PROJECT INTERIM REPORT

Grant Agreement number:	**HEALTH-2010-261123**
Project acronym:	**GEUVADIS**
Project title:	**Genetic European Variation in Disease**
Funding Scheme:	**Coordination and Support Action**
Start date of project:	**01.10.2010**
Duration: to	**31.09.2013**
Interim report:	**1st**
Period covered:	from **01.10.2010** to **01.11.2011**

This interim report will be provided to our SAB as well as to our EU Project Officer

Summary of Work progress and achievements from October 2010 to November 2011

Note that none of the partners except partner 7 reported any deviations from the tasks planned in Annex I - Description of work.

Oral and poster presentations are detailed in each partner's report. All journal and book publications are reported at the end of the section.

Partner 1, Center for Genomic Regulation, CRG Project coordinator, leader of WP1

a) Summary of tasks performed

WP1 Coordination and Management (WPL)
We ensured the daily management and coordination of the project. We monitored all activities through an efficient communication strategy (Described in Annex 3 Deliverable D1.3). The main tools we put in place are:
- Six specific Mailing Lists, trough which one can target the recipient of specific information within one WP.
- Regular Teleconferences, with updates on the project's activities and discussion on WP tasks implementation and future collaborations.
- A Website, comprising of a public page (including 'resources' sections dedicated to the general public, as well as a podcast explaining the main concept of the project) and an intranet accessible to consortium members only, and gathering all useful information on the project; as well as a Wiki - created by UNIGE team to monitor all activities within WP4- (Website described in Annex 4 Deliverable D1.4
- A Newsletter (biannual)
- Press releases (the first one was published during the KO meeting, we plan to launch the second one once main results are obtained in the different WP.

WP2 Quality control of sequence Data
The CRG team participated in the elaboration of the questionnaire prepared by UU to assess the QC methods employed in the different labs.

WP4 Handling, analysis and interpretation of RNA-sequence data
CRG actively participated in this WP, being involved in the setup as well as in the implementation of the different activities of this WP.
- Pilot study
- 500 samples study
    - miRNA library prep

CRG collaborated with UNIGE in establishing a suitable protocol to extract total RNA including the micro RNA fraction. The CRG also participated in the discussion to decide the minimum quality criteria to be met by all teams regarding the preparation and sequencing of small RNA libraries.

WP5 Biological and medical interpretation of sequence data for rare variants
The CRG is participating in the WP5 subproject on chronic inflammatory disorders, together with CAU and LUMC. The CRG will mainly contribute to the data analysis for this subproject. We are also actively involved in the Parkinson's disease subproject together with HMGU. The CRG will sequence the DNA from 50 blood samples from unrelated cases, as well as RNA and DNA from 15 brain samples, and participate to the data analysis.

WP6 Ethical, legal and social issues (ELSI)
The CRG actively participated in the creation of the material used for the discussion on "Am I fine with having my genome sequenced and put in a database" monitored by Inserm. A discussion was organised in the CRG with both Roderic Guigo and Xavier Estivill and their lab members, with a total of 30 participants. The discussion was recorded and a thorough feedback was provided to Inserm together with 20 filled questionnaires. Moreover, this discussion was run in a familial context. This same discussion was also chaired by the GEUVADIS project manager during the European Summer School of health law and bioethics organised in Toulouse in July 2011. The CRG is also involved in the analysis and dissemination of the results from this study.

WP7 Dissemination and training
The CRG has been actively involved in this WP, suggesting workshop ideas and fostering the involvement of the GEUVADIS consortium in events organised by the partners.
- Roderic Guigó gave a talk on "preclinical and clinical application of RNA-seq" during the ESHG satellite meeting co-organised by GEUVADIS and ENGAGE.
- Gabrielle Bertier presented GEUVADIS, and organised an experts' discussion on 'Am I fine with having my genome sequenced and put in a database" during the Toulouse European Summer School organised by Inserm.
- Esther Lizano participated to the course on RNA-seq data analysis organised at the EBI in Hinxton. Several R-packages were presented in the course for quality of sequencing data assessment, alignment to a reference genome and differential expression analysis. Since some of these R-packages were presented by their developers, the discussions around them was very valuable. More details on these events can be found in Partner 11's report. (Leader of WP7)

<u>Presentations:</u>
The Coordinator and his team participated to the following congresses where Geuvadis was acknowledged (in a presentation and/or a poster):

- “Towards achieving individual medical genomic data", Barcelona, Spain, 04-nov-11
- 12th ICHG Montreal- The American Society of Human Genetics 61st Annual Meeting, Montreal, Canada, 11-oct-11
- 3rd meeting of the IRDiRC (International Rare Diseases Research Consortium) Montreal, Canada, 08-oct-11
- Epigenomics of Common Diseases conference Hinxton, UK, 13-sep-11
- The Genomics of Common Diseases 2011, Hinxton, UK, 30-ago-11
- European Human Genetics Conference 2011, Amsterdam, The Netherlands, 28-may-11
- The Biology of Genomes,Cold Spring Harbor Laboratory, US, 10-may-11
- OCFGC Meeting, Vancouver, Canada, 16-apr-11

- The NVHG Spring meeting, Veldhoven, The Netherlands, 31-mar-11
- Genomic Disorders 2011 - The Genomics of Rare Diseases, Cambridge, Hinxton, 23-mar-11
- Alzheimer's and Parkinson's Diseases: Advances, Concepts and New Challenges The 10th International Conference AD/PD 2011, Barcelona, Spain, 09-mar-11

d) Main tasks planned for the next year

- **WP1**: Complete all deliverables and reports planned in Annex III.
- **WP4**: CRG will prepare 96 small RNA libraries as well as collaborating with the CNAG to prepare mRNA-seq libraries with the same samples. Regarding data analysis, the CRG will participate actively in the evaluation of the sequencing data quality and analysis, including RNA expression quantification, differential expression analysis and correlations between mRNA and microRNA data.
- **WP5**: Analyse Geuvadis produced data from chronic inflammatory disorders. Perform exome sequencing of DNA and RNA from blood and brain samples of Parkinson patients, and initiate data analysis.
- **WP6**: Participate to the dissemination of the results of the study on "Am I fine with having my genome...".
- **WP7**: Achieve a dissemination plan for 2012 including training activities and expert's workshop, together with the Consortium members.

Partner 2, University of Geneva, leader of WP4

a) Summary of tasks performed.

WP4 Handling, analysis and interpretation of RNA-sequence data (WPL)
In the past year we have coordinated the purchase, growth and RNA extraction of 500 LCLs from the 1000 genomes project samples (400 from Europe and 100 form Africa). These are now distributed to the various labs for mRNAseq and miRNAseq and we expect all sequencing to be done by end of February 2012.

b) Main tasks planned for the next year.

Finishing of sequencing and statistical analysis of RNAseq and miRNAseq data for the 500 individuals above and submission of main analysis paper for this project

Partner 3, Institute of Human Genetics, Helmholtz Zentrum München, HMGU, Leader of WP5

a) Summary of tasks performed.

During the first 12 months of project, the twelve sequencing centers involved have developed routines for large scale sequencing projects for rare variants (objective #1). Large scale sequencing projects at the centers can be divided into three categories, i) whole genome, ii) whole exome and iii) targeted exon sequences. Analysis pipelines have been established to investigate sequence datasets for rare variants. The selection of pilot samples from existing collections has been initiated (objective #2). Phenotypes chosen include those originally planned (mitochondrial disorders (n=50), cardiovascular malformations (n=20), mental retardation (n=300) and cases with extreme glucose levels). In addition samples from patients with neurological disorders have been selected.
We decided to launch three sub-projects in the context of this WP:
We are coordinating two sequencing projects, and will perform a joint analysis of the sequences produced. 1) Parkinson's disease cases (n=450), and 2) Chronic Inflammatory disorders – Crohn's disease (CD), ulcerative colitis (UC), rheumatoid arthritis (RA) and celiac disease (CEL) - (n=50). Selection criteria included availability of samples and involvement of several sequencing centers. The coordinator maintains a project observatory which is available to all participating centers and which keeps track of the selection procedure (D5.1. Database for pilot project).

1) Parkinson's disease

Deviating from the original proposal, most of these medical samples will not have both DNA and RNA sequences available for analysis. For a subset of samples of Parkinson's disease, RNA and DNA will be sequenced from brain samples. This deviation from the original plan is well founded. First, it seemed more appropriate to invest all the resources available for RNA sequencing to general population samples (HapMap samples, WP4) where integration with whole-exome and whole-genome sequencing data is possible. Second, the selection process for a joint GEUVADIS sequencing project led us to Parkinson's disease, a disorder where RNA samples are only available from brains obtained from autopsy cases. Exome sequencing of Parkinson's samples has started already. We aim to solve sequence production and data exchange issues during the next 12 months in order to make the pilot sample accessible through a central database by month 24 (D5.2. Exome sequences of pilot samples accessible through central database).

