GEUVADIS-Sequencing Survey

Q1.1 Introduction:   THE GEUVADIS PROJECT  The latest high-throughput next-generation sequencing technologies allow investigators to sequence entire human genomes and transcriptomes at an affordable price and within a short time frame. An increasing number of research centers in Europe have access to these technologies, in-house or through regional, national and international infrastructures. Storing, disseminating and analysing the large amount of data produced generate major challenges. Tackling these challenges requires extensive exchange of data, information and knowledge between sequencing centers, bio-informatics networks, the medical research community and the industry at the European level. The GEUVADIS (Genetic EUropean VAriation in DISease) Consortium has four main aims: Develop standards in quality control and assessment of sequence data Develop models for sequencing data storage, access and exchange Develop standards for the handling, analysis and interpretation of sequencing data from DNA (and RNA) Develop guidelines on the handling of ethical, legal and social implications of phenotype prediction from sequence variation More information on our website: www.geuvadis.eu   PURPOSE OF SURVEY Collect quantitative and qualitative information on the current status of DNA sequence production, storage, exchange and analysis in Europe. Collect feedback from research/clinical centers on their main challenges and difficulties regarding the management of these large datasets potentially containing important medical information Collect information on local standardization efforts, and avoid duplication of efforts throughout Europe Create a roadmap/policy document outlining the necessary steps to take national standards to the European level  POTENTIAL PUBLICATION OF RESULTS Results of this survey will be presented at an internal GEUVADIS workshop, on October 30-31st 2012. We will disseminate the roadmap to potential funders (private of public), including the European Commission and other possible national public funders. We will submit an abstract to the Joint Conference of Human Genome Meeting 2013 and 21st International Congress of Genetics. DEADLINE: 30.09.2012 The survey should not take you more than 15 mins. It is structured as follows: - Introduction - Confidentiality agreement - Contact details - Main characteristics of the Institution - Technology - Data Use - Patients/participants and you (origin of samples, consent forms, incidental findings...) - Data storage - Data analysis - Clinical data management - Problems and suggestions

Q2.1 2) Confidentiality Agreement: Please enter your preferences below

* I want my name and personal information to be kept confidential
* I want the name of my Institution to be kept confidential
* I want to be informed of all future publications of the results

Q3.1 3) Contact Details: First name

Q3.2 Last Name

Q3.3 Email address

Q4.1 4) Main characteristics of the Institution:   Your Institution is:

* Research Center
* Clinical Center
* Hospital Department
* University
* Private Company

Q4.2 Institution Name

Q4.3 City

Q4.4 Country

Q4.5 Main source of funding for your institution

* Public funding
* Private funding

Q5.1 5) Technology enquiry:     Please indicate if the sequencing is processed:

* Locally - In-house facilities
* Externally

Q5.2 If 'Externally', please indicate if the sequencing data is produced:

* Data produced by a company
* Data produced by another center

Q5.3 Specify name of the center/company

Q5.4 Technology available

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Illumina MiSeq | Illumina GAII | IIlumina HiScanSQ | Illumina HiSeq1000 | Illumina HiSeq2000 | SOLiD 4 | SOLiD 5500 | SOLiD 5500xl | Roche GS Junior | Roche 454FLX+ | Ion Torrent Ion PGM | Ion Torrent Ion Proton | Other | Other |
| Number of machines |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Average runs/year |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Available automation of processes (Y/N) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Q5.5 Sequencing Capture method: please indicate the one(s) used in your institution

* Long-PCR
* Access Array System (Fluidigm)
* Microdroplet PCR (Raindance)
* AmpliSeq technology (Life Technologies)
* TrueSeq Amplicon Kit (Illumina)
* HaloPlex (Agilent)
* SureSelect (Agilent)
* SeqCap EZ (Nimblegen)
* TrueSeq Enrichment Kit (Illumina)
* TargetSeq (Life Technologies)
* Other

Q5.6 Please specify which other capture method you are using

Q6.1 6) Data Use:   Please indicate if the NGS data in your institution is used for:

* Research
* Clinical/Diagnostic

Q6.2 Type of data used for Research

* Whole genome sequencing
* Whole exome sequencing
* Targeted resequencing (custom panels)

Q6.3 Type of data used for Clinical/Diagnostic purposes

* Whole genome sequencing
* Whole exome sequencing
* Targeted resequencing (custom panels)

Answer If Data Use Research Is Selected

Q7.1 7) Research NSG: Participants and you     Origin of samples for research: Please indicate if the samples used for research in your institution are:

* samples from a Hospital/Clinic
* samples from other research projects
* specific collection of samples from research participants
* local biobank
* international biobank

Q7.2 Please specify from which biobank the samples for research are from:

Answer If Data Use Research Is Selected

Q7.3 Consent form:   Do you have a consent form template for your NGS research activities ?

