HUMAN TECHNOPOLE **NATIONAL FACILITY FOR GENOMICS** **CALL FOR ACCESS** **24-G-PILOT**

Table of Contents

1. INTRODUCTION

The Access of Researchers affiliated with Universities, *Istituti di Ricovero e Cura a Carattere Scientifico* (IRCCS), and Public Research Entities to Fondazione Human Technopole (HT) National Facilities (NFs) is regulated by the NF Access rules available on the NFs dedicated webpage (link).

Services offered by NFs are available through regular open calls for Access that are published yearly on the HT website (link) and are free of charge for the project (or aspects of the project) approved for Access.

The open call for Access is aimed at supporting Access to the technologies offered by the NFs and it is not meant to provide direct funding to the User. The costs for the activities to be performed at the NFs will be fully covered, including shipment of relevant material from and to the User's laboratory as well as travel and accommodation for the User while accessing the NF. Project-related costs (personnel, consumables, and other costs) at the User laboratory are not funded.

The User Access workflow comprises different steps, spanning from the initial submission of the application to evaluation and Access approval, Access to the performance of the service(s) and Access conclusion. A detailed description of the workflow is available on the NFs dedicated webpage (link).

1.1 Access modalities

Three different Access modalities can be requested. Their availability will vary, based on the service specifics of each NF:

- **"Simple" Access to NF or individual instruments thereof**: this modality is intended for Users involved in projects requiring technologies that are available at the NF for direct Access by User. This Access modality requires prior expertise with the technology of interest. After an initial introductory training aimed at defining the level of expertise of the User, the use of the instrument with limited supervision by NF staff is authorised. For defined NFs/ instruments/ services this Access modality may be restricted or not available.

- **Access to NF services**: This procedure entails the provision of services performed by NF staff on behalf of the User. NF services may include both standard services as well as, when foreseen by the technology development specifics of each NF, bespoke services conceived and discussed with the User. To allow the NF staff to best align the experimental activity to the research objective, the User may be invited, if needed, to assist the NF staff while performing the project or aspects of it.

- **Access to NF services including training**: This procedure entails training by NF staff to provide Users, in addition to or alternatively to the services described in the previous modality, with training courses and/or programs, aimed at transferring the expertise necessary for the independent use of the specific technology. In this case, technical and/or experimental activities are conducted with the active participation of the User. This type of Access is also aimed at researchers who want to acquire expertise for subsequent independent use of a specific technology in other laboratories.

2. TERMS AND DEFINITIONS

2.1 Access

"Access" refers to the authorized use of the NF and of the services offered. Such Access can be granted for sample preparation, set-up, execution and dismantling of experiments, education and training, expert support and analytical services, among others. Access to the NFs includes all infrastructural, logistical, technical and scientific support (including training) that is necessary to perform the aspects of the project approved for Access.

2.2 Researcher

"Researcher" is a professional engaged in the conception or creation of scientific knowledge. They conduct research and improve or develop concepts, theories, models, techniques, instrumentation, software or operational methods.

2.3 Principal Investigator

"Principal Investigator" (PI) is the Researcher affiliated with an eligible Institution with the role of independent Group Leader, who is responsible for coordinating the research activities conducted within the framework of the submitted project.

The PI shall hold a primary appointment as Group Leader at an eligible Institution, with the following requisites:

- Coordinate an independent research team.

- Have a supervisory role towards junior and/ or senior Researchers.

- Their Group has an autonomous budget sufficient to cover their current research expenses.

- Be the recipient of independent research funding as PI or co-PI.

Junior PI: Up to 6 years from their first appointment in an independent Group Leader position.

The period specified above may be extended beyond 6 years in the event of adequately documented career breaks, occurring before the submission of the application and resulting from:

i. Maternity leave: The time limit is increased by 18 months for each child born after their first appointment in an independent group leader position; if the Applicant is able to document a longer total maternity leave, the period of eligibility will be extended by a period equal to the documented leave, taken before the submission of the application. Maternity status must be documented by submitting the birth certificate of the child or children;

ii. Paternity leave: The time limit is increased by the actual amount of paternity leave taken before the application submission deadline for each child born after their first appointment in an independent group leader position. Paternity status must be documented by submitting the birth certificate of the child or children;

iii. Long-term illness of more than 90 days, or national service: The time limit is increased, for each eligible event occurring after their first appointment in an independent group leader position, by the actual amount of leave from which the Applicant has benefited prior to the application submission deadline.