2) Chronic Inflammatory disorders

Inflammation is a key biological mechanism for maintaining the immunological integrity of higher organisms. While phylogenetically designed to protect the body from infections, uncontrolled chronic inflammatory reactions may also be detrimental and represent the underlying cause of numerous socio-economically relevant disorders (e.g. rheumatoid arthritis, inflammatory bowel disease and celiac disease). In this sub-project, thanks to the unique access of the participants to disease samples, we will use Crohn's disease (CD), ulcerative colitis (UC), rheumatoid arthritis (RA) and celiac disease (CEL) as prototype traits for deciphering complex architectural changes in the transcriptome in chronic inflammation.

The data set will serve as an entry point to a functional understanding of the genetic risk maps put forward in each of the diseases over the last 3 years. It may help to translate the shared and unique risk genes into common and specific cellular response programs across the chronic inflammatory disorders and to identify biomarkers for diagnostics and therapeutic guidance.

3) Mental retardation

Taking advantage of their extensive experience in exome sequencing for mental retardation, and its use in the clinic, the Nijmegen group leads a subproject on this subject. It will mainly consist in sharing expertise and producing guidelines on the production, analysis and use of exome-seq data in both the research and the clinical settings.

The schedule for a review of sequence analysis tools available stays unchanged (D5.3. Provide review of sequence analysis tools available).

b) Summary of tasks performed.

WP2 Quality control of sequence data

To date we have produced in the order of 160 exomes for different paroxysmal neurological diseases. This has required working with the sequencing production pipelines, looking at the quality control of the sequence to define the quality control standards required for exome sequence with a view to detecting rare variants. Specifically we have been looking at variation in sequence quality and how this affects variation in the number of variants called. In addition we have been working on improved variant calling methods and developing workflows for using different calling algorithms including some methods developed in-house. We have validated variants using Sanger sequencing to determine error rates for the different calling methods.

c) Main tasks planned for the next year.

- To increase the number of exomes sequenced.
- Continued development of variant calling with improved filters for false positives looking at gene annotation and variant frequencies in different populations and available control sets.
- Continued emphasis on assessment and development of methods for indel calling

Partner 5, National Genotyping Center, CNG

a) Summary of tasks performed.

WP2 Quality control of sequence data
Participation in survey on quality control of exome sequencing.

WP5 Biological and medical interpretation of sequence data for rare variants
- Pilot clinical sequencing to identify mutations in known genes for maturitry onset diabetes of the young (MODY) on the Roche 454 Junior NGS platform.
- Together with the other partners we have selected 50 individuals with Parkinson's disease for exome sequencing at CNG.

b) Main tasks planned for the next year.

WP2: Implementation of best practice procedures for exome sequencing (Task 2.2)

WP5:
- Pilot exome sequencing in Parkinson’s disease.
- Merging and joint analysis of exome sequences from the participating centers. Formulation of standard operating procedures for clinical exome sequencing.

Partner 6, Max Planck Institute for Molecular Genetics, MPIMG

a) Summary of tasks performed.

The MPIMG is together with partner 2 (UNIGE) lead participant of WP4: Handling, Analysis and Interpretation of RNA-Sequence data and other functional datasets. The MPIMG participates in WP2 Quality Control of Sequence Data, WP3 Data Exchange, Archiving and Access, and WP7 Dissemination and Training.

The main tasks were the completion of the consortial RNAseq experiment addressing task 2.1 comparison of already established procedures for quality assessment of the sequence data among the participating laboratories with respect to transcriptome sequencing using NGS platforms available within the consortium, and task 2.2 defining best-practice procedures for quality control of transcriptome sequence data from the existing sequencing platforms, task 3.4 providing a report on standard processing pipeline formats and analysis components, task 4.1 pilot to RNA sequencing in 500 pilot samples, task 4.2 Coordinated evaluation of analysis tools and recommendation for translation

In the course of WP2 and WP4 the MPIMG participated in the gEUVADIS pilot project 1 (RNAseq). Five CEU cell line samples, which were used in the 1000 Genomes project pilot phase were ordered from Coriell by partner UNIGE. After RNA isolation at UNIGE the samples were distributed to different partners. After initial QC, we have performed sequencing using two different protocols on the Illumina platform: 1/ using the Illumina RNAseq standard protocol, and 2/ using the protocol for strand-specific sequencing of complementary DNA (see Transcriptome analysis by strand-specific sequencing of complementary DNA. Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A. Nucleic Acids Res. 2009 Oct;37(18):e123). After primary data analysis at the MPIMG, data were transferred to partner UNIGE for a common data analysis of the QC exercise. The MPIMG participated in the data analysis group.

Together with partners UNIGE and CAU, the MPIMG is performing a technology comparison study in order to better understand limitations and challenges in strandedness RNA seq. The data have been generated using the 5 CEU samples and will be correlated with GRO (global run on) seq and ChIPSeq data aiming at a full resolution of determinants of directionality of transcription and maturation kinetics of distinct RNA species.

Presentation
Ralf Sudbrak***: A map of human genome variation from population scale sequencing*** presentation at the National Genome Research Network (NGFN) Annual Meeting, Berlin NOV 27, 2010. Subject: 1000 Genome Project and GEUVADIS

b) Main tasks planned for the next year.

Task 4.1 participating in RNA sequencing in 500 pilot samples.

Partner 7, National Center for Genomic Analysis, CNAG

a) Summary of tasks performed.

In collaboration with partner P1 (CRG) we have been sequencing the exomes of 144 patients with Fibromyalgia. These exomes were sequenced at >40x coverage using 75 bases paired-end reads on the Illumina HiSeq2000 and Genome Analyzer IIx platforms (4 samples per lane). All sequences have been transferred to CRG for further downstream analysis and archiving. The sequencing data will allow identification of Single Nucleotide Variants (SNVs) and small insertions and deletions (indels) with respect to the reference sequence and to identify the relevant alterations that confer susceptibility to the disease.

CNAG is involved in the multicenter RNA-Seq experiment that aims to study gene expression in 465 RNA samples from HapMap3 and 1000GP projects. At CNAG we will perform RNA-Seq in 96 samples using 75-bp paired-end sequencing to at least 10 million mapped and properly paired-read pairs. We have participated in the experiment design, which has been led by partner P2 (UNIGE).

Deviation:
The main deviation from Annexe I pertains to the use of funds. Whereas the initial plan was to use the attributed funds for salaries and consumables, the increase in ambition of the project will lead us to a substantial overspend.

b) Main tasks planned for the next year (until November 2012).

- In collaboration with partner P1 (CRG), we will perform RNA-Seq with 96 RNA samples from HapMap3 and 1000GP projects.
- This work will include samples preparation and sequencing.
- We will also take charge of the sequencing of the small RNA libraries that will be prepared by P1 (CRG).

Partner 8, Uppsala University, UU, Leader of WP2

a) Summary of tasks performed:

WP2 Sequence data quality: The aim of WP2 is to establish and disseminate standards for quality controls (QC) for Next Generation Sequencing of RNA and Exomes. As a first step, UU has gathered input from the other sequencing partners and defined and listed important QC parameters currently used amongst the partners (D2.1). Based on this list, a survey form will later be created and circulated to the partners for Deliverable D2.2.

WP4 RNA sequencing: In the pilot study, UU sequenced and analysed 5 RNA samples provided by UNIGE. The data was made accessible to the consortium by uploading it to EBI's (P13) data server. P8 has participated actively in the preparations for the RNA sequencing project. UU is ready to sequence 48 RNA samples that will arrive by November.

WP6 Ethical, Legal and Social Issues: P8's main contribution to WP6 has been the organization of a local discussion with the topic *Ethical issues of whole genome sequencing*. A questionnaire provided by P14 (WP6 leader) was answered by a total of 22 colleagues and family members that participated in the discussion. The answers were sent to P14 together with a summary of the discussion.

Meetings: During the first year P8 has participated in the Kick-off meeting (Barcelona, Dec 2010, A-C Syvänen and Ulf Gyllensten), and A-C Syvänen and Mathias Bränvall) will attend the 1st Annual meeting (Toulouse, Nov 2011). UU has participated in a total of 10 teleconferences.

a) Main tasks planned for the next year.

- **WP2:** The survey on sequencing quality controls mentioned above will be circulated to the partners and a summary of the results will be included in D2.2. The QC parameters concerning the data analysis may also be included in WP3.
- **WP4**: We will use the protocols agreed upon to sequence small RNA and mRNA from a subgroup (n=48) of samples from the 1000 genome project. The results will be analysed and fed in to WP5 to be compared with the results from similar sequencing done by the other partners. The combined results will then be exploited to improve the appropriate QC- and data processing recommendations (WP2 and WP3).

