that is closest to the desired Tm value. Primers can be designed on both strands of DNA or only for either one.
As a conclusion of tasks 4.5 and 4.6, the co-workers at Tartu University developed successfully statistical models for prediction of call rate and signal intensities of extension primers on microarrays. Corresponding software for re-sequencing primer design was therefore developed and is now available.

Deliverables done

D10 (**14 person-months allocated, 21 person-months spent**). Month 11. Methodology and software for primer design of SBS microarrays. **Done Month 11.**

D15 **(12 person-months allocated, 19,5 person-months spent).** Month 14. Software prototype for handling image data from SBS experiments. **Done Month 14.**

D22 (**16 person-months allocated, 42 spent**). **Month 24.** A software for primer design and software for handling image data from SBS experiments. **Done Month 24.**

Milestones

The expected result at the end of the second project year is a well functioning software tool for array based sequencing-by-synthesis and primer design. This milestone was successfully completed at month 24 due to deliverable D22.

DEVELOPMENT OF A RE-SEQUENCING ASSAY

Workpackage number	**WP 5**	**Start date or starting event:**			0		
Participant Short Name	Asper	Tartu	JWG U	Quiatech	Fermentas	Silex	
Person-months per participant	30	4	4	1	2	1	

Objectives and tasks

To develop an assay for the p53 gene coding sequence (1218 bp) in a microfluidic reaction device and to validate it by comparing the results with the ones produced by APEX and Sanger sequencing.

task5.1: Development of SBS method on APEX chips (D16)

task5.2: Implement first SBS setting (D16)

task5.3: Implement second SBS settings (D20)

task5.4: Implementation of p53 re-sequencing (D20)

Achievements on tasks

task5.1:

Asper's current APEX-based TP53 re-sequencing system glass microarrays with fluorescently-labelled dideoxy terminators and ThermoSequenase DNA polymerase was used to evaluate the origin of possible variations in re-sequencing results. The DNA fragments used for APEX were amplified from one individual DNA three times in order to create sufficient amount of material for three independent analyses. The amplifications, purifications and pre-APEX treatments were carried out on different days. The APEX reactions were performed using Asper's standard conditions and the analyses of results are currently in progress at Tartu University Bioinformatics department in order to help the selection of best possible re-sequencing primer candidates.

Asper also ran tests in the lab to understand the factors influencing the design of the ArraySBS system microchips by applying automated hybridization station HS-400 from Tecan AB.

In this period Asper Biotech has been mostly doing research for the deliverables to be completed in second period and working together with other partners, mainly Silex AB from Sweden and University of Tartu.

The first prototype of microfluidic device was worked out in collaboration with Silex AB and was provided by them. From different coating materials of microfluidic chamber offered by Silex AB, the Teflon was first selected based on its chemical resistance.

task5.2:

In the first project year, Asper worked mostly with the APEX slide system. Due to the development regarding the chemistry with its glass-etching fluoride-based cleaving conditions, a suitable DNA chip material which is resistant to this deprotection chemistry and which allows a stable immobilization of the oligonucleotides was needed. For this purpose, different DNA chip carrier materials were tested at Asper´s lab:

1. SAL-0.1 DNA microarray glass slides produced by Asper Biotech. The slides were spotted with 3'-Cy-3-labeled and 5-amino- modified oligonucleotides using standard protocols and scanned with Affymetrix 428 microarray scanner. The slides were treated with 1 M Tetra-*n*-butylammonium fluoride in tetrahydrofuran (TBAF in THF) for 5 min at room temperature, washed once with THF and scanned. The results showed that after 5 min treatment in 1 M TBAF in THF about 55-60% of the applied oligonucleotides remained immobilized.
2. Silicon slides protected with gold layer provided by Silex were also tested. 5'-thiol- modified oligonucleotides labeled with 3'-Cy3 were immobilized onto slides. The slides were treated with 1M TBAF for 20 min at room temperature, washed once with THF and scanned. The results showed that after 20 min treatment in 1 M TBAF in THF only about 10% of the applied oligonucleotides remained immobilized and the protective effect of gold layer was not sufficient.
3. CodeLink™ Activated Slides from GE Healthcare (formerly Amersham BioSciences) were also

tested. The slides were spotted with 3'-Cy-3-labelled and 5'-amino-modified oligonucleotides using protocol suggested by vendor and scanned with Affymetrix 428 microarray scanner. The slides were treated with 1 M TBAF for 15 min at room temperature, washed once with THF and scanned. The results showed that after 15 min treatment in 1 M TBAF in THF no detectable amount of immobilized oligonucleotide molecules were cleaved off from the glass.

In the following experiments Cy-5-labelled oligonucleotide probes, fully complementary to the immobilized probe sequence were used for hybridization experiments. The oligonucleotide duplexes were treated with 1 M TBAF as shown above and scanned again to measure how much of the DNA target will remain hybridized. We found that about 28% of labeled oligonucleotide template stayed in duplex with complementary oligonucleotide probe after TBAF treatment.

Based on these experiments the CodeLink™ Activated Slides were chosen for further experiments.

To perform primer extension experiments due to the low stability of the duplex during the TBAF/THF treatment, special hairpin oligonucleotide probes were designed by Tartu University partner for testing of different single strand part length and incorporation efficiency and to avoid DNA template stability issues. Oligonucleotide probes with both 12 and 17 nucleotide linker were designed, ordered and spotted onto microarray slides. In addition, four branched oligonucleotide probes were designed and synthesized by Oligovation using their proprietary synthesis technology, in order to find out whether the orientation (towards glass surface of upwards) of free 3'-OH group will play any role for primer extension reactions. These branched oligonucleotide probes were immobilized onto slide surface from the middle of oligonucleotide probe using the same chemistry as normal oligonucleotide probes.

Our results showed that hairpin oligonucleotide probes carrying 17 nucleotide linker worked better for primer extension than probes with 12 nucleotide linker. Therefore, for the further experiments we decided to use probes with 17 nucleotide linker. On the other hand, there was no difference in primer extension efficiency between normal and branched oligonucleotide probes. Since the synthesis procedure of branched oligonucleotides is much more complicated than ordinary DNA probe synthesis, it was decided to continue with normal DNA probes.

task5.3 and task5.4:
In the second stage, 64 oligonucleotide probes for testing up to three consecutive cycles of terminator incorporation/removal were designed by Tartu University partner. There is a 3 nucleotide motif after each hairpin. All oligonucleotides have been tested for proper secondary structure. These oligonucleotide probes were used to spot 100 primer microarrays to perform proof- of principle experiments of array SBS procedure.