Established PI: More than 6 years from their first appointment in an independent group leader position.

2.4 Applicant

"Applicant" is the Principal Investigator who applies to a NF open call for Access and who is responsible for the submitted project. They can be of any nationality and must be affiliated with an eligible Italian Institution, as detailed in section 4.

2.5 User

A "User" is intended as a Researcher affiliated with an eligible Institution who accesses the NFs to perform the approved activities or to support the National Facility staff while performing the approved service.

If requested by the Applicant, the User of the NF can also be a separate member of their research team.

3. APPLICATION TYPE

Applicants shall select the type of application they want to submit, choosing between two options:

a. **Standard** application for projects that are technically mature.

b. **Proof-of-concept** application for:

i. Projects with high scientific potential but with insufficient technical maturity or preliminary data.

ii. Projects aimed at setting up the experimental conditions required for a standard project, including methods or technology development projects.

iii. Time-limited Access projects (e.g., to acquire data to complete a manuscript, or preliminary data needed for a grant application, or single microscopy session).

4. ELIGIBILITY AND ADMISSIBILITY

PIs, as defined in section 2.3 of this call, affiliated with an eligible Institution are eligible to apply. The Applicant's role as a PI shall be confirmed by their Institution in a mandatory letter of Institutional endorsement.

Applications from Researchers who are not independent should be submitted by their group leader. Applicants are strongly encouraged to support NF Access by young Researchers (R1 and R2 profiles of the European Framework for Research Careers, link) who are part of their group. In this case, the Applicant shall indicate in the application form that the NF User is a member of their group, specifying User's career stage.

Below are the links to the relevant lists of eligible Institutions:

Universities: This category includes Institutions recognized by the Ministry of University and Research (link). In detail:

i. State funded public universities, listed under the following link.

ii. Specialized superior graduate schools or Institutions, listed under the following link.

iii. Legally recognized non-public universities, listed under the following link.

iv. On-line universities, listed under the following link.

Istituti di Ricerca e Cura a Carattere Scientifico (IRCCS): this category includes Institutions recognized by the Ministry of Health and listed at the following link.

Public research entities: this category includes Institutions recognized by the Ministry of University and Research and listed at the following link.

Eligible Institutions are strongly encouraged to limit the number of applications to the very best 2/ NF, with at least 50% coming from Junior PIs.

Applicants shall declare that they have not received funding to perform the submitted project (limited to the aspects included for Access to the NF) in their own laboratory, home Institution or elsewhere. Applicants shall confirm the economic and scientific feasibility for the aspects of the project to be performed outside the NFs.

Applicants cannot request Access for the same service if an approved Access is ongoing. Before submitting a new application for the same service, Applicant shall consult with the NF staff and confirm that the ongoing Access will be completed before the end of the next evaluation round. A clear motivation for the request must be provided.

During the same application window, a PI can submit only one application to the NFs (i.e., participate to only one call for Access). If more than one application is submitted, all will be rejected during administrative review. Applicants who have an application under evaluation are not allowed to submit another application before receiving notification of the results.

Applications must be written in English and must be complete (i.e., consist of all the requested elements and information). Incomplete applications will be considered not eligible and will be rejected at the administrative review stage.

The application form consists of six components:

1. Applicant's general information.
2. Justification for requesting Access to the NF.
3. Abstract.
4. Project proposal, including:

   a. *Significance.*

   b. *Innovation.*

   c. *Approach, including aims, preliminary data in support of the proposed experiments, experimental design and anticipated results.*

   d. *Environment, including facilities and resources available to support the aspects of the project to be performed elsewhere (i.e., outside the NF).*

5. Applicant's CV in NIH biosketch format.

6. Letter of Institutional Endorsement, addressing the following points:

   a. *Confirmation of the Applicant's role at their Institution, and their eligibility under the category of PI (see section 2.3).*

   b. *Confirmation that relevant authorisations, declarations and accreditation from the competent authority(ies) have been obtained in order to process samples and data through the NFs.*

   c. *Justification of the request for Access – including a statement on why the project cannot be performed at the Applicant's Institution.*

   d. *Confirmation that the Applicant has not received funding for performing the submitted project, for the aspects to be performed at the NFs, in their own laboratory, home Institution, or elsewhere.*

   e. *Confirmation of the project's economic and scientific feasibility for the aspects to be performed at the host Institution.*

   f. *Acceptance of NF Access terms and conditions.*

   The facsimile available as <u>Annex II</u> of this call shall be used as template.