Partner 9, Institute of Clinical Molecular Biology, CAU

a) Summary of tasks performed

CAU has participated in the consortial RNAseq efforts (gEUVADIS pilot project 1) in order to perform a standardized experiment of the 1000 genomes cell lines across all centres. RNA samples were distributed by UNIGE to all partners. After initial QC, we have performed sequencing using two different protocols on the SoliD platform (ds cDNA custom/ rRNA depleted WTAK) (see also Klostermeier et al., BMC Genomics 2011) in the five distributed CEU cell line samples from the 1000 Genomes project pilot phase. Mapped and raw data were both analyzed locally and with the partner University of Geneva in in order to generate proof-of-concept that such sequencing endeavours can be performed in a decentralized way. Together with MPIMG and Geneva, CAU is performing a technology comparison study in order to better understand limitations and challenges in strandedness RNA seq. The data have been generated using the 5 CEU samples and will be correlated with GRO (global run on) seq and ChIPSeq data aiming at a full resolution of determinants of directionality of transcription and maturation kinetics of distinct RNA species.

b) Main tasks planned for the next year.

The main tasks are the completion of the consortial RNAseq experiment and the organization of at least one other collaborative experiment derived from the disease indication groups. Here, a first principal outline for a disease-related experiment in different disorders of the intestine has been agreed upon between three gEUVADIS partners (CRG/LUMC/CAU).

Partner 10, Radboud University Nijmegen Medical Centre, RUNMC

a) Summary of tasks performed.

RUNMC is mostly involved in **WP5**, pilot sequencing project for quality control and disease mutation discovery. Within this WP we have obtained extensive experience in exome sequencing for Mendelian diseases associated with intellectual disability in the last 12 months, with a specific focus on the detection of rare *de novo* mutations in sporadic patients by sequencing patient-parent trios. Within the context of GEUVADIS we have performed the following actions:

- A detailed comparison of the performance of different exome enrichment assays from the following companies: Agilent, NimbleGen and Life Technologies (WP5).
- Optimization of *de novo* variant calling. This is not trivial as the focus on de novo mutations enriches for sequencing artefacts. For this we have developed a bioinformatic tool, DeNovoCheck, that performs its prediction based on data available in the parental BAM files (WP5).
- Established the informed consent procedure for diagnostic exome sequencing (WP6).
- Directed a novel course under the umbrella of the European School of Genetic Medicine entitled “Next generation sequencing for rare and common genetic disorders”, Bologna, Italy. For this course we invited both prof.dr. M. Dermitzakis as well as prof.dr. A. Cambon-Thomsen from the GEUVADIS consortium (WP7).

Presentations:
Invited Presentations on the use of next generation sequencing for Mendelian disease gene identification, in particular for the detection of de novo mutations, in which GEUVADIS has been acknowledged:

- Invited research seminar, Baylor college of Medicine, Houston, USA (2010).
- Swiss Society of Medical Genetics, Geneva, Switzerland (2011).
- Sao Paulo School of Advanced Science, Sao Paulo, Brasil (2011).
- Lausanne Genomics Day, Lausanne, Switzerland (2011).
- Invited research seminar, Exeter, England (2011).
- Joint Eurogentest & TECHGENE meeting, Leuven, Belgium (2011).
- 11th int. symposium on Mutations in the Genome, Santorini, Greece (2011).
- Workshop on Complex Genomics, Tartu, Estonia (2011).
- Second annual “Beyond the genome” conference, Washington, USA (2011).
- Fourth annual “Personal Genomes” meeting, Cold Spring Harbor, USA (2011).
- Symposium “Next Generation Sequencing in a clinical setting”, Uppsala University, Sweden (2011).
- International Congress of Human Genetics, Montreal, Canada (2011).

In addition, partner 10 directed a novel course for the European School of Genetic Medicine entitled “Next generation sequencing for rare and common genetic disorders”, in Bologna, Italy (2011). Partner 10 presented the Introduction of this course on “next generation sequencing basics” as well as a lecture on “whole genome *vs.* targeted next generation sequencing in the clinic”. In addition, partner 10 organized the hands-on data-analysis workshops in which participants learned how

to identify and interpret genomic variation in targeted as well as whole exome sequencing datasets. GEUVADIS is widely acknowledged during this course.

b) Main tasks planned for the next year.

- Development of guidelines for diagnostic exome sequencing, in close collaboration with GEUVADIS members (WP5).
- Report on the validation rate of de novo mutations called from exome sequencing data (WP5).
- Report on the diagnostic application of exome sequencing for genetically heterogeneous disorders (WP5).
- Evaluation of the informed consent procedure for diagnostic exome sequencing (WP6).
- Promote the discussion on clinical applications of next generation sequencing at meetings and in the media (WP7).

Partner 11, Leiden University Medical Center, LUMC, leader of WP7

a) Summary of tasks performed

WP4 RNA sequencing

LUMC was involved in the sequencing and data analysis of the pilot RNA-seq project. We compared different analysis pipelines, implemented some pipelines in the user-friendly workflow system Galaxy, and generated a new analysis pipeline dubbed PASSION (Zhang et al., submitted). PASSION uses a pattern-growth algorithm for improved splice junction detection. Shared developed workflows for transcriptomics (DeepSAGE) analysis via the Galaxy work flow system: http://galaxy.nbic.nl/

WP7 Dissemination (WP leader)

List of events sponsored/supported by Geuvadis:

	Event1	Event 2	Event3	Event4
Partner	RUNMC	LUMC, UNIGE	Inserm	EBI
Organiser	J. Veltman	P. Hoen S. Montgomery T. Giger	A.M Duguet A. Cambon Thomsen	A. Brazma
Date	14-17 April 2011	28 May 2011	7-8 July 2011	12-13 Sept. 2011
Venue	Bologna	Amsterdam	Toulouse	Hinxton
Type	Course	ESHG satellite meeting	European Summer School	Engage/Geuvadis Course
General Topic	Next Generation Sequencing for rare and common genetic disorders	RNA-Seq: from sample preparation to data analysis	Health law and bioethics Academic session	RNA-seq data analysis
Targeted Public	Students	PhD students, Preclinical and Clinical Researchers	PhD students, MDs, Researchers…	PhD/Post docs
Speakers	G. Romeo M. Hurles R. Casadio P. Robinson G. Matthijs R. Hennekam J. Lunshof	S.Montgomery A. Gonçalves	M. Breuning V. Nigro H. Yntema E. Rial Sebbag N. Haoulia (France) M. H. Zawati	A Goncalves M. Gonzales Porta T.J. Hardcastle G.Rustici
Geuvadis Speakers	J. Veltman A. Cambon-Thomsen M. Dermitzakis	P. A.C. 't Hoen R.Guigo	A. cambon Thomsen A. Soulier G. Bertier X. Fernandez	P.A.C Hoen
Geuvadis participants *Indicative*	10/90	10/25	5/20	4/26
Geuvadis budget *Indicative*	-	5000	4000 - 5000	-

Oral Presentations:

- P.A.C. 't Hoen: NGS Benelux meeting (Utrecht, June 2011): ***Statistical methods for analysis of RNA-Seq data***
- P.A.C. 't Hoen: GEUVADIS/ENGAGE satellite meeting (Amsterdam, May 2011): ***Introduction to RNA-seq technology***
- P.A.C. 't Hoen: INSERM (Nantes, April 2011): ***Statistical methods for analysis of RNA-Seq data***
- P.A.C. 't Hoen: Dutch Society for Human Genetics (Veldhoven, March 2011): ***Next generation sequencing-based mRNA profiling of total blood in a large human cohort***

Poster presentations:

P.A.C. 't Hoen: HitSeq / ISMB (Vienna, July 2011***): Multifactorial analysis of digital gene expression data from a large human cohort***

Courses:

P.A.C. 't Hoen: ENGAGE / GEUVADIS course on RNA-seq analysis (Hinxton, September 2011): ***Introduction to RNA-seq technology***

P.A.C. 't Hoen: First LUMC/NBIC course on advanced RNA-seq data analysis (Amsterdam, August 2011) (coordinator)

b) Main tasks planned for the next year.

- *Organization of further dissemination activities, including at least two RNA-seq and exome sequencing data analysis courses.
- *RNA-sequencing of specific subpopulations of T-lymphocytes from the gastrointestinal tract and other body locations
- *Participation in RNA-sequencing of HapMap cell lines (coordinated by UNIGE)

Partner 12, University of Santiago de Compostela, USC

a) Summary of tasks performed

WP2 T2.1:
During this period we worked on the establishment and evaluation of quality controls for exome sequencing. We established three types of quality controls, two types during sample processing and the third one during data analysis. During sample processing there are some quality controls that allow us to detect factors that can introduce some bias on final data (on library preparation: fragmentation of DNA, amplification, etc; enrichment: uniformity, coverage, etc) and other ones to check the performance of some steps of the protocol in order to avoid sequencing samples that not meet the minimum criteria just with the purpose of saving money and time. Regarding data analysis we evaluated the performance of some variant calling algorithm comparing exome sequencing data with Sanger sequencing data and genotyping data. Some errors and bugs have been detected and we are now trying to define the best parameters for reducing errors to the minimum. All this efforts have been done with SOLiD 4 platform and needs now to be compared with other platforms.

Presentation:
Angel Carracedo, ***Del GWAS al Deep-sequencing***. Valencia 29 Abril 2011.

b) Main tasks planned for the next year.