Deliverables done

D16 **(16 person-months allocated, 67.5 person-months spent)**. Month 14. Proof-of-principle of array SBS procedure with up to 100 primers. **New delivery Month 30, in progress.**

D20 **(26 person-months allocated, 20 spent)**. Month 24. The whole p53 coding sequence microarray tested with array SBS procedure using 3 labelled dNTPs. **In progress.**

Problems

The research for microarray carrier material was much more complicated than expected. The results of the work done at Quiatech and Fermentas significantly influenced the concept of the ArraySBS system and changed it significantly from the first plans. This has posted significant challenges to Asper to find surfaces and technologies working with the enzymes and reversible terminator technology developed by the other partners. From silicon chips coated by Silex AB with the following protective materials: thermal oxide, nitride, C_4F_8, aluminium and gold, the gold-coated chips were selected for further tests in regard to 5'-thiol group modified oligonucleotide primer binding properties. All these tests led to a big time delay as well as the late delivery of the three reversible terminators for performing the SBS re- sequencing.

PROJECT MANAGEMENT

Workpackage number	**WP 6**	**Start date or starting event:**					0
Participant Short Name	JWGU						
Person-months per participant	4						

Objectives and tasks

To provide efficient coordination of all scientific and administrative work, including regular assessment of project work.

task6.1: daily general project management issues

task6.2: Establish a web page (D1, D2)

task6.3: Design and publication of a project leaflet (D6)

task6.4: Preparing an intermediate plan for using and disseminating knowledge (D13)

task6.5: Writing final public scientific report and a plan for disseminating knowledge (D23 and D24)

Achievements on tasks

task6.1:

The main daily project management issues are the organisation of the communication and other interactions between the partners and the collection of all information that is exchanged between the different work packages. Further the coordinator has to be addressable for all participants in cases of problems or questions of all kinds.

Further three project meetings were organized and the coordinator was responsible for making the presentations held during this meetings and the minutes available on the web page.

The coordinator also was responsible for the organisation of the dissemination of the knowledge in the patent filed.

task6.2:

A functioning web page as well as a functioning internal web page had been established, see deliverables list below.

task6.3:

This task and its deliverable has been dropped. The explanation for this can be found in section "Problems" below.

task6.4:

An intermediate plan for using and disseminating knowledge had been delivered by the project coordinator in month 12 and can be found in the Annex of this periodic activity report.

task6.5:

As it was already described in the first PAR, an intermediate plan for using and disseminating knowledge had been delivered by the project coordinator in month 12 and can be found in the Annex. The final public scientific report could not be written within the timeframe of the project due to missing results from WP1, WP2 and WP5. A scientific publication covering the chemistry, the polymerase acceptance and incorporation will follow after the end of the project.

Deliverables done

D1. (**0.1 person-months allocated, 0.25 person-months spent**) Month 1. A functioning web page established, including a public Project Presentation. **Done Month 1.**

D2. (**0.1 person-months allocated, 0.25 person-months spent**) Month 1. A functioning internal web page established. **Done Month 1.**

D6. (**0.15 person-months allocated, 0 person-months spent**) Month 3. A project leaflet. **Cancelled.**

D13. (**0.25 person-months allocated, 0.75 person-months spent**) Month 12. An intermediate plan for using and disseminating knowledge. **Done Month 12.**

D17. (**1 person-months allocated, 1 person-months spent**). Month 14. First non-confidential project summary report. **Done Month 14.**

D23. (**1 person-months allocated, 1 person-months spent**). Month 25. Final non-confidential project summary report. **Done Month 31.**

D24 (**1 person-months allocated, 1 person-months spent**). Month 25. Plan for using and disseminating knowledge. **Done month 31.**

Problems

The objective for deliverable was to spread the message about the project to a public audience in order to find potential commercialization partners. The project was announced with the press release 24th August 2005 at the Quiatech website and by Tartu University. It was also sent to several genomic journals and it was published at the Genome Website.

Due to this well spread of the activities two third parties contacted Quiatech and began to discuss business opportunities. With ongoing business discussion it would have been wrong to publish a leaflet. Thus no project leaflet was written.

The contact person at Quiatech has left the company. The new coordinator, Prof. Engels (JWG U) is elected.

3 Consortium management

4.1 Organization of scientific collaboration

As many of the persons as possible working for the different parties were involved by inviting them to the kick-off meeting held in Frankfurt as well as to the other project meetings held in Vilnius and Uppsala. Each project meeting was divided into two parts, during the first part one person representing each WP presented the planned work and achieved results with all persons present. All presentations were uploaded to the ArraySBS website for later reading. The second part was held with the Project Management Committee (PMC) members in order to discuss related issues to the presentations and decide on coming actions and activities. The decisions in the PMC were based on majority votes. In case of a conflict because of equal numbers on each side the Project Coordinator (PC) will cast the decisive vote.

4.1.1 Kick-off meeting in Frankfurt/Germany (31.08.2005 - 02.09.2005)

13 persons participated in the kick-off meeting.

The following topics were covered and discussed:

- The PMC group members and tasks
- Contracts
- Reports
- EU contacts Scientific and Financial Officers
- Time lines
- Deliverables
- Budget
- How we should communicate

Decisions were taken unanimously on the following topics by the PMC:

- If a party would like to release a local Press Release in the local language they are free to do that. A copy of the PR should be sent to the coordinator. Additional information can be added after communication with the coordinator.
- Publications should be discussed as early as possible in order to not delay any publication. Submission should be approved by the SME group.
- IP: If needed, a patent lawyer may be invited to explain inventions etc.
- All communication should be with e-mails and phone. If needed for specific actions visits should be organized.

4.1.2 Second ArraySBS project meeting in Vilnius/Lithuania (01.03.2006 - 03.03.2006)

22 persons participated in the second project meeting.

Decisions on the research activities taken unanimously by the PMC:

- **Regarding WP2:** To modify the DoW by stopping the screening of polymerases from crude extracts, since a good candidate exists. Focus on molecular evolution. Fermentas will implement this.
- **Regarding WP1:** JWG U has to send the amino-linked 3'-*O*-CE-thymidinyl-5'-*O*-triphosphate as soon as possible to Fermentas after the choice of CEM or CE in April.
- **Regarding WP1:** To send the 3'-*O*-CE-thymidine synthesised at Quiatech to Engels and to Fermentas to make the triphosphate as soon as possible.
- **Regarding WP2:** During the coming polymerase tests at Fermentas, the potential reversible terminators should be evaluated with the following priority order: 1/ CEM, 2/ CE, 3/ DTM

and 4/ MTM. The MTM group will not be further tested until it can be unblocked. This will be further tested by Marek Kwiatkowski.

- **Regarding WP1:** The next base to be synthesised should be dATP with CEM/CE at the 3'-end (both at JWG U labs). From both 3'-*O*-protection groups the amidite and triphosphate should be made (Quiatech and JWG U). A primer ending with A- CEM/CE should be sent to Fermentas during May (Quiatech).
- **Regarding WP1 and WP5:** To use preferentially Cy3. Asper has to approve other alternative dyes.

