Publishable summary 1st Periodic Report (01/06/2008 – 31/05/2009)

www.cng.fr/READNA

The READNA (REvolutionary Approaches and Devices for Nucleic Acid analysis) project got off to a flying start with all of the activities reaching full operation within a few months. READNA is structured in five work packages with the scientific goals to aid the implementation of the current generation of nucleic acid sequencing technology, the advancement of nucleic acid analysis technology towards the implementation for diagnostics, the development of 3rd generation nucleic acid sequencing technology and the development of methods to tackle the needs of nucleic acid analysis that are currently not met. The scientific work packages are supported by a work package that takes care of the management of the consortium and a work package on the dissemination of project results. In this first reporting period substantial milestones and deliverables have been achieved and numerous results have been published, have been submitted for publication or filed for patent protection.

WP1: Near Term Innovations: Ancillary elements for 2nd generation DNA sequencers

Work package 1 has evolved after its first year to now concentrate primarily on how best to amplify or enrich genome sub-regions so that they provide input material for 2nd generation DNA sequencing. This work spans activities 1-3 (of 7 activities in this WP) To synergise and initiate optimally these tasks we organised a 2-day workshop on the topic. This helped all Partners understand the state-of-the-art and share their ideas for solutions. This joint effort is now being extended by a dedicated email and forum managed in conjunction with the GEN2PHEN project (www.gen2phen.org). Lab work has begun on activities 1-3, and useful progress made with respect to 'MegaPlex PCR', 'Selectors', and some new protocols for conducting 'Hybridisation-Based DNA Enrichment' upon arrays and in solution. Considerable benchmarking and utility comparisons of external methods has also been undertaken. Other work in this WP is either not yet due to begin (Activity 4: 'Rare Mutation Enrichment') or completed ahead of schedule (Activities 5 & 7, concerning 'Amplified Single-Molecule Arrays' and 'Ethical Standards'). Activity 6, which deals with 'Data Complexity Reduction by Standardizing Data Output' has been de-prioritised after liaison with international standards consortia who are already dealing with that objective.

WP2: Near Term Innovations: Improvement and extension of existing methods

This WP elaborates on protocols that were well advanced before the start of the READNA project. The objective of this work is to, improve, stabilize and implement these protocols with a view to make them amenable to clinical application and diagnostics. Five distinct tasks are undertaken: 1) the development of ribo-PCR sequencing, which is a single reaction step DNA sequencing protocol starting from genomic DNA as a template that uses MALDI mass spectrometry for readout, 2) Implementation of HLA-typing based on µ-haplotyping, where short, phased sequence elements are analysed at selected positions in the highly variable HLA genes. The readout is made by MALDI mass spectrometry that is well adapted to the analysis of many different distinct species in a single analysis, 3) Translation of Dynamic Allele-

Specific Hybridisation (DASH) into a version that can be used for resequencing. The concept is that a reference sequence is prepared on an array, a tester sample hybridized to it and dynamically released from it by heating the array, 4) the implementation of the Datta-array picolitre reactors for counting-based DNA analysis applications, 5) the combination of multiple levels of biomolecule analysis (DNA polymorphism, copy number, transcript, protein profile) on a single device and joining the information for a diagnostic.

All of the concepts are working according to the initial plan and progress has been faster than anticipated.

WP3: Fluorescence-based Single Molecule Sequencing

Work package 3 aims to develop innovations based on single molecule fluorescence analysis that will result in a next (3rd) generation of sequencing technologies. The main objectives of the work package are to (**1**) develop massively scalable techniques that further reduce the cost and increase the throughput of DNA sequencing and (**2**) that provide sequence reads that are long or are obtained within their long-range context. During the first year of the project substantial progress has been made in integrating the interdisciplinary research efforts of the project partners towards the objectives of the work package. This effort has enabled progress to be made towards the early objectives of the work package. On one front, we have set up model systems and read-out technologies that have allowed us to obtain our first measurements of the novel sequencing biochemistries being developed whilst on another front we have designed and constructed micro/nanofluidic chips that have allowed progress to be made towards handling and visualizing very long pieces of genomic DNA. One paper has been published and others are under review.

WP4: Nanopore Sequencing

Work package 4 targets technologies that will enable the study of single molecules of DNA as they interact with nanometer-scale holes or 'nanopores', through which an ionic current is being passed. The WP breaks down into four areas of research with all showing measureable progress during the first year of the Project. We reported on the identification and selection of an auxiliary protein which processes DNA into its compontent parts to aid control of movement through a nanopore, and we published a paper in Nature Nanotechnology which demonstrated that a nanopore can identify these component parts with accuracy. We also made progress with the development of stable membranes to support nanopore proteins, a key activity for building a commercially viable product. One method we have investigated is the integration of protein nanopores into an interface between two water droplets in an oil/lipid mixture. Droplets have been brought into direct contact with each other, forming a lipid bilayer at their interface and a nanopore protein has been demonstrated to reliably insert to form a highly stable connection between them. This enables many new experiments to be envisaged and investigated to see how such connections might be exploited. We continue to work on the identification of modified DNA bases and on the integration of protein nanopores into apertures fashioned in sheets of solid material.

WP5: New Genotyping Challenges

The work in this WP is devoted to solving genotyping and epi-genotyping challenges, where the performance of current technology is not satisfactory. Existing nucleic acids analysis methods are adapted for DNA methylation analysis and copy number variation analysis where

analysis has to deliver very precise quantitative results. Methods are also develop to analyse mutant nucleic acid molecules under challenging conditions such as in situ in histological sections or in a high background of wildtype molecules. The technologies will serve for genome-wide-, targeted multiplexed-, and clinical analysis needs. The work is progressing very well and is over all following the original plan. The work in the WP is well integrated and particularly the SME Olink is actively collaborating with several partners in the consortium. One paper has been published, and several manuscripts have been submitted and one patent has been filed based on work in the WP during the first 12 months.

WP6: Training and dissemination

Several project meetings and workshops have been held. READNA contributed to the public "3rd Paris Workshop on Genomic Epidemiology" in Paris (25.-27. March 2009) where a number of READNA partners presented their latest developments in DNA analysis technology. This 3-day workshop had close to 200 participants from 20 countries and was great success. A smaller semi-public workshop that dealt with sequencing standards and the issues surrounding targeted enrichment was held in Berlin 15./16. January 2009 with 35 participants of which 25 were from the READNA consortium and 10 were external to the project. It managed to highlight many of the issues that scientists are facing in applying the current 2nd generation of DNA sequencers. A manuscript summarizing the discussions is currently being prepared. READNA also sponsored speakers for the "10th International Symposium on Mutations in the Genome". The mobility between different partners has been initiated through two mobility awards. A READNA website has been established that provides descriptions of the various activities and progress towards the objectives.

WP7: Management

The structure of the project has been set and a mode of operation achieved that allows smooth and constructive interaction of the partners with each other and the project management. The academic and industrial partners are interacting well. Each of the work packages has held regular meetings and the necessary interactions and bonds have been created. Mechanisms for tracking and guiding the project have been established.