- On the next year, we plan to continue working on quality controls for exome sequencing together with partners involved in WP2.
- We upgraded the SOLiD 4 to 5500xl, so we need to adapt the quality controls to the new platform and re-evaluate them.
- New enrichment methods have been launched on the lasts month. We plan to compare them with current methods on December 2012.
- Regarding sample data available for the project, on one hand, in the coming months we plan to increase the number of samples and pathologies for exome sequencing. It will be very interesting to discuss the possibility of including them in the exome sample data set of WP5.
- On the other hand, on the following weeks we will decide if we go for whole genome sequencing of 50 samples of colorectal carcinoma. If we go ahead with this option, we could explore the convenience of including them in the project.

Partner 13, European Bioinformatics Institute, EBI, leader WP3

a) Summary of tasks performed.

The procedures of handling high-throughput sequencing experiments in EBI archives have been established.

High-throughput sequencing data is handled at EBI in three major repositories: ArrayExpress, European Nucleotide Archive (ENA) and European Genome-phenome Archive (EGA). The ENA accepts data generated by high-throughput sequencing platforms such as Roche 454, Illumina Genome Analyzer and ABI SOLiD. ENA works in close collaboration with the NCBI and DDBJ as part of the International Nucleotide Sequence Database Collaboration (INSDC). All submitted public data is exchanged between the partners on a daily basis. ENA stores sequence read data and technical meta-data describing how a sequencing experiment has been processed, for example base coordinates, library layout (single or paired) and spot length. Raw data can be submitted in different formats, but all data are converted into fastq files in the archive. As a result, users can access both originally submitted raw data and processed raw data in fastq format.

ArrayExpress, on the other hand, is a database of functional genomics experiments including those based on high-throughput sequencing technology. It stores experiment meta-data including description of samples, variable factors, and experimental design, as well as processed data such as mapping files in bam format, and references to the raw data stored in the ENA. If a submission contains human identifiable data, it will go through the European Genome-phenome archive, which provides secure storage and works with data access committees for controlled access when that is required. Data generated in Geuvadis project such as protocols comparison in different labs can be easily submitted into EBI archives. At the moment EBI provides two possibilities for high-throughput sequencing data submitters:

1. Submit data into ArrayExpress in MAGE-TAB format. In such a case the raw data will be passed on to ENA, while metad-data and processed data will be stored in ArrayExpress.
2. Submit data into ENA, get accession numbers for study and submit meta-data together with processed data into ArrayExpress including ENA accessions into MAGE-TAB file.

Creation of MINSEQE standard

The described data flow where ArrayExpress stores meta-data and processed data and ENA stores the raw sequence read data fully supports the MINSEQE (Minimum information about a high-throughput SeQuencing Experiment) standard. MINSEQE describes guidelines for the minimum information that should be reported about high-throughput sequencing experiments to enable unambiguous interpretation (see http://www.mged.org/minseqe/).

Ontology development for annotating highthroughput sequencing experiments

The Experimental Factor Ontology (EFO) was initially developed in order to annotate experimental data in ArrayExpress and in other databases at EBI. The methodology used when developing the EFO involves transforming terms in an experimental variable vocabulary into an ontological representation, and forming classes and

relationships between those classes. In close collaboration with EBI European Nucleotide Archive (ENA) a number of terms specific for high-throughput sequencing experiments have been added to EFO. New classes and groupings have been created, and synonyms, external references, axioms and definitions have been added or revised in order to make EFO terms easily accessible. EFO is used for data annotation and submission in the EBI ArrayExpress and ENA databases.

Presentations:

- Angela Gonçalves, ***Challenges in RNA-seq data analysis,*** GEUVADIS and ENGAGE ESHG satellite meeting: Introduction to RNA-seq technology. Sunday, May 29 2011.
- Xose Fernandez***, Genomics: A View from Hinxton,*** European summer school on medical law and bioethics academic session. Special workshop on Bioinformatics and Sequencing platforms of use in Europe for genetic research and diagnostics. July 8 2011 Toulouse.
- 8th Georgia Tech-Emory International Conference on Bioinformatics, November 10 - 12, 2011, Ferst Center for the Arts, Georgia Tech.

Training:

ENGAGE/GEUVADIS RNA-seq Course has been organised at EBI 12th - 13th September 2011. The course covered data analysis of RNA-Seq experiments. 30 people attended. 4 registered from the GEUVADIS project. Feedback on the event was collected from all participants and was generally very positive.
http://www.ebi.ac.uk/training/onsite/110912-ENGAGE_Workshop.html

b) Plans for the next year.

- Hosting of Geuvadis processed data and users support.
- Meta-data management.
- Evaluation of existing at EBI high-throughput sequencing data processing pipelines.
- Participation in Geuvadis data analysis.
- We are going to use EFO ontology widely for high-throughput sequencing data annotation and supplement the ontology if it is needed.

Partner 14, National Health and Medical Research Institute Inserm, leader of WP6

a) Summary of tasks performed.

WP6 Ethical, legal and social issues
Task 6.1: Ensure the ethics management is effective: discussions at kick off, survey of issues and update at each steering teleconference
Task 6.2: Establish a network of ELSI experts and initiatives : started
Task 6.3: Produce a synthetic document on specific ELSI aspects in relevant to phenotype prediction from sequence variation in various clinical situations. One chapter published; several presentations; decision to make a research survey within the project in addition to bibliography analysis:

Research survey*: "Am I fine with having my genome sequenced and put in a database?"* Undertaken with the collaboration of Geuvadis partners (and their families)
- Questionnaires: 67 respondents (14 nationalities)
- Groups of discussions: 5 (5 countries)

Task 6.4: Hold a 2 days expert workshop of about 30 people with key other ELSI initiatives and produce a consensus document on ELSI related policy guidelines for implementation of phenotype prediction from sequence variation, Planned for 2nd part of 2012; coordination with ESGI.
Task 6.5: To organize a repertoire of relevant ELSI events: started.
Task 6.6: To make a survey on the web about sequencing proposed directly to consumer, with bibliographic work on the subject and write a synthesis for publication: started; 1 presentation in Innsbruck symposium.
Task 6.7: To organize an expert meeting to draft a position paper and policy recommendations proposal on ELSI related policy guidelines regarding sequences proposed direct to consumer; 1st discussion planned in the ELSI session of Nov 29 Geuvadis meeting.
Task 6.8: to propose to WP7 relevant ELSI training modules incorporated to other training actions or as stand alone events: 3 actions in 2010-11: Teaching lecture in:
- The 4th Paris Workshop on Genomic Epidemiology May 30 – June 1, 2011 Maison de la Chimie, Paris.
- European School of Genetic Medicine: Course in Next Generation Sequencing for rare and common genetic disorders EuroMediterranean University Centre of Ronzano, Bologna, Italy - 14th -17th April 2011.
- European school on bioethics and medical law, Toulouse, 4-7 July 2011 (with Geuvadis funds participation).

Posters presentations:
- A. Cambon-Thomsen, A. Soulier, G. Bertier, S. Leonard, S. Julia, GEUVADIS consortium. ***Professional and family attitudes regarding large scale genetic information generated through next generation sequencing in research***. 12th International Congress of Human Genetics, Montreal, October 2011; 61st meeting of The American society of human genetics, Montreal, October 2011.
- S Julia, A Soulier, A. Cambon-Thomsen. ***Are genetic testing technologies driving clinical practice?*** ESHG Amsterdam, May 2011.

- E Rial-Sebbag, A Cambon-Thomsen. ***Biobanks: a step forward database infrastructure ethical and legal challenges in the context of next generation sequencing***. ICGC meeting, Kyoto, July 2011.

Oral presentations:

- A Soulier, G Bertier, A. Cambon-Thomsen ***"Am I fine with having my genome sequenced and put in a database. Investigation in a research environment"*** European summer school of medical law and bioethics, Toulouse, July 2011.
- S Leonard, A Soulier, A. Cambon-Thomsen, ***No Man is an Island – A look at the wider implications of DTC genome sequencing***. Interdisciplinary Symposium "Genetics as Culture in a Consumerist Age", Innsbruck, Austria, 27-29 Oct 2011.
- Cambon-Thomsen A (invited presentations).
- ***Public health and societal issues***. The 4th Paris Workshop on Genomic Epidemiology May 30 – June 1, 2011 Maison de la Chimie, Paris.
- ***Data driven research and large scale studies in biomedical research: what consequences for data sharing and bioethics in human genetics?*** 14th congress of logic, methodology and philosophy of science (clmps) July 19-26, 2011, Nancy.
- ***Ethical and legal frameworks applying for personal genome data.*** European School of Genetic Medicine: Course in Next Generation Sequencing for rare and common genetic disorders Euro-Mediterranean University Centre of Ronzano, Bologna, Italy - 14th -17th April 2011.
- ***Privacy in the context of high throughput technologies in genetics.*** UNESCO Tenth Meeting of the UN Interagency Committee on Bioethics. « Genetic privacy and non discrimination" Paris, 4/3/2011.

b) Main tasks planned for the next year.