7. Technical information, including:

   a. *Requested service(s), as described in <u>Annex I</u> of this call.*

   b. *Sample technical information.*

   c. *Requested preliminary data for technical feasibility analysis (if applicable).*

   d. *Whether the entire sample set is already available (otherwise indicate the date of availability of the entire sample set).*

   e. *Resources and expertise to receive and process the products – data (e.g. Cryo-EM micrographs) or reagents (e.g. human iPSCs) – generated by the NF.*

   f. *Research data management plan and bioinformatics support for data analysis, specifying:*

      i. *How the bioinformatics analysis of the data generated by the NF will be performed (if such analysis is not provided by the NF for Data Handling and Analysis).*

      ii. *How the data generated by the NF will be handled during and after the end of the project.*

      iii. *Whether and how the data will be shared/ made Open Access.*

      iv. *How data will be curated and preserved, including after the end of the project.*

Information provided in sections 1 and 6 are used for the eligibility and admissibility check.

Information provided in section 7 is used for assessing the technical feasibility of the aspects of the project to be performed at the NF.

The entire application is evaluated by the Standing Independent Evaluation Committee (SIEC) for assessing scientific merit.

5. APPLICATION SUBMISSION METHODS, CALL DEADLINE AND EVALUATION PERIODS

Applications shall be submitted exclusively through the web-based procedure managed by CINECA and accessible at this link, according to the terms and methods there indicated.

Applications are accepted throughout the year and assessed in three evaluation rounds per year. The time between one evaluation round and the next one is considered an application window.

- Evaluation round 1: applications submitted between the 1st of January and the 30th of April will enter the first evaluation round (May/ June).

- Evaluation round 2: applications submitted between the 1st of May and the 31st of August will enter the second evaluation round (September/ October).

- Evaluation round 3: applications submitted between the 1st of September and the 31st of December will enter the third evaluation round (January/ February).

This first call for Access (Call ID: 2024-G-Pilot) will be **open from the 10th of June 2024 (1 pm) to the 31st of December 2024 (1 pm)** and will involve 2 rounds of evaluation (September 2024 and January 2025).

A comprehensive list of services, available equipment and the technical requirements for Access as well as terms and conditions are available on the dedicated NFs webpage (link).

The complete list of offered services and technical requirements are available in the Annex I of this call.

6. EVALUATION OF APPLICATION

The evaluation procedure is conducted by the SIEC that is supported by a Panel of independent external Reviewers (Review Panel) selected by the SIEC on the basis of their scientific expertise.

Each Review Panel is composed of 2 SIEC members, who will act as Chairs, plus 10 appointed external Reviewers, with the relevant expertise.

Below is a scheme describing evaluation steps and timeline.

There are four application categories that are evaluated and ranked separately:

- Junior PI – Standard application

- Established PI – Standard application
- Junior PI – Proof of Concept application
- Established PI – Proof of Concept application

The NF User Access Office first performs an administrative review of the application to ensure that all the requested components have been provided, and that all eligibility criteria have been met.

The application is then sent to the Review Panel for assessing scientific merit and technical feasibility.

If the number of applications exceeds by a factor of 4 the estimated capacity of the NF, a triage will be applied within each application category by the relevant Review Panel.

Triage criteria will include:

    a. Justification for requesting Access to the NF.

    b. Field-Weighted Citation Impact (FWCI).

    c. Track record in securing research funding.

The application will remain confidential throughout the entire evaluation process. Reviewers will be asked to declare that they do not have any conflict of interest and they will be bound by a Confidentiality Agreement.

The application will be individually evaluated by three Reviewers who are part of the relevant Review Panel.

Proposals will be evaluated and ranked based on their average score, within each category.

An on-line meeting of the Review Panel may be requested by the Chairs if deemed necessary (for example to discuss proposals with highly discrepant scores).

At least 50% of the available Access will be allocated to applications from the two Junior PI categories.

6.1 Evaluation criteria

The scientific merit of the project is assessed based on the following criteria:

- **Significance**: Overall scientific merit of the proposed research. If all the experiments proposed are successful, how will the resulting knowledge advance the field?

- **Innovation**: Degree of innovation (conceptual and/ or technological), and ambition of the proposed study compared to the state-of-the-art in the relevant field.