* Yes
* No

If No Is Selected, Then Skip To Incidental medical findings

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Research Is Selected

Q7.4 Consent includes right to withdraw from the research

* Yes
* No

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Research Is Selected

Q7.5 Consent includes possibility to use samples/data for future research

* Yes
* Only if chosen by the patient
* No

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Research Is Selected

Q7.6 Open consent, for a variety of unspecific research projects

* Yes
* No

Answer If Data Use Research Is Selected

Q7.7 Incidental medical findings   Please indicate if incidental medical findings from your research activities are:

* never reported
* possibly reported if chosen by the patient
* always reported

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Research Is Selected

Q7.8 If incidental medical findings are sometimes reported, how is the medical relevance evaluated?

* on a case per case basis
* • by the Investigator
* • by the patient's medical doctor
* • by a committee including investigators and medical doctors
* only a set of specific diseases are reported

Q7.9 If only a set of diseases are reported, please indicate below which diseases:

Answer If Data Use Clinical/Diagnostoc Is Selected

Q7.10 7) Clinical NGS: Patients and you     Do you have a consent form template for your NGS clinical/diagnostic activities ?

* Yes
* No

If No Is Selected, Then Skip To End of Block

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Clinical/Diagnostic Is Selected

Q7.11 Main items of the consent form:

* Consent includes possibility to use the samples for research
* Consent includes possibility to use the data for research
* Consent includes possibility to use the samples and data for research

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Clinical/Diagnostic Is Selected

Q7.12 Incidental findings (not on the disease in question): Please indicate if medical incidental findings on characteristics that were not looked for are:

* never looked at: data is filtered out to avoid this possibility
* never reported
* possibly reported if chosen by the patient
* always reported

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Clinical/Diagnostic Is Selected

Q7.13 If reported how is the medical relevance evaluated?

* on a case per case basis
* • by the Investigator
* • by the patient's medical doctor
* • by a committee including investigators and medical doctors
* only a set of specific diseases are reported

Q7.14 Please describe here how you filter out your NGS data to avoid medical incidental findings:

Q7.15 If only a set of specific diseases are reported, please indicate below which diseases:

Q8.1 8) Data storageRaw data storage

* Local database
* Shared database
* Cloud

Q8.2 Please indicate if the local database for raw data storage is located in:

* Lab
* Hospital

Q8.3 How is the database shared?

* by research project
* with local investigators
* with local and international collaborators

Q9.1 9) Data Analysis:   Please indicate if the NGS data is processed in your institution by:

* Local cluster
* Supercomputing center

Q9.2 Name of the Supercomputing center where data is processed:

Q9.3 Do you have an agreement with local center, or company for raw NGS data analysis?

* Yes
* No

Q9.4 Please specify parties and nature of the agreement

Q9.5 Analysis pipeline(s)please indicate the type of pipeline(s) you are using

* Pipeline designed locally
* Pipeline designed by the sequencer's company
* Using a publically available or shared analysis pipeline
* Data processed by a specialized company

Q9.6 Please specify which pipeline you're using:

Q9.7 Specify which company processes your data:

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Clinical/Diagnostic Is Selected

Q10.1 10) Clinical data managementSource of clinical information

* from the patient/participant
* from the patient MD
* other...

Q10.2 Other source of clinical information, please specify:

Answer If 6) Data Use: &nbsp; Please indicate if the NGS data i... Clinical/Diagnostic Is Selected

Q10.3 Reporting of clinical results: Please indicate how NGS clinical results are reported to the patients

* report directly sent to the patient
* report sent to MD
* other...

Q10.4 If Other, please specify how the clinical NGS results are reported:

Q11.1 11) Problems and suggestions: Please tell us more about your sequencing experience here, focusing on main challenges and main suggestions to overcome these challenges:     Data production (Standardization of protocols, kits versions management, cost of sequencing)

Q11.2 Data storage (Storage capacity, data security, data access policy...)

Q11.3 Data exchange (Definition of access policy, protection of confidentiality of clinical and research findings)

Q11.4 Data analysis (Pipeline design, bioinformatics capacity...)