- Analysis of the survey material collected and writing of articles based on this research.
- Poster submitted: ***"Professional and family attitudes regarding large scale genetic information generated through next generation sequencing in research".*** ESRC Genomics Network Conference: 'Genomics in Society: Facts, Fictions and Cultures' London, April 2012
- Organization of Geuvadis annual meeting in Toulouse on 28/29 November 2011. Organisation of ELSI workshop during the annual meeting –Geuvadis annual meeting.
- 2 abstracts submitted 'Professional and family attitudes regarding large scale genetic information generated through next generation sequencing in research'
- Questionnaire statistics and responses; Group discussion analysis and identification of themes.
- Abstract submitted on 'Revision of Data Protection Direction 95/46/EC'
- Perspectives to explore on ethical aspects of whole genome sequencing.
- Revision "Directive data protection".
- Ethical analysis of an innovative use of whole genome sequencing, when proposed to consumers on Internet.

Project publications

1. Accepted publications:

Rami Abou Jamra, Orianne Philippe, Annick Raas-Rothschild, Sebastian H. Eck, Elisabeth Graf, Rebecca Buchert, Guntram Borck, Arif Ekici, Felix F. Brockschmidt, Markus M. Nothen, Arnold Munnich, Tim M. Strom, Andre Reis, and Laurence Colleaux, ***Adaptor Protein Complex 4 Deficiency Causes Severe Autosomal-Recessive Intellectual Disability, Progressive Spastic Paraplegia, Shy Character, and Short Stature,*** The American Journal of Human Genetics doi:10.1016/j.ajhg.2011.04.019

Kapushesky M, Adamusiak T, Burdett T, Culhane A, Farne A, Filippov A, Holloway E, Klebanov A, Kryvych N, Kurbatova N, Kurnosov P, Malone J, Melnichuk O, Petryszak R, Pultsin N, Rustici G, Tikhonov A, Travillian RS, Williams E, Zorin A, Parkinson H, Brazma A. ***Gene Expression Atlas update--a value-added database of microarray and sequencing-based functional genomics experiments****,* Nucleic Acids Res., 2011, PMID: 22064864

Soulier A, Julia S. Cambon-thomsen A. ***A review of ethical questions raised by the transfer into clinics of massive parallel sequencing technologies.*** Chapter published: in «Droits des patients, mobilité et accès aux soins», Les Etudes Hospitalières. Collection «Séminaire d'actualité de droit médical», Bordeaux, 2011, 257-266

2. Submitted publications:

Yanju Zhang, Eric-Wubbo Lameijer, Peter A.C. 't Hoen, Zemin Ning, P. Eline Slagboom, and Kai Ye ***PASSion: A Pattern Growth Algorithm Based Pipeline for Splice Junction Detection in Paired-end RNA-Seq Data.*** Bioinformatics

Joep de Ligt, Lisenka E.L.M. Vissers, Christian Gilissen, Joris A. Veltman and Jayne Y. Hehir-Kwa. DeNovoCheck: ***Inheritance analysis for NGS trio data***.

Soulier A. and Cambon-Thomsen A. ***Transferring technologies, transferring values.***

Report Annexes:

- Annex 1: List of participants and institutions acronyms
- Annex 2: List of Work Packages and Work Package Leaders (WPL)
- Annex 3: Deliverable D1.4: Communication
- Annex 4: Deliverable D1.3: Website evaluation

Annex 1: List of participants and institutions acronyms

P1 **WP1 leader**	Centre for Genomic Regulation (CRG)	Xavier Estivill Roderic Guigo
P2 **WP5 leader**	Université de Genève (UNIGE)	Emmanouil Dermitzakis Stylianos Antonarakis Tuuli Lappalainen
P3 **WP4 leader**	Helmholtz Zentrum München (HMGU)	Thomas Meitinger Tim M Strom
P4	Wellcome Trust Sanger Institute (WTSI)	Aarno Palotie
P5	Centre National de la Recherche Génomique (CNG)	Mark Lathrop Jorg Hager
P6	Max-Planck-Gesellschaft zur Förderung der Wissenschaften (MPI-MG)	Hans Lehrach Ralf Sudbrak
P7	Centro Nacional Avanzado de Genómica (PCB-CNAG)	Ivo Gut Mónica Bayès
P8 **WP2 leader**	Uppsala University (UU)	Ann-Christine Syvanen Kerstin Lindblad-Toh Ulf Gyllensten Matthias Bränval
P9	Institute of Clinical Molecular Biology, Christian-Albrechts University of Kiel (CAU)	Stefan Schreiber Philip Rosenstiel
P10	Radboud University Nijmegen Medical Centre (RUN-MC)	Joris Veltman Han Brunner
P11 **WP7 leader**	Academisch Ziekenhuis Leiden - Leids Universitair Medisch Centrum (LUMC)	Gert J. van Ommen Peter t Hoen
P12	Universidad de Santiago de Compostela (USC)	Angel Carracedo Beatriz Sobrino
P13 **WP3 leader**	European Molecular Biology Organization / European Bioinformatics Institute (EMBL/EBI)	Alvis Brazma Paul Flicek Natalja Kurbatova
P14 WP6 leader	Institut National de la Santé et de la Recherche Médicale (INSERM)	Anne Cambon-Thomsen Jane Miller
P15	Applied Biosystems Deutschland GmbH (AB)	Jonathan Mangion
P16	Illumina Cambridge Limited (Illumina)	David Bentley
P17	Johns Hopkins University School of Medicine (JHU/OMIM)	Ada Hamosh

Annex 2: List of Work Packages and Work Package Leaders (WPL)

Title	**WPL**
WP1: Coordination	**P1**: CRG
WP2: Quality control of sequence Data	**P8**: UU
WP3: Data storage access and exchange	**P13**: EBI
WP4: Handling, analysis and interpretation of RNA-sequence data	**P2**: UNIGE
WP5: Biological and interpretation of sequence data for rare variants	**P3**: HMGU
WP6: Ethical, legal and social issues (ELSI)	**P14**: Inserm
WP7: Dissemination and training	**P11**: LUMC

Annex 3: Deliverable D1.4: Communication

DELIVERABLE REPORT

Deliverable D1.4
GEUVADIS Communication

Grant Agreement number: **HEALTH-2010-261123**

Project title: **Genetic European Variation in Disease**

Funding Scheme: **Coordination and Support Action**

Due date of deliverable: month 14

Actual submission date: 02.12.2010

Start date of project: 01.10.2010 **Duration: to 31.09.2013**

Organisation name of lead contractor for this deliverable CRG

Project co-funded by the European Commission within the 7th Framework Programme		
Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

1. OVERVIEW

During this first year of the project, we have put in place an efficient communication policy, within the consortium, as well as from the project to the outside world -from the research community to the general public. This communication strategy has been implemented through a number of tools:

- Public website (Cf D1.3)
- Intranet (Cf D1.3)
- Specific mailing lists
    - Geuvadis_partners list
    - Geuvadis_admin
    - Geuvadis_RNAseq
    - Geuvadis_exome
    - Geuvadis_wpl
- Regular Telephone Conferences (TCs)
- Meetings
- Newsletter
- Press Release

In the following sections, we'll give a more detailed description of the tools and how we implemented them. For more details on the website and intranet, please refer do D1.3.

2. MAILING LISTS

We created a total of 5 mailing lists to facilitate distribution and dissemination of relevant information throughout the project:

- **Geuvadis_wpl**: All Work Package leaders, composing the Management Committee (MC). This list is used to facilitate decision-making, to organise TCs with the MC, distribute and edit TC minutes, etc...
- **Geuvadis_partners**: All partners in the projects. This list is used to distribute all information relevant to our entire Consortium: Annual Meeting date and venue, Meetings and Workshops, important publications, decisions taken by the MC, etc...
- **Geuvadis_admin**: This list regroups all our administrative contacts in the different groups.
- **Geuvadis_RNAseq**
- **Geuvadis_exome**: These two lists are used to communicate information specific to WP4 (Handling, analysis and interpretation of RNA-Sequence data and other functional datasets) and WP5 (Biological and Medical interpretation of sequence data for rare variants)

A new mailing list for the analysis of RNAseq experiment data is under construction.

3. TELEPHONE CONFERENCES (TC)

Apart from email communication through the mailing lists and individual emails - which was indented to be as efficient and as low in volume as possible, we used TCs to communicate directly on specific aspects, from regular updates on project activities, to organisation of meetings and workshops, to important decision-making.

All TC agendas and minutes are available to project members in the intranet a few days after the TC. Work Package Leaders TCs were organised twice a month, and every time a specific decision had to be made we organised extra conferences with the relevant partners.

All TCs were organised by the Coordination team. All partners can suggest to organise a Tele Conference when they feel that they need a direct discussion within the consortium or with external partners to explore collaborations.