- **Approach**: Appropriateness of proposed methodology, preliminary data in support of proposed experiments, and project feasibility.

- **Environment**: Facilities and resources available to support the aspects of the project to be performed elsewhere (i.e., outside the NF).

- **Justification for requesting Access to the NF**: Explanation on why the service cannot be performed at the host Institution, at a cost which is deemed affordable for the applicant.

- **Applicant**: PI's scientific background and expertise.

6.2 Scoring system

A numeric score between 1 (exceptional) and 9 (poor) is provided for each of the six evaluation criteria. Moreover, an overall project score including a short descriptive comment is provided as feedback to the Applicant.

- **HIGH**:
  - **Score 1 (Outstanding)** – The proposal successfully addresses all relevant aspects of the criterion. There are no weaknesses.
  - **Score 2-3 (Excellent - Very Good)** – The proposal addresses the criterion exceptionally well, aside from a small number of minor weaknesses.

- **MEDIUM**:
  - **Score 4-6 (Very good - Good)** – The proposal addresses the criterion well, but a number of weaknesses are present.

- **LOW**:
  - **Score 7-8 (Fair - Poor)** – The proposal broadly addresses the criterion, but there are significant weaknesses.
  - **Score 9 (Poor)** – The criterion is inadequately addressed, or there are serious inherent weaknesses.

6.3 Technical feasibility analysis

During the evaluation, the relevant experts from SIEC will receive a report from NF staff who will perform a comprehensive analysis of the proposed project's technical feasibility. Technical feasibility also includes an evaluation of the fulfilment of the technical requirements in terms of capacity to receive and process the research data generated by the NF, as described in the research data management plan. This latter evaluation is performed in consultation with the NF for Data Handling and Analysis.

Based on the technical maturity of the project, the application can be assessed as Feasible/ Not Feasible/ Pilot study required (switch from Standard to "Proof-of-Concept application track).

6.4 Evaluation results and Access approval

NF staff provides the SIEC with information on the resources needed (cost and time) to perform the highest ranked projects. The most positively evaluated applications that fulfil all technical requirements are approved for Access by SIEC, based on the capacity of the NF. NF staff schedules Access. A selected number of applications may be placed on a waiting list (in case of cancellations).

Evaluation results – Access approved, Access waitlisted, Access not approved – are communicated to the Applicant through the Access portal.

Applicants whose applications are placed on the waiting list will receive additional information advising whether the project can be Access approved or should be resubmitted within the subsequent application window.

7. AFTER ACCESS HAS BEEN APPROVED

After Access approval, a kick-off meeting is organized by the NF User Access Office and the Applicant is invited to meet NF staff to discuss the experimental design of the project and to finalize the project plan.

Once the project plan has been agreed, the NF User Access Office coordinates the signature of the required formal Agreements (e.g., Access Agreement, Collaboration Agreement, other) and the project can start.

8. AFTER ACCESS HAS BEEN COMPLETED

At the end of the activities carried out at the NF, and not later than 3 months thereafter if not differently agreed with the NF User Access Office, the User must submit a short report on the results obtained and the impact of the service on their research. Moreover, a final report to be published on the NFs website and describing the impact of the Access to the NF on the research project for which the service has been requested shall be provided upon publication of the relevant results. Users who will not be able to demonstrate the consistency and relevance of the activities carried out at the NF with the research project for which Access was requested will be considered not eligible to participate in the subsequent calls for Access.

Moreover, the User will be asked to fill in a brief, mandatory survey regarding their experience, providing feedback and suggestions for further service improvement.

The User must communicate to the NF User Access Office (via email to national.facilities@fht.org) any publication acknowledging the NF.

Research data obtained during Access shall be made available to the scientific community following the FAIR principles. User must inform the NF User Access Office (via email to national.facilities@fht.org) when and how the data are made public.

9. CONTACTS

Requests for information and/or clarifications concerning the application procedure may be sent to the dedicated e-mail address national.facilities@fht.org, indicating the call ID in the subject line.

10. REFERENCES

NF Access workflow_Convenzione (link)

NF Access rules_Convenzione (link)

NF Access Agreement_Convenzione (link)

11.   CHANGES TO THE CALL

Any changes or additions to this notice will be communicated through publication on the NFs website (link).