Decisions on the project management taken unanimously by the PMC:

- All parties agreed that equipment and travel money can be used for payment of staff as long as it is within the budget.

Decisions on the dissemination of knowledge taken unanimously by the PMC:

- The patent situation was discussed during the Second ArraySBS Meeting in Vilnius. It was concluded that both CE/CEM have a good potential to be the base for a new patent application. The content of a patent application was discussed, and it was concluded that it should include the blocking group, deblocking process and polymerase, optionally the detection. Owners are Fermentas, Quiatech and Asper. Aspers participation in this proposed application has to be further discussed when results exist, since it will be dependent on the claim set.
- An agreement between Asper, Fermentas and Quiatech must be set up within 3 months after filing of a patent application. The application should be filed during May.
- No publicity of the news in the project before a patent application is filed. The target time for filling it should be in May, but it depends very much on the results that are available at that time. All supported the idea to go for a publication of the results and decide of content in August.

4.1.3 Third ArraySBS project meeting in Uppsala/Sweden (23.08.2006 - 25.08.2006)

14 persons participated in the second project meeting.

Because of the very positive polymerase-incorporation test results by using the new 3'-*O*-CEM protection group, and even better incorporation results with the 3'-*O*-CE-protection group, the PMC members made the following decisions.

Decisions on the research activities taken unanimously by the PMC:

- **Regarding WP1:** The CE modification at the 3'-*O*-position should be the only used group. The linker between the base and the dye should preferably be fluoride sensitive, since then the linker can be cleaved and the 3'-*O*-protection can be converted to the hydroxyl group in one step.
- **Regarding WP1:** The dye that should be used for labelling the reversible terminator is fluorescein (the previous decision to use Cy3 had been modified due to the very high cost of this dye).
- **Regarding WP1:** A dATP-CE with a linker and fluorescein should preferably be ready within 3 months.
- **Regarding WP2:** The biotinylated dTTP-CE which was synthesised in the first year of the project should be used for the molecular evolution experiments.
- **Regarding WP2:** Due to the changes in the ownership of the intellectual property rights, the work with the 3'-*O*-DTM-modification should from now and onwards be stopped and no additional work should be performed on this compound by the parties within the framework of the ArraySBS project.
- **Regarding WP1 and WP2:** Marek Kwiatkowski should synthesise an oligonucleotide with a 3'-end CE and provide it to Fermentas.

- **Regarding WP3:** Silex should send Au-coated Si-slides to Asper for binding of thiol tagged primers and test if these are stable against 1 M TBAF.
- **Regarding WP3:** Marek Kwiatkowski should contact friends at Ångström lab to check for other coating options.
- **Regarding WP3:** Per-Johann Ulfendahl should contact Åmic AB to discuss plastic options for slides and suppliers.
- **Regarding WP3:** Per-Johann Ulfendahl should contact a former colleague in order to check if spare parts for microfluidic can be bought cheaply.

Comments on the decisions concerning the research activities:
At the moment the new reversible terminator bearing the 3'-*O*-CE-blocking group needs fluoride ions (TBAF) in tetrahydrofuran (THF) for complete removal of the blocking group after incorporation. The drawback with fluoride ions is that they etch any surface containing silicon (Si). This causes problems to the primary plans of the contractors of WP4 and WP5 to use glass surfaces. Now they need an alternative material (like Teflon) or a coating for the glass slides that is resistant against fluoride ions. They will be concerned with finding a solution for this problem during the second working period of the project (see work package description form of WP4 and WP5). One easy accessible option is a standard gold (Au) coating from Silex that may work with both fluoride ions and binding of DNA oligonucleotides to the surface. If a gold surface is used oligonucleotides with a 5'-end thiol can be used to link them to the gold surface.

More work is needed to verify that the 3'-*O*-CE protection can be removed. For that work an oligonucleotide with a CE at the 3'-end is needed. Further other possibilities than fluoride ions should be investigated using a suitable model system.

Asper needs additional support in finding ways to make a microfluidic device, since Silex can only make the slides. One option is to use an external company to assemble this device. The external advisor of the consortium Dr. Andrew Griffiths will be contacted for helping with outside contacts.

Deviation from the project work programme and its impact:
Only one activity that has an impact on the timetables of other work packages is delayed according to the original plan. It is the synthesis of the dye labelled reversible terminator it will be provided approximately 3 months later (month 15 instead of month 12). The consequence of this delay is that Asper will get shorter time for the final re-sequencing assay. This can be handled since the assay is fast and they will get additional time for verifying the microfluidic device in combination with the detection platform.
So after the first year working period it was found that the work plan and timetable are still applicable.

Decisions on the project management taken unanimously by the PMC:
- As the new Project Coordinator (start month 13) Prof. Dr. J. Engels was elected.
- The first year report should be sent to the EC Officers at or before 30 September.
- A phone conference should be held 20 November at 13:00 C.E.T.
- Next project meeting should be held in Tartu, Estonia 08.03.2007-09.03.2007.

4.1.4 Fourth ArraySBS meeting in Tartu/Estonia (08.03.2007 - 09.03.2007)

16 persons participated in the 4th ArraySBS meeting.

The following topics were covered and discussed:

- **WP 1-5** program for 2007 (extension of 6 months)
- Tasks
- Deliverables

- Milestones
- Patent
- Exit Scenario
- Next Meeting

Workpackages 1 – 5:
Decision: The new coordinator of the ArraySBS project management consortium, Prof. Dr. J. W. Engels, will make an application for extension of time for the ArraySBS project. An extension time of six months should be sufficient to fulfil all deliverables regarding the DoW. The project would finish then at the end of February 2008. The extension of time had been agreed by all project members.

Tasks, deliverables and milestones:
The deliverables list from the DoW has been updated: Regarding D12 and D21, G was excluded from the list.

Patent:
Decision: Until now, it has not been decided, if another patent regarding the reversible terminator chemistry and Fermentas' polymerase should be filed. This decision is to be discussed within the last period of the project.

Exit scenario:
As a positive outcome of this project, all four terminators with the 3'-*O*-CE-blocking group, a cleavable linker and dye will be synthesized and will have led to the evolution of a selective and efficient polymerase for the SBS procedure.
This will enable the workpartners from JWG University, Fermentas, Asper and Oligovation to publish the proof-of-principle results in *Nature Biotechnology* or similar scientific magazines. The intellectual property may also be attractive for a third company, as well as the development of a final microfluidic device would be possible for SBS-interested companies.

Next Meeting:
Decision: The 5th project meeting should be at the end of August 2007 in Frankfurt. Staff at the JWG University/Frankfurt will be responsible for the arrangements.