During this project first year, we organised a total of 11 TCs:

TC	08/10/2010	WPL
TC	27/10/2010	WP4 RNAseq Pilot study
TC	16/11/2010	WPL
TC	01/12/2010	WPL
TC	07/02/2011	WPL
TC	18/04/2011	WPL
TC	24/05/2011	WP5 Intestinal inflammatory disorders Subgroup
TC	21/07/2011	WPL
TC	13/09/2011	WP4 RNAseq Results
TC	03/10/2011	WPL
TC	03/11/2011	WP4 RNAseq study design+wiki

4. MEETINGS

Taking advantage of the regular national or international meetings where several Geuvadis partners were present, the community managed to meet several times this year, for instance:
- during the CRG 2010 Annual Symposium, which was held before the project kick-off meeting.
- during the European Society of Human Genetics 2011 Meeting in Amsterdam, where a project workshop was held
- during the 4th Paris Workshop on Genomic Epidemiology
- during the Leena Peeltonen School of Human Genomics
-...

5. NEWSLETTER

We wrote a first newsletter on April 2011, and we plan to create the following ones on a bi-annual basis. This Newsletter was elaborated by the coordination team, with the contribution of all relevant partners, and was approved by all partners before being distributed within the consortium.

The Newsletter is available in Annex 1.

6. PRESS RELEASE

The first project press release was elaborated for the project KO meeting.
A first draft was produced by the coordinator, and sent around for comments and edits to all partners and their communication departments. Once the final draft established, it was sent to the press, on the same day as the KO meeting.

The press release is available in Annex 2.

ANNEX 1 : FIRST GEUVADIS NEWSLETTER

GEUVADIS NEWSLETTER 1

April 2011
Course on Next Generation Sequencing, Bologna, Italy

In this issue…

KO Meeting in Barcelona

The Project officially started on October 2010, and the annual meeting was held in Barcelona on 17th December.

All partners discussed the plan for the next three years, and Work Package Leaders presented their ideas on the implementation of the tasks described in the grant.

3 Geuvadis Workshops in 6 months!

Since the beginning of the project, partners have been very active in organizing workshops and events linked with the project:
- In April 2011, Joris Veltman organised a Course in Next Generation Sequencing for rare and common genetic disorders in Bologna, Italy. Manolis Dermitziakis and Anne Cambon-Thomsen were invited speakers to this course. (see frontpage picture)
- Peter Hoen set up a Geuvdadis/Engage satellite meeting during the last meeting of the European Society of Human Genetics, to which Stephen Montgomery and Roderic Guigó participated.
- Third, Anne Cambon-Thomsen and Xose Fernández from EBI are organizing a hands-on Workshop on databases available at the EBI during the European Summer School of Health Law and Bioethics in July 2011. The results of the discussion[1] on 'Am I fine with having my genome sequenced and put in a database?' will be presented.

[1] See the Highlight section, p2.

RNA sequencing: lessons from the pilot study

Before the project started, RNA from 5 different Central European samples was sent to 7 participating labs for sequencing and analysis. The results of this first Geuvadis pilot study were presented by Thomas Giger, UNIGE, during the KO meeting, and by Stephen Montgomery during the ESHG Satellite meeting. All partners agreed that the poor comparability of the results that were obtained in the different labs was due in great part to the different protocols that had been used for library preparation, and sequencing.

Manolis Dermitziakis, Work Package Leader for RNA sequencing, designed together with his team and in collaboration with 12 Geuvadis partners, a larger experiment workflow: using the very same protocols, 12 labs will sequence a total of 500 randomly distributed samples from the HapMap Cell lines. Data will be stored in an EBI database, and analysed by Manolis' team.

Disease samples sequencing on track

During the KO meeting, partners decided to set up three working groups for WP5 on disease samples sequencing:

The subgroup on cardiovascular diseases is lead by Thomas Meitinger, the one on mental retardation by Joris Veltman, and finally the one on inflammatory disorders is lead by Stefan Schreiber and GertJan Van Ommen.

These working groups have organized specific teleconferences and started setting up a work plan.

Website, leaflet and poster

At the start of the project, the CRG management team created a dedicated website, which is regularly updated with news and events, and also includes a private section where all useful documents of the project are deposited. In the following months we will have to assess the use of the website, by distributing a survey questionnaire to all partners. All comments and suggestions are very welcome to make the website as useful as possible for communications and collaborations through the project.

We also created a leaflet for the project to inform potential new collaborators and communicate main goals and features of Geuvadis.

A poster was also created and presented at the ESHG meeting in Amsterdam. Note that all this material can be found on the website's intranet.

Press room: 7 articles just in 2011!

Since January 2011, the project has received a significant press coverage. Articles mentioning the KO meeting, the main objectives of the project and quoting our press release, and interviewing several partners were published in 6 different press journals and specialized websites. Links to these articles are available in our online press room page.

Highlight: "Am I fine with having my genome sequenced and put in a database?"

As suggested by Thomas Meitinger during the KO meeting, Anne Cambon-Thomsen, as WPL for Ethical Legal and Social Issues prepared a series of documents to help the partners to organize a discussion on the theme: "Am I fine with having my genome sequenced and put in a database?". This discussion was held in partners´ families and in their labs. Feedback from the discussions, as well as questionnaires filled in by participants, were gathered, and the preliminary results of these discussions will be presented during the European Summer School in Toulouse in July 2011.

New Faces: Mathias Bränvall, Uppsala

Please don't hesitate to let us know if new people in your team are involved in the GEUVADIS project, so we can add them to the mailing list and announce their involvement to all partners.

Mathias Bränvall, PhD
Project Coordinator
Molecular Medicine
Dept. of Medical Sciences
Uppsala University

After a PhD degree at Uppsala University, focusing on the role of metal ions in the ribozyme-substrate interaction, Mathias worked and studied in Leif Kirsebom's lab, investigating a catalytic RNA, bacterial RNase P RNA, and then worked for 5 years for Progenika -a Spanish biotech company involved in several EU-funded projects. He started in Ann-Christine Syvänen's group in April this year (2011), foremost hired to coordinate the EU-funded projects running in the group, including GEUVADIS. His role will be to keep an eye on the deadlines, track the progress of work, write WP and Deliverable reports and make sure UU fulfill its part of the project.

" As I have been involved in technical product validation, I find the coming comparison of results from different labs, using the same protocols and the same/similar samples will be very interesting. I think project will contribute to and also improve the quality and generation of sequencing data for the benefit to everybody working with NGS."

Links and Announcements

- **Annual Meeting 2011**

Our next annual meeting will take place in Toulouse on November 28 and 29, 2011. Anne Cambon-Thomsen and the management team will organise the meeting to which the EU officer and our SAB have been invited. *All suggestions and ideas for a potential one day workshop are welcome.*

- Link from Ada Hamosh:

"Please make sure that all the members of the Geuvadis consortium know that there is a new website for OMIM: www.omim.org"

- Wellcome Trust Advanced courses: Next Generation Sequencing.

Supported by the European Sequencing and Genotyping Infrastructure (FP7 Infrastructure project). 2-10 October 2011; Wellcome Trust Genome Campus, Hinxton, Cambridge; Deadline for applications: 24 June 2011

ANNEX 2 : FIRST GEUVADIS PRESS RELEASE

PRESS RELEASE

DECIPHERING HUMAN GENOME SEQUENCES

TOP EUROPEAN SCIENTISTS ESTABLISH A NEW BASIS TO STANDARDIZE ANALYSIS AND PROMOTE GOOD PRACTICE IN HUMAN GENOME SEQUENCING.

The GEUVADIS (Genetic European Variation in Disease) project brings together Europe's leading medical genome sequencing laboratories to define technological and ethical standards, and to promote multidisciplinary training for the global scientific and medical community.
GEUVADIS, coordinated by Dr Xavier Estivill of the Centre for Genomic Regulation in Barcelona, Spain and supported by the European Commission, includes 17 international partners throughout Europe and the United States.

Our genes determine how our body functions – or malfunctions. Understanding genes will help us to identify the best way to treat people for disease, and to discover new medicines. But obtaining, reading and interpreting the 3,000 million-letter code of the human genome has been a difficult journey. First, researchers pointed out millions of variable "spots" in our genome, and discovered their association with disease. Today, advanced generation sequencing technologies have revolutionised the field by allowing us to read in detail and analyse the complete sequence of individual genomes, in an ever-faster and cheaper manner.
A growing number of research projects have flourished in response to the increasingly rapid evolution of these technologies, which has led to an unprecedented surge in new biological data. There are now several large-scale sequencing projects like the **1000 Genomes project** and the **International Cancer Genomics Consortium** that are analysing thousands of samples from different populations and disease status. GEUVADIS investigators are partners of these large-scales projects. The production of this large amount of data poses major challenges that the GEUVADIS consortium is going to tackle in Europe.
The amount of data produced is large, but its quality and accuracy needs to be thoroughly determined. The correct interpretation of the data, and its use in diagnosis and treatment is hence complex. The release of all this data to the research community, and – in a global format - to the public, raises the scale of ethical, legal and social reflection: How can we ensure that the privacy of individual patients is protected as sequencing becomes cheaper and more available? How does the availability of this technology affect freedom of choice and mutual respect?
GEUVADIS takes a coherent and tightly coordinated approach to addressing these and other important questions. Dr Estivill, coordinator of the **Genes and Disease Programme** at the Centre for Genomic Regulation in Barcelona and coordinator of the project, envisions GEUVADIS as "a strong European framework that can be used to dissect the genetics of disease and to implement genomics responsibly in the medical setting". In the framework of Spain, there are ongoing discussions to sequence the genomes of thousands of subjects affected by the most common diseases. This will boost basic and pharmaceutical research in a short time frame.
Anne Cambon-Thomsen of Toulouse (France), in charge of the ethical, legal and social aspects of the project, adds "To fill the gap between the technological fascination and speed, and the responsible implementation of genome sequencing, it is crucial that scientists participate early in the analysis of the ethical and societal dimensions of their work. This dimension is a lively axis of the work in GEUVADIS".