ANNEX I: NATIONAL FACILITY FOR GENOMICS: SERVICE LIST

HUMAN TECHNOPOLE

NATIONAL FACILITY FOR GENOMICS

CALL FOR ACCESS

24-G-PILOT

SERVICE LIST

TABLE OF CONTENTS

1. INTRODUCTION

The NF for Genomics provides state-of-the-art and innovative services in genomics. Its core mission is to implement robust experimental and analytical workflows to probe all major domains of genomic exploration, including but not limited to the analysis of DNA, RNA, chromatin, and other markers of epigenetic and regulatory activity. These techniques can be applied to diverse areas of biology, with resolution spanning to whole organisms, to tissues or individual cells. Overall, the NF for Genomics aims to empower research in all domains of genomics for the Italian scientific community at large.

The NF for Genomics includes four Infrastructural Units (IU):

UI1 High Throughput Sequencing mainly focused on providing state-of-the-art high-throughput sequencing services. This dedicated unit specializes in delivering top-tier genomics, transcriptomics, and epigenomics analyses, employing the latest protocols for sample processing and sequencing.

UI2 Multi-Omics Technologies specialized in multi-omics technologies. Its focus extends to providing cutting-edge services in single-cell and spatial multi-omics analysis, as well as long-read sequencing. This Unit closely collaborates with the Tissue Processing Infrastructural Unit part of the NF for Light Imaging, specifically dedicated to implementing spatial transcriptomics protocols.

UI3 Computational Genomics is the computational core of the NF for Genomics, the team is dedicated to developing, implementing, and maintaining automated pipelines for the pre-processing and primary data analysis of sequencing data. This unit closely collaborates with the NF for Data Handling and Analysis.

UI4 Technology Development stays at the forefront of innovation, this dedicated team is committed to methods and technology development. The highly qualified staff is continuously engaged in the evaluation of new technologies and instruments, considering them for acquisition into our NF. This dedicated unit specializes in the optimization and scale-up of custom experimental protocols, offering these services to the broader scientific community.

Below a list of available services with a detailed description of the service and technical requirements.

Appendix I below summarized the technical requirements for each available service.

2. SERVICE LIST

SID: G-001 – Whole Genome Sequencing (WGS)

Services description:

Whole Genome Sequencing (WGS) is a comprehensive and high-throughput technique that enables the complete DNA sequence of an organism's entire genome. Whole Genome Sequencing is a powerful tool with applications in various fields, including genomics research, personalized medicine, and clinical diagnostics. It provides a comprehensive view of an organism's genetic makeup, enabling a deeper understanding of genetic variations, evolution, and the genetic basis of diseases.

Here is an overview of the whole genome sequencing process:

DNA Extraction: The first step involves extracting genomic DNA (gDNA) from the biological sample, which could be cells, tissues, or even an entire organism. The goal is to obtain a high-quality and pure DNA sample (this task must be undertaken by the Users).

Library Preparation: The extracted DNA is then fragmented into smaller, manageable pieces. Adapters are added to these fragments to allow the subsequent sequencing process.

Sequencing: The prepared DNA library is subjected to high-throughput sequencing technologies, such as next-generation sequencing (NGS) platforms. These technologies generate short DNA sequences, or reads, from the DNA fragments.

Bioinformatic analysis of data from WGS can consist of several steps depending on the biological question:

- Read Mapping: When a reference sequence for the organism being sequenced is available (common case for model organisms), the reads are mapped to the reference generating an alignment.

- Genome Assembly: The generated sequencing data can be computationally assembled to reconstruct the complete genome sequence. This process involves overlapping the short reads to create longer contiguous sequences, or contigs.

- Genome Annotation: The assembled genome is annotated to identify and characterize various genomic features, including genes, regulatory elements, and non-coding regions. Annotation helps in understanding the functional elements within the genome.

- Variant Calling: The sequenced reads are compared to the reference genome to identify variations, such as single nucleotide polymorphisms (SNPs), insertions, deletions, and structural variations. This information is crucial for studying genetic diversity, population genetics, and identifying potential disease-causing mutations.

Bioinformatic analysis of WGS data can be provided as a combined service by the NF for Data Handling and Analysis. SID: NF62.001

Library preparation protocol:

Libraries will be prepared by following the protocol:

Illumina DNA PCR-Free Library Prep Reference Guide (1000000086922)

Illumina DNA PCR-Free offers a unique combination of benefits from on-bead tagmentation and PCR-free chemistry steps. On-bead tagmentation supports bead-based normalization, easy volume-based library pooling, and elimination of pre- and post-library quantification steps. The PCR-free