Comments on the decisions concerning the research activities:

The chemists showed very good progress in synthesizing the complete reversible terminators: 3'-*O*-CE-dTTP containing an allyl-glycol linker with Fluorescein was delivered as the first complete reversible terminator. Regarding the second complete reversible terminator, the synthesis of the Texas Red labelled 3'-*O*-CE dATP using the same allyl-glycol linker was currently under way.
Besides the allyl-glycol linker system, which is cleavable with a Palladium catalyst, the chemists at JWG University were developing a fluoride cleavable dye-linker system as an alternative. At the 4th meeting no preferences for the use of one of these two dye-linker systems were decided.
Besides that, it could also be proven that the cyanoethyl(CE)-function as 3'-*O*-protection group for the reversible terminators is incorporated by the polymerase and removed by using 1 M TBAF/THF at 60°C within 5 minutes. This was supported by data presented by JWG University and Fermentas.
Fermentas showed good polymerase incorporation results using 3'-*O*-CE-dATP, 3'-*O*-CE-dCTP and 3'-*O*-dTTP as substrates. The only tested substrate showing no terminating ability, due to its insufficient purity, was 3'-*O*-CE-FITC-dTTP.

Decisions on the research activities taken unanimously by the PMC:

- **Regarding WP1**: The 3'-*O*-CE-dTTP containing the propargylamine moiety will be synthesised at JWG university in larger amounts (> 50 mg) and sent to Fermentas. In Fermentas' lab, the compound will be labeled with Fluorescein and used for repeated incorporation tests and enzyme evolution.

- **Regarding WP2:** Fermentas will strongly focus on the optimization of their enzyme by evolution of the polymerase within the next months: For this purpose, at least three different reversible terminators (dATP, dCTP and dTTP containing a cleavable linker-dye system as well as the 3'-*O*-CE-blocking group) will be needed.

- **Regarding WP1:** Because of the problematic synthesis of 7-deaza-7-iodo-guanine, the complete dGTP-terminator will be delivered as the latest of all four reversible terminators. Nevertheless it is necessary for the enzyme evolution to have the triphosphate of guanine containing the 3'-*O*-CE-group and to test this compound

- **Regarding WP1:** JWG University will focus on completing the synthesis of the three reversible terminators dATP, dCTP and dTTP containing the 3'-*O*-CE-group, the cleavable linker attached to the base moiety and the dye (Fluorescein, Texas Red and Cy3 or Cy5 (optional)). Those three compounds will be sent to Fermentas where enzyme evolution is to be continued. Marek Kwiatkowski (Oligovation) will synthesize 3'-*O*-CE-dGTP and provide it to Fermentas for the enzyme evolution. He will also synthesize several oligomers containing terminal 3'-*O*-CE-dTTP or 3'-*O*-CE-dGTP and test their stability on an array. He will also check the stability of the oligo containing the terminal 3'-*O*-CE group in solution.

- **Regarding WP3:** Silex has to design a third microfluidic device prototype with integrated features adapted to the existing Asper platform.

- **Regarding WP1:** Asper developed a method and software for designing suitable probes for the SBS procedure as well as a base calling software. Silex provided a gold layer-covered silicon chip to Asper. After cleaning and conditioning the surface, thiol-modified oligonucleotides were immobilized onto the chip. The slide was then treated with 1M TBAF/THF solution. Asper showed that the fluoride ions from the TBAF/THF solution cleave the templates from the chip even by using different coatings like the SAL-1 microarray. This led to the decision that Marek Kwiatkowski (Oligovation) will check the availability of suitable glass slides.

- **Regarding WP4:** Tartu University will design a suitable hairpin type oligonucleotide and provide it to Fermentas and to Asper. Tartu also will create a method for the sequencing of hairpins based on their already achieved work (base calling software etc.).

- **Regarding WP5:** Asper makes hairpin experiments with 100 hairpins. The optimal sequences for those hairpins will be provided by the bioinformatics team from Tartu University.

- **Regarding WP5:** Asper also will check single strand and double strand stability of the templates under the deprotection conditions by several array experiments.

4.1.5 Fifth ArraySBS project meeting in Frankfurt/Germany (31.08.2007 - 01.09.2007)

11 persons participated in the 5th project meeting.

Decisions regarding the Project management:

Decision: Every project member of each work package, named as the following:

- Asper
- Fermentas
- JWG University
- Oligovation (former Quiatech)
- Silex
- Tartu University

has to write a report of their spending and send it to the coordinator, Prof. Dr. J. W. Engels until the 30th September 2007.

Besides that, the following topics were discussed at 1st September 2007 by all project members:

- Deliverables list
- Patent
- Last ArraySBS Meeting

Deliverables list:
The deliverables list was updated during the 5th meeting: Regarding WP2, the deliverable D19 was postponed from month 24 to month 28 (see attached deliverables list)

Patent:
Fermentas and Oligovation will decide at the end of the project, if a second patent containing Fermentas' polymerase and the reversible terminator chemistry will be filed. Both companies (Fermentas and Oligovation) would be the owners of this patent then.

Last PMC Meeting of the ArraySBS project:
Decision: The 6th project meeting will be held in Copenhagen/Denmark at the 11th January 2008. Asper will be responsible for the arrangements.

Comments on the decisions concerning the research activities:

As the chemists had already reported at the 4th meeting, the 3'-*O*-CE-dTTP containing an allyl-glycol linker with Fluoresceine was delivered as the first complete reversible terminator. It was reported then that the allyl-containing linker system is already used by other SBS-researchers (e.g. Burgess *et al.*). So the project members decided to drop the idea of a palladium-cleavable linker. They started exclusively using their own invented, fluoride cleavable cyanoethyl (CE) - linker for the rest of the SBS project. Compared to the allyl-linker, the fluoride-cleavable linker has also the advantage of being simultaneously cleaved with the 3'-*O*-CE-blocking group via TBAF/THF-treatment.
Fermentas confirmed by presenting their experimental data from the polymerase acceptance tests the successful incorporation of all 3'-*O*-CE-dNTPs (A, C, G and T *without* linker and dye-system) at 37°C within 5 minutes in buffered solution.
The chemists reported that they synthesized a very small amount (less than 30 mg) of 7-deaza-7-iodo-2'-deoxy-guanosine. For passing the next synthetic steps up to the complete reversible terminator of G, roughly 2 to 5 g of this nucleoside as starting material are needed.
Fermentas also tested the incorporation of modified nucleotides in a microarray mode. They used SAL-type slides (as already known by Asper) as well as CodeLink slides (better suitable for this