Notes for editors:
The GEUVADIS Project, (Genetic European Variation in Disease) is a Coordination Action supported by the European Commission under the 7th framework programme, and

coordinated by Xavier Estivill at the Center for Genomic Regulation in Barcelona.
Geuvadis participants: Top sequencing centers in Europe

- Centre for Genomic Regulation, Spain
- University of Geneva, Switzerland
- Helmholtz Zentrum München – German Research Center for Environmental HealthHealth, Germany
- Wellcome Trust Sanger Institute, United Kingdom
- National Genotyping Center, France
- Max-Planck Institute of Human Genetics, Germany
- National Centre for Genomic Analysis, Spain
- Uppsala University, Sweden
- Christian-Albrechts University of Kiel, Germany
- Radboud University Nijmegen Medical Centre, The Netherlands
- Leiden University Medical Center, The Netherlands
- University of Santiago de Compostela, Spain
- European Bioinformatics Institute, United Kingdom
- National Institute for Health and Medical Research, France
- Applied Biosystems Deutschland GmbH, Germany
- Illumina Cambridge Limited, United Kingdom
- Johns Hopkins University School of Medicine, United States of America

For further information: **Laia Cendrós**, Communication & PPRR Dept. Centre for Genomic Regulation (CRG), Dr. Aiguader, 88 – Edif. PRBB, 08003 Barcelona. Tel. +3493 3160237 .

Annex 4: Deliverable D1.3: Website evaluation

DELIVERABLE REPORT

Deliverable D1.3
GEUVADIS Website Evaluation

Grant Agreement number: **HEALTH-2010-261123**

Project title: **Genetic European Variation in Disease**

Funding Scheme: **Coordination and Support Action**

Due date of deliverable: month 14

Actual submission date: 02.12.2010

Start date of project: 01.10.2010 **Duration: to 31.09.2013**

Organisation name of lead contractor for this deliverable CRG

Project co-funded by the European Commission within the 7th Framework Programme		
Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

1. INTRODUCTION

The project's website is publically accessible since October 2010 at www.geuvadis.eu.
It provides an overview on the project, its origin, goals, members and ambition; as well as general public targeted resources, such as a podcast on the project, and basic information on genomic medicine, sequencing and Ethical, Legal and Social (ELSI) issues. This core information is fixed and remains unchanged.

On the other hand, we perform regular updates, and aliment the website, mainly with:

- News: Important publications, events or updates on our activities.
- Events we organize. (Training, meetings or workshops)
- Events related to the project (such as the European society of Human Genetics meeting…)
- Publications: all scientific publications liked to the project, and produced through the project are regularly updated and acknowledged. We also indicate all general press articles linked with GEUVADIS on the 'Press room' page.

A "Private" intranet section is also accessible to all consortium members through an individual username/password. For more details see section 2.

To evaluate its use and usefulness for project members, we conducted a short survey. As summary of responses to this survey is provided in the **Annex** of this Deliverable Report.

To improve the content of the website, notably on the general presentation and language, we also consulted members of Partner Institutions' Communication Departments.

Since its original publication in October 2010, we have significantly completed and improved the content of our website:

Original Site Map

- HOME
- Project
  - Background
  - Aims
  - Management
  - Work plan
- Partners
  - Map
  - Community
  - Scientific Advisory Board
- News & Events
  - News
  - Events
- Press room
- Publications
- Facts & Figures
  - Sequencing our Genome
  - Genomic Medicine
- Related Projects
  - European Projects
  - International Initiatives
  - Infrastructures
- PRIVATE

Updated Site Map (Additional sections highlighted)

- HOME
- Project
  - PODCAST
  - Background
  - Aims
  - Management
  - Workplan
- Partners
  - Map
  - Community
  - Scientific Advisory Board
- News & Events
  - News
  - Events
- Publications
  - Scientific Publications
  - Press room
- Resources
  - Ethical Social Legal Issues
  - Sequencing our Genome
  - Genomic Medicine
- Related Projects
  - European Projects
  - International Initiatives
  - Infrastructures
- PRIVATE

2. GEUVADIS INTRANET

The GEUVADIS intranet (cf screenshot below) gathers all information useful to partners in the course of the project. All official documents, TeleConferences and meeting minutes, or documents linked to specific workpackages (results of pilot studies) and discussions on the project are accessible to each partners thanks to a unique username and password. A Wiki page specific to WP4 RNaseq has been created by the University of Geneva team (WP4 leader).

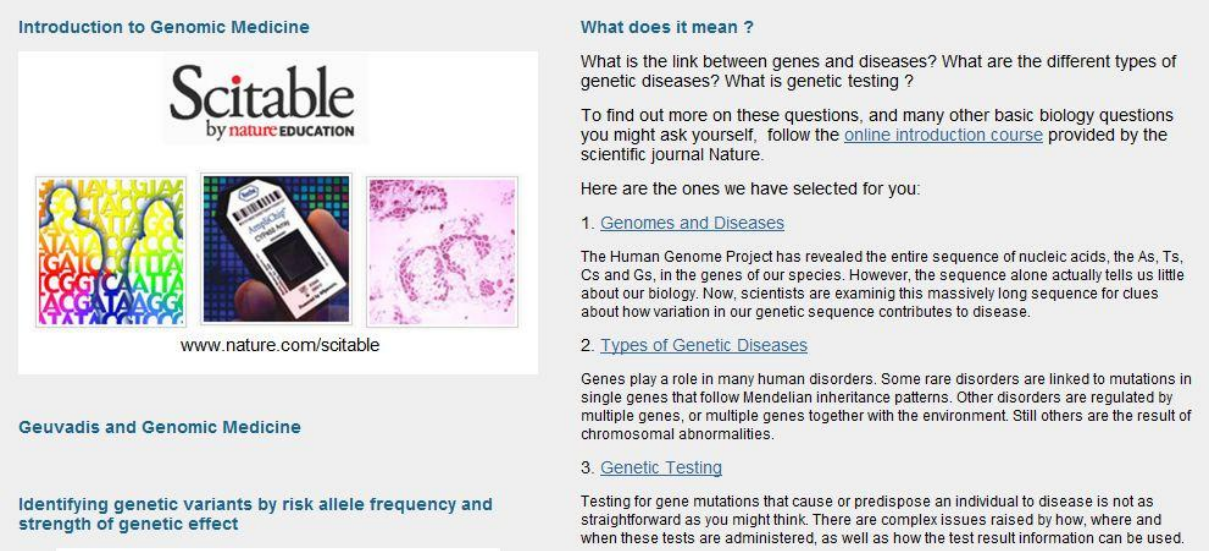
Structure of the intranet:

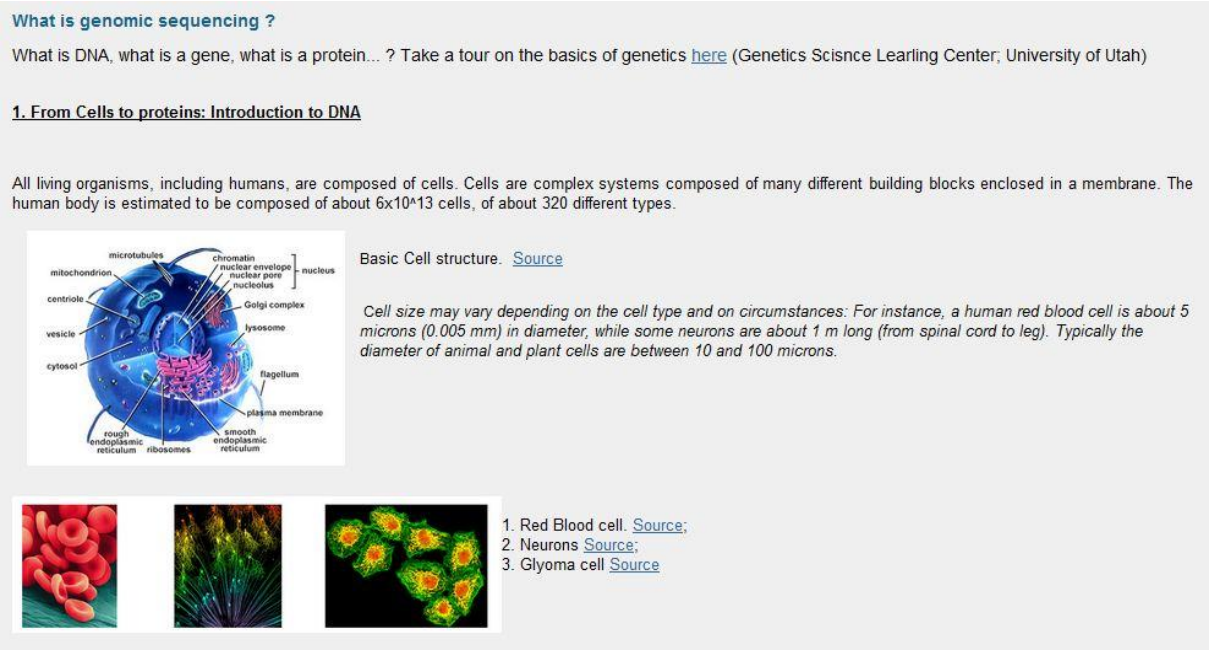
- **Documents**
  - Budget GEUVADIS
  - Deliverables
  - KO meeting 17.12.2010
  - Logo Flyer Poster
  - News
  - Official Documentation
  - Periodic Reports
  - Podcast
  - Presentation Documents
  - Press Releases Newsletters
  - Tele Conferences
  - Work in progress
  - WP4 RNAseq
  - WP6 ELSI issues
  - Subfolders: Discussion and Questionnaire on Genomic Sequencing
  - WP7 Dissemination Training
- **Debates and discussions**
  - HapMap Pilot study
  - Press Release Geuvadis
  - Procedures for Quality control of Sequence Data
  - RNA and Exome Seq protocols
- **Wikis**
  - WP4 RNAseq
- **Annual Meeting 2011**
  - Agenda
  - Accommodation
  - Registration
  - Location
- **Calendar**
- **Public Home**

3. FOCUS: OUR COMMUNICATION ACTIONS TO THE GENERAL PUBLIC

A. OUR 'RESOURCES' PAGES

The GEUVADIS project being a coordinated action aiming at standardizing use of sequencing technologies in Europe from research to the clinic, we take it as an important goal to communicate on what we do to the general public. Indeed, although the progresses of medical sequencing can potentially affect the whole society, it often appears as a complicated technological issue, which understanding is restricted to a few specialists. Through our 'Resources' pages, on ELSI issues, Sequencing our genome and genomic medicine (screenshots below), we gathered basic information to inform the general public on these issues, and presented them in a simple and handy way. We took this opportunity to link these pages with extremely good educational material, notably created by Nature.





B. GEUVADIS PODCAST

In the same course of action, we decided to interview two actors from the project: Xavier Estivill, coordinator, and Esther Lizano, post-doc in Estivill's lab, for them to explain what they do in GEUVADIS, and why it is important to them. The 'podcast' sound-only material has been chosen for its ability to transfer core information while stimulating imagination, unlike video support.

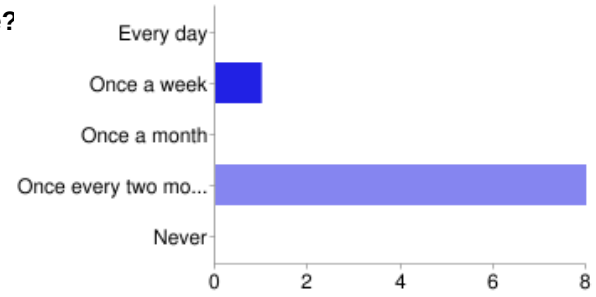
On our webpage, you can follow the podcast outline and see the speakers pictures. (See screenshot below). We are planning to push this idea further and realise more interviews during the course of the project; dedicated to the general public, but also to a more specialised public, on our main results. We created the 'Geuvadis Podcast logo' (Detail screenshot below) for website visitors to quickly identify this material throughout the pages.

ANNEX : RESULTS OF WEBSITE EVALUATION SURVEY

Total Respondants: 9 (50% of the groups)

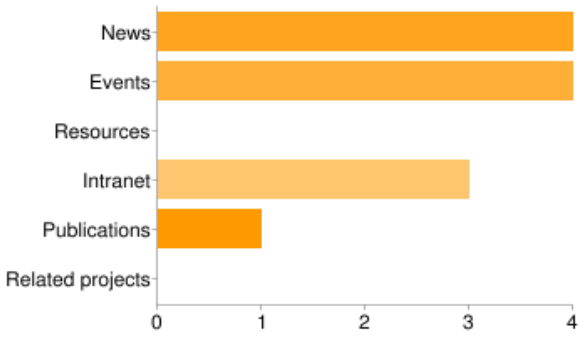
1) How often do you visit the project website?

Every day	0	0%
Once a week	1	11%
Once a month	0	0%
Once every two months	8	89%
Never	0	0%



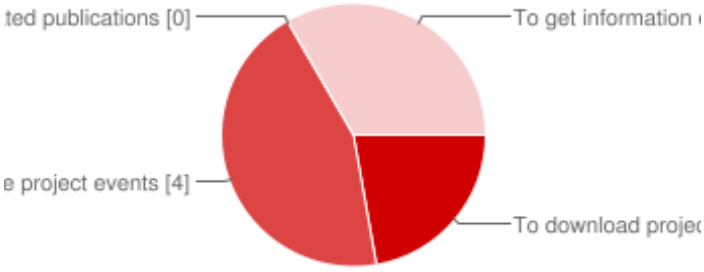
2) What page of the website do you visit most often?
People may select more than one checkbox, so percentages may add up to more than 100%.

News	4	44%
Events	4	44%
Resources	0	0%
Intranet	3	33%
Publications	1	11%
Related projects	0	0%



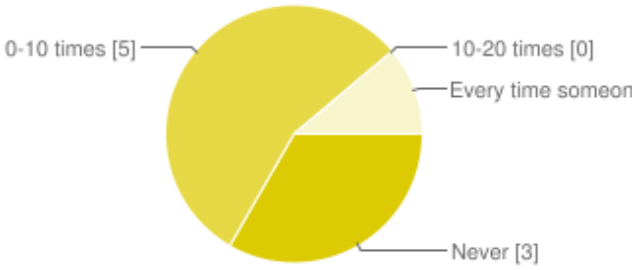
3) Why do you mainly go on the website ?

To download project documents (TC minutes, Event presentations, project flyer, reports)	2	22%
To get information on the project events	4	44%
To get information on project publications or related publications	0	0%
To get information on the project itself (partners, objectives, deadlines)	3	33%



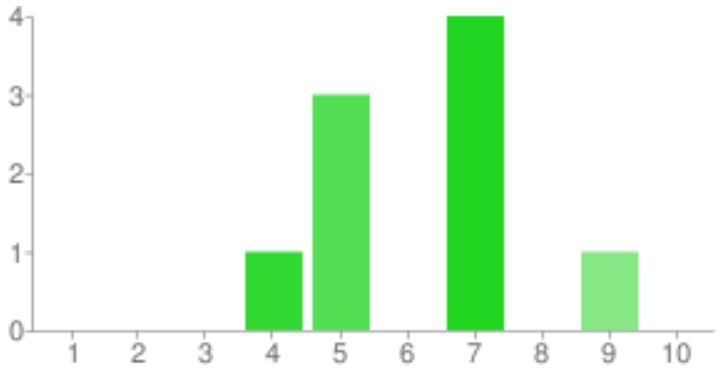
4) How often did you forward the website's URL to collaborators in order to give them an idea about the project?

Never	3	33%
0-10 times	5	56%
10-20 times	0	0%
Every time someone is interested in the project	1	11%



5) Scale the usefulness of the website

1	0	0%
2	0	0%
3	0	0%
4	1	11%
5	3	33%
6	0	0%
7	4	44%
8	0	0%
9	1	11%
10	0	0%



6) Your suggestions on how to improve the website: -content and design-

You'll find notes when corrective actions have been taken after the survey.

- I personally think that a wiki site type of resource would be sufficient, where necessary documents and events are available. This is not a very large project, so the main thing is to have handy site where to share internal info.
- One important thing that is missing, is a clear focus and target group. I would say we mostly aim our communication at other scientists, or even other genomic scientists. In my view, aspects for the general public should come into play as soon as there is something spectacular to tell them.
- add pictures and contact address of partners?
    - ***Note: this is available on the 'Partners'- 'Community' page***
- Although I have not used the Geuvadis homepage so much, I dare to say it is one of the better ones I have seen in similar projects. One glitch I have found is that the link in the name of each partner's institution doesn't work (the link works for the logo images)
    - ***Note: this has been corrected***
- I haven't used it much, but haven't really had a need for it either.
- I find both the content and the design very useful.
- We still have to link from our site to the GEUVADIS website. That would help me visiting the site more regularly and others getting introduced to the GEUVADIS initiative. We are in the middle of restructuring our websites but we will do so when completed.
- Better organize the documents in intranet. For example, include a document with the last agreement for the project to perform RNA-Seq in Hapmap samples....
    - ***Note: this has been corrected***