chemistry) for attaching their oligonucleotides, designed with the help of Tartu University. These prepared slides were used for incorporation tests using Fluorescein-labelled-3'-*O*-CE-dTTP and the three unlabelled 3'-*O*-dNTPs (A, C and G). Although the incorporation worked on this microarray, it was very slow (60 minutes each cycle). To fasten the incorporation reaction, it was suggested to use a 17 nucleotide linker instead of the 12 nucleotide linker or to attach the hairpin templates to the slide in the loop region. The idea behind this was that this geometry would offer more space for the polymerase, so that it would be less sterically hindered during the incorporation reaction.
Another general problem with immobilized, double-stranded templates appeared during the CE-cleavage tests done at Asper: The double-stranded structure of the template was destroyed during the treatment with 1M TBAF/THF solution. Regarding this problem Asper should check if *hairpin-structured* oligonucleotides are stable during the TBAF/THF treatment.
The third microfluidic device was delivered to Asper and Tartu University in three different volumes and the coating for its wafer surface was Teflon for each device. The reaction chamber was more rounded than the chamber of the 2nd device prototype. Still a detailed and complete design of that device containing valves, mixers, heaters and pumps was missing. The integration of these features (pumps, valves etc.) would be possible by established information about the reaction conditions (especially the fluoride cleavages) of the microarray inside the chamber as well as certain knowledge from Asper regarding the fluorescence measurement technique.
At the bioinformatics group from Tartu University, two statistical models to predict the relative signal intensity and the call rate of each primer candidate for hairpin sequences were developed. Both statistical models were tested on non-training data- sets. Besides that, Remms group developed a software prototype for the automatic design of primers for re-sequencing. This program is able to predict primer sequences and the call rate for the region of interest on an oligomer.

Decisions on the research activities taken unanimously by the PMC:

- **Regarding WP1**: At least three complete reversible terminators, A, C and T containing the fluoride cleavable linker, a base-specific dye and the cyanoethyl-function as the 3'-*O*-blocking group, should be available for Fermentas tests at the end of the year (December 2007). The chemists at JWG University have to synthesize the three complete reversible terminators and send them to Fermentas. The reversible terminators of the nucleobases A and T will be the first ones, followed by C.

- **Regarding WP1**: JWG University will buy 2 g of 7-deaza-7-iodo-2'-deoxy-guanosine after the successful synthesis of all three complete reversible terminators A, C and T. With the experience gained by the preparation of these compounds, the chemists at JWG University will try to synthesize (hopefully within the timeframe of this project) the last complete reversible terminator (G).

- **Regarding WP1**: Oligovation will synthesize hairpin-templates that can be attached in the loop- region to the slide. The sequences of these hairpins are already delivered by WP4. The chemists also have to check the DNA-structure stability of the duplex during the TBAF/THF treatment using different conditions.

- **Regarding WP1, WP2, WP4 and WP5:** If the three reversible terminators of A,C and T are delivered, Fermentas will perform in cooperation with Asper the proof-of-principle of re-sequencing with these terminators. To support this, Remm's group will have to design 64 hairpins to make 3 SBS-cycles possible.

- **Regarding WP2**: Because of the delayed availability of the three terminators, deliverable D19 (the final enzyme improvement steps) is shifted to month 28.

- **Regarding WP1, WP3 and WP5**: In cooperation with Asper and Silex, Oligovation will care about integration of tools like valves, mixers and pumps to the existing 3rd microfluidic device.

- **Regarding WP4**: With their software and the statistical models, the co-workers from Tartu University have to design at least 64 hairpin primers for the proof-of-principle.

4.1.6 Sixth ArraySBS project meeting in Copenhagen/Denmark (11.01.2008)

9 persons participated in the 6th project meeting.

During the final meeting, the latest results from all work packages were presented. The final report was discussed as well as the planned publication of the scientific results achieved during the project.

Decisions regarding the Project management:

Final Report

Decision: Every project member of each work package, named as the following:

- Asper
- Fermentas
- JWG University
- Oligovation (former Quiatech)
- Silex
- Tartu University

has to contribute their part regarding the PAR, PMR, form C and audit certificate by sending it to the project coordinator at the end of the project.
For the delivery of the Periodic Activity Report and the Periodic Management Report the coordinator had set the deadline up to the 7th March 2008.
For the delivery of the form C and the audit certificate the coordinator had set the deadline up to the 21th March 2008.

Scientific Outlook

Decision: Complete reversibly terminating, dye-labelled dGTP will not be available within the remaining timeframe of the project. It will be synthesised by JWG University and provided to Fermentas for incorporation experiments after the end of the project.

Decision: The aim of the project is a cooperate publication in e.g. *Nature Biotechnology,* therefore also data about the dGTP is needed.

Decisions on the research activities taken unanimously by the PMC:

- **Regarding WP1:** Providing 3'-*O*-CE-dATP with fluoride cleavable linker and TexasRed attached until 21st January.
- **Regarding WP1:** Providing 3'-*O*-CE-dCTP with fluoride cleavable linker and Cy3 attached until first week of February.
- **Regarding WP1:** Supporting Asper in performing cleavage tests by providing all data about cleavage conditions tested at JWG University.
- **Regarding WP1:** Oligovation should accomplish the experiments for testing the duplex stability under cleavage conditions (1 M TBAF in dry THF at 58°C) as soon as possible. Oligovation should also write a non-confidential project summary.
- **Regarding WP2 and WP5:** Both contractors should perform incorporation/cleavage experiments to determine incorporation and cleavage efficiency.

- **Regarding WP5:** The contractor has to do the proof-of-principle using normal 5'-attached hairpins with 17 nt spacer designed at Tartu University and with the use of three reversible terminators as soon as they are available and also the best polymerase from Fermentas.
- **Regarding WP4:** Tartu University has to finish its software prototype for handling image data from SBS experiments independently.
- **Regarding WP3:** It is not necessary to go on working on refining the third microfluidic device since the point to test it will not be reached within the timeframe of the project.

4.1.7 Visits and communication

No specific monitoring was done by the coordinator regarding the use of man power or other resources until the end of the whole project period. The parties in the project have been well interacting with each other the RTD parties (Tartu U., Prof. Engels group and Silex) have all contributed a lot to the increased value for the SME's. Six project meetings have been held, the kick off meeting in Frankfurt, 31.08. - 02.09.2005, the 2nd meeting in Vilnius, 01.03. - 03.03.2006, the 3rd in Uppsala, 23.08. - 25.08.2006, the 4th meeting in Tartu (Estonia), 08.03. - 09.03.2007, the 5th meeting in Frankfurt (Germany), 31.08. - 01.09.2007 and the final (6th meeting) in Copenhagen (Denmark), 11.01.2008. Besides these project meetings, the contractors often held telephone conferences via the web-phone program "Skype" which is known as very cheap communication medium. Frequently held telephone conferences and vivid email contact caused that most of the scientific problems could be discussed and solved besides the conferences. All important chemical devices, i. e. the reversible terminators demanded for the polymerase evolution tests, were shipped via express service to Fermentas so no extra traveling costs were needed.

4.1.8 Consortium Management – Problems

During spring 2006 Quiatech got into financial trouble and the company was reorganized, so all staff got a notice to leave the company and most did. The coordinator person left Quiatech late August. From the first of September there is only one person responsible at Quiatech. As it was decided at the 4th meeting in Uppsala (Sweden) then, Prof. Dr. J. Engels was elected as the new coordinator and contact person. During the refinancing of Quiatech the organisation number and the company name was changed, so Quiatech is known as Oligovation since then. Visits and communication, the organization of the scientific collaboration and other management issues like writing the final report were coordinated at Frankfurt University. Changes in the consortium management as well as delayed results in chemistry due to the former used thiol-chemistry, which had been sold to a third party, led to a time-delay in fulfilling some project tasks. Therefore the new coordinator applied for a time extension of the project of 6 months.

4.1.9 Major achievements concerning dissemination of knowledge

The major achievement is the new patent application filed on the 1st of September 2006 with inventors from the Engels group (RTD), and the SME's Fermentas and Oligovation (former Quiatech), where both these SME's will benefit from the possibilities of commercialization of this invention.

4.2 Use of person-month and resources

Person-month per Work package tasks:

WP / task names	**Person-months planned**[1]	**Person-months spent**[1]
WP1 Synthesis of dye-labelled reversible terminator	**36**	**100.5**
task1.1: Synthesis of first 3'MRT (D3)	2	2
task1.2: Synthesis of biotinylated 3'MRT (D5)	4	4
task1.3: Synthesis of second 3'MRT (D8)	7	9
task1.4: Synthesis of third 3'MRT (D8)	7	8
task1.5: Optionally synthesis of fourth 3'MRT (D8)	not planned (optionally)	5
task1.6: Synthesis of 3'MRT with dye (D12)	4	20
task1.7: Synthesis of the remaining 3'-*O*-modified reversible terminators with dyes (D21)	6	31
task1.8: Synthesis of (optionally) fourth 3'-*O*-modified reversible terminators with dyes (D21)	3	10.5
task1.9: Set up large-scale and purification methods of the 3'-*O*-modified reversible terminator (D21)	3	11
WP2 Identification and improvement of DNA polymerase	**116**	**116**
task2.1: Assay development (D4)	4	4
task2.2: Id DNA polymerase (D9)	25	25
task2.3: Purification and cloning of best polymerase (D14)	13	0, 13 person-month shifted to tasks of D19[2]
task2.4: Technique development for improvement (D19)	20 + 13 from D19	38
task2.5: Improving the DNA polymerase by gene modification (D19)	39	36
task2.6: Purification of the improved polymerase (D19)	15	13
WP3 Development of a microfluidic device for reaction	**11**	**11.4**
task3.1: Draw first prototype (D7)	1	1
task3.2: Make first prototype (D7)	2	2
task3.3: Draw second prototype (D11)	1	0.75
task3.4: Make second prototype (D11)	4	3.65
task3.5: Draw final prototype (D18)	0.5	1
task3.6: Make final prototype (D18)	2.5	3
WP4 Development of bioinformatics tools and selection of primers	**42**	**82.5**
task4.1: Method development step1 (D10)	5	8
task4.2: Implement primer design software (D10)	9	13
task4.3: Improve parameter settings (D15)	12	19.5

task4.4: Modify Base Caller for ArraySBS application (D22)	6	22
task4.5: Identify primers for p53 re-sequencing (D22)	5	12.5
task4.6: Analyse data from instrument (D22)	5	7.5
WP5 Development of a re-sequencing assay	**42**	**87.5**
task5.1: Development of SBS method on APEX chips (D16)	6	20
task5.2: Implement first SBS setting (D16)	10	47.5
task5.3: Implement 2nd SBS settings (D20)	9	10, in progress[3]
task5.4: Implementation of p53 re-sequencing (D20)	17	10, in progress[4]
WP6 Project administration	**4**	**4.9**
task6.1: daily general project management issues	0.4	0.5
task6.2: Establish a web page (D1, D2)	0.2	0.4
task6.3: Design and publication of a project leaflet (D6)	0.15	0, cancelled[5]
task6.4: Preparing an intermediate plan for using and disseminating knowledge (D13) and first non-confidential project summary (D17)	1.25	2
task6.5: Writing final public scientific report and a plan for disseminating knowledge (D23 and D24)	2	2

[1] The contribution to one task can be a combination of work from the contractors of the corresponding work package.

[2] Because of the availability of the required genes the work at task2.3 (D14) was stopped. The time and person-month were used for starting the work at task2.4 (D19) earlier than planned.

[3] Because of significant changes from initial plans regarding the chemistry used for terminator and fluorescence group removal the time delay for this deliverable was first estimated to month 18. Further stability problems of slides, immobilization and duplexes caused an additional delay to month 30. Now the experiments are still in progress.

[4] Significant changes from initial plans in regard to chemistry used for terminator and dye-linker group removal led to a big time delay which disabled the re-sequencing of the p53 gene within the timeframe of the project.

[5] The objective for this task/deliverable was to spread the message about the project to a public audience in order to find potential commercialization partners. The project was announced with the press release 24th August 2005 at the Quiatech website and by Tartu University. It was also sent to several genomic journals and it was published at the Genome Website. Due to this well spread of the activities two third parties contacted Quiatech and began to discuss business opportunities. With ongoing business discussion it would have been wrong to publish a leaflet. Thus no project leaflet was written.

In the technical annex it was planned for the full duration of the project, that the person-month to be spent by the RTD's should be approximately 33% of the total person-month. It turned out to be about 48% instead.

The major reason for the discrepancy between the foreseen and the spent person-month is the strong contribution from the RTD's, JWG U (Prof. Engels group) who have employed three persons, one diploma student (without payment), later also PhD student, one PhD student and one Post Doc and Tartu University (Prof. Remm group) who also has employed one PhD student and 4 diploma students (low payment) in the first year and the double number of PhD students as planned in the second year.

The reason for the high contribution of JWG U was the transfer of work from the SME partner Oligovation (former Quiatech). Since Quiatech got into financial problems and is now named Oligovation only few employees are left in this company. This also lowers the contribution of the SME partners to 52% instead of 67%.
The high discrepancy in use of person-month in WP5 is due to the change in terminator chemistry and the resulting problems which are discussed in the main achievements and problems parts in the work package description form of WP5 (section 3). To solve these problems more person-month than planned are necessary.
As the work done by diploma or PhD students is without or with only low payment, the costs for the salaries stayed within the planned frame and the increase of person-month has no impact on resources.

Costs and resources for the full project period:

Workpartner	Type of Expenditure	Actual Costs (EUR)		
		Period 1	Period 2	Total
JWG University	Personnel	54,844.62	132,189.21	187,033.83
	Consumables	23,995.98	33,060.63	57,056.61
	Durable equipment	0.00	0.00	0.00
	Travel	3,123.14[1]	4,719.04[4]	7,842.18
	Other costs	-	-	-
	Management / Subcontracting	-	3,517.96 / 0.00	3,517.96
	Sub-Total	81,963.74	173,486.84	255,450.58
	Overheads	16,392.75	34,397.87	50,790.62
	Total Costs	**98,356.48**	**207,884.73**	**306,241.21**
Oligovation (former Quiatech)	Personnel	26,733.75	34,584.00	61,317.75
	Consumables	7,229.87	5,481.00	12,710.87
	Durable equipment	5,356.81[3]	0.00	5,356.81
	Travel	5,625.95[1]	669.00[4]	6,294.95
	Other costs	-	988.00	988.00
	Management / Subcontracting	18,806.33 / 2,000.00	3,037.00 / 2,000.00	25,843.33
	Sub-Total	65,752.71	46,726.00	112,478.71
	Overheads	12,750.55	8,945.00	21,695.55
	Total Costs	**78,503.26**	**55,671.00**	**134,174.26**
Fermentas UAB	Personnel	81,390.02	149,732.72	231,122.74[5]
	Consumables	26,536.43	61,677.63	88,214.06
	Durable equipment	7,025.84	10,555.66	17,581.50[6]
	Travel	5,605.38[1]	5,456.89[4]	11,062.27
	Other costs	-	-	-
	Management / Subcontracting	-	0.00 / 2,600.00	2,600.00
	Sub-Total	120,557.67	230,022.90	350,580.57
	Overheads	24,111.53	45,484.58	69,596.11
	Total Costs	**144,669.20**	**275,507.48**	**420,176.68**
Asper	Personnel	37,946.99	66,397.95	104,344.94
	Consumables	52,046.69	14,964.25	67,010.94
	Durable equipment	15,169.02	13,548.14	28,717.16
	Travel	-[2]	1,508.24[4]	1,508.24
	Other costs	-	-	-
	Management / Subcontracting	0.00 / 319.56	0.00 / 479.34	798.90
	Sub-Total	105,482.26	96,897.92	202,380.18
	Overheads	21,032.54	19,283.72	40,316.26
	Total Costs	**126,514.80**	**116,181.67**	**242,696.44**

Workpartner	Type of Expenditure	Actual Costs (EUR)		
		Period 1	Period 2	Total
Tartu University	Personnel	32,793.36	58,543.14	91,336.50
	Consumables	4,608.87	3,117.19	7,726.06
	Durable equipment	0.00	0.00	0.00
	Travel	12,452.53[1]	1,944.88[4]	14,397.41
	Other costs	-	-	-
	Management / Subcontracting	-	0.00 / 1,789.53	1,789.53
	Sub-Total	49,854.76	65,394.75	115,249.51
	Overheads	9,970.95	12,721.04	22,691.99
	Total Costs	**59,825.71**	**78,115.79**	**137,941.50**
Silex	Personnel	35,016.00	50,482.00	85,498.00
	Consumables	2,427.00	2,496.00	4,923.00
	Durable equipment	0.00	0.00	0.00
	Travel	2,015.00[1]	1,055.00[4]	3,070.00
	Other costs[7]	26,582.00	25,656.00	52,238.00
	Management / Subcontracting	-	-	-
	Sub-Total	66,040.00	79,689.00	145,729.00
	Overheads	13,208.00	15,938.00	29,146.00
	Total Costs	**79,248.00**	**95,627.00**	**174,875.00**
Total Project Costs	**Total Eligible Costs**	**587,117.45**	**828,987.67**	**1,416,105.12**
	Requested EC Contribution	**385,093.20**	**561,855.57**	**874,089.00**

[1] The travel costs from Quiatech, JWG U and Fermentas were exclusively used for the three project meetings. One additional visit took place between Asper and Silex (see 4.1.4 visits and communication). Tartu U used the travel expenses for the three project meeting and additional meetings in Leicester, UK (SNP2005 conference), Madrid, Spain (Bio sapiens 3rd Permanent School in Bioinformatics) and Arlington, US (Nucleic Acid-based Technologies 2006).

[2] Unfortunately Asper did not show the travel costs in detail which are here part of the consumables.

[3] The money for the Equipment was spent for the items listed in the DoW. Quiatech acquired a semi-preparative HPLC system and Asper acquired different items for its laser setup and a hybridisation station.

[4] The travel costs were exclusively used for the three ArraySBS project meetings that took place during the second project period (Tartu, Frankfurt, Copenhagen).

[5] The deviation from planned budget regarding personnel cost is approx. 18%, due to approx. 24% increase in Fermentas' staff salaries during the period between years 2004 and 2007.

[6] Costs for equipment rose from 6,100 EUR to 17,581.50 EUR mainly due to acquiring of new instrumentation specific for project demands (two temperature-controlling vertical electrophoresis systems) not foreseen during the preparation of project application. In addition, expenses for durable equipment increased considerably due to the extension of the project (30 months instead of 24).

[7] "Other costs" in this case represent manufacturing costs, i.e. all expenses related to the fabrication of the components, including material used for production like chemicals or gases, use of equipments and cleanroom facility. Also personal costs related to operators, process engineers and technicians are included.

During the whole duration of the project 1,416,105.12 € were spent by all partners. The RTD performers spent approximately 44% and the SME's 56%. The requested EC contribution amounts 874,089.00 €.

The RTD's spent 619,057.71 € which is 54,935.31 € more than planned eligible costs in the technical annex of the DoW (564,122.40 €). This is due to the elongation of the project for additional 6 month and caused by additional personnel costs and consumables.

The SME's spent 797,047.41 € which is 21,076.99 € less than the planned eligible costs in the technical annex of the DoW (818,124.40 €). This is due to the reduction of employees and done work at Quiatech/Oligovation.

5. Other issues

5.1. Safety

In all cases the scientists in the project assumed the necessary responsibilities for safety. The partners ensure that all National and EU guidelines are followed.
All partners undertake to respect international safety regulations, including European standards and national regulations concerning research, safety of workers and researchers and waste disposal at laboratories. All work was carried out in accordance with the Control of Substances Hazardous to Health (COSHH) act 1990 and all other National and Safety at Work legislation.
The parties in the project have been well interacting with each other the RTD parties (Tartu U., Prof. Engels group and Silex) have all contributed a lot to the increased value for the SME's. The major achievement is the new patent application filed with inventors from the RTD, Engels group, and the SME's Fermentas and Quiatech, where both these SME's will benefit from the possibilities of commercialization of this invention.

5.2. Gender issues

5.2.1 First year

Two of the new employees at JWG U are females. Of two new employees at Fermentas one is male and one is female.

		Males	Females
SME	Asper	5	2
	Quiatech	4	1
	Fermentas	4	3
RTD	JWG U	1	2
	Silex	5	1
	Tartu U	5	1
	Totally	**24**	**10**
		70%	**29%**

5.2.2 Second year

During the second project year, one new female employee joined the researcher group at JWG University.

		Males	Females
SME	Asper	5	2
	Quiatech	4	1
	Fermentas	4	3
RTD	JWG U	1	3
	Silex	5	1
	Tartu U	5	1
	Totally	**24**	**11**
		69%	**31%**

Annex – Plan for using and disseminating the knowledge

Section 1: Exploitable knowledge and its use

Exploitable Knowledge (description)	Exploitable product(s) or measure(s)	Sector(s) of application	Timetable for commercial use	Patents or other IPR protection	Owner and Other partners involved
1. Enzymatic incorporation of 3'-*O*-CE-dNTPs	Enzyme (DNA polymerase)	1. Academic 2. Industrial	Too early for estimation	Not yet	Fermentas
2. Improved enzymatic incorporation of 3'-*O*-CE-dNTPs	Mutant DNA polymerase which incorporates better (faster)	1. Academic 2. Industrial	Too early for estimation	Not yet	Fermentas
3. Reversible terminators useful for sequencing-by-synthesis	Three fluorescently labelled dNTPs carrying removable 3'-*O*-CE-modification	1. Academic 2. Industrial	Since spring 2007, we are in contact with a major Biotech company (not part of the Consortium) who wished to evaluate our reversible terminators	Patent application filed	Oligovation, Fermentas, JWG U
4. System for DNA sequence determination using sequencing-by-synthesis approach	The Kit composed of 4 fluorescently labelled nucleotides and a polymerase which incorporates these nucleotides	1. Academic 2. Industrial	Too early for estimation	Not yet	Fermentas, Oligovation
5. Software for re-sequencing and primer design applicable for the ArraySBS procedure	A well functioning software for re-sequencing	1. Academic 2. Industrial	Too early for estimation	Not yet	Tartu University

1. (Fermentas)

The DNA polymerase which incorporates nucleotides bearing the fluoride cleavable 3'-*O*-CE-group could be used in sequencing-by-synthesis experiments. The Search for such enzymes which incorporate 3'-*O*-CE-dTTP (and 3'-*O*-CEM-dTTP) resulted in several hits. During the second project period, the incorporation rate could be improved by using the three complete dye-labelled 3'-*O*-CE-dNTPs (with N = A, C, T) for the incorporation tests. At the moment it is still not proven if the identified polymerases are suitable for incorporation of the fourth reversible terminator, Cy5-labelled 3'-*O*-CE-dGTP,.which will be tested after the project end due to the late availability of the terminator.

2. (Fermentas)

The DNA polymerase, which incorporates modified nucleotides faster and better, would shorten each step of sequencing-by-synthesis procedure. To evolve such improved polymerase, the efficient evolution *in vitro* scheme needed to be elaborated and applied. In the first year such a scheme (which is unique and combines features of several evolution *in vitro* approaches) was developed, and now we are adopting this scheme for isolation of polymerase mutants which incorporate biotinylated nucleotides with the CE block on their 3'-ends.

3. (Oligovation, JWG University, Fermentas)

As it was found that terminators bearing the cyanoethyl-function as 3'-*O*-modification were incorporated best it was strongly focused on the CE-group for further project tasks. Besides that, a lot of time-demanding work was put in engineering a CE based cleavable linker connecting the respective fluorophore linked to the base. The chemistry about the cleavable linker and the complete synthesis of the first complete reversible terminator is already filed as a patent and might be commercially used by a third party.

4. (Oligovation, Fermentas)

One of presumable final products of this project is a kit which contains modified nucleotides and a polymerase which incorporates these nucleotides, instructions how to perform incorporation, and instructions how to remove both the 3'-*O*-CE group and the fluorescent label from extended DNA strands. The final polymerase optimization will be done when the fourth complete reversible terminator (Cy5-labelled 3'-*O*-CE-dGTP) will be delivered to Fermentas. This issue could not be achieved within the timeframe of the project, but the still existing collaboration between Frankfurt University and Fermentas will fulfil this task which might lead to a commercially useful sequencing kit.

5. (Tartu University)

After the second project year a well functioning software tool for array based sequencing-by-synthesis and primer design has been established. Several algorithms have been tested for this purpose and it was found that the most appropriate algorithm for genotyping is the EM algorithm, which uses equal shape and volume of genotype clusters. Named algorithms were implemented into BsPred software and adapted for base calling of re-sequencing data. The web interface, that means the graphical user interface for base calling for BsPred was implemented with BsPred_common model and is available at http://bioinfo.ut.ee/bspred/.

Section 2: Dissemination of knowledge and publishable results

Year	Type	Type of audience	Countries addressed	Partner responsible/ involved
Media and Press Dissemination				
Sept. 05	Press release	Journalists and newspapers	Europe	Quiatech
Sept. 05	(Radio interviews)	General audience	Sweden	Quiatech
Sept. 05	(News articles)	General audience	Estonia	Asper
Conferences organized by Array SBS				
31.08.-02.09.05	1st ArraySBS conference (kick-off), Frankfurt	Scientists	Sweden, Germany, Estonia, Lithuania	JWG University
01.03.-03.03.06	2nd ArraySBS conference, Vilnius/Lithuania	Scientists	Sweden, Germany, Estonia, Lithuania	Fermentas
23.08.-25.08.06	3rd ArraySBS conference, Uppsala/Sweden	Scientists	Sweden, Germany, Estonia, Lithuania	Quiatech
08.03.-09.03.07	4th ArraySBS conference, Tartu/Estonia	Scientists	Sweden, Germany, Estonia, Lithuania	Tartu University
31.08.-01.09.07	5th ArraySBS conference, Frankfurt/Germany	Scientists	Sweden, Germany, Estonia, Lithuania	Frankfurt University
11.01.08	6th ArraySBS conference, Copenhagen/Den-mark	Scientists	Sweden, Germany, Estonia, Lithuania	Asper
Website				
Sept. 05	Project website http://bioinfo.ebc.ee/ArraySBS/	General public	All	Oligovation (Quiatech), Tartu U
Publications and patents				
04.09.06	Patent, filed	US patent 60/824453, Title: "Reversible Terminators for efficient Sequencing by Synthesis"	All, not published yet	Oligovation (Quiatech)
Mar. 07	Research article, accepted 06.12.2006	"The fluoride cleavable CEM group as reversible 3'-*O*-terminator for DNA sequencing-by-synthesis – Synthesis, Incorporation and Cleavage", Nucleosides, Nucleotides and Nucleic Acids **2006,** *26,* 271-275	Sweden, Germany, Lithuania	Frankfurt University
First half 08	Research article, planned in a scientific journal like *Nature Biotechnology*	Combination of chemistry, polymerase and experiments on the slides (proof-of-principle)	Sweden, Germany, Lithuania, Estonia	Frankfurt University

Conference presentations by ArraySBS partners on ArraySBS research				
03.09.-07.09.06	17. Int. Roundtable on Nucleosides, Nucleotides and Nucleic Acids	Scientists	Bern/Switzerland	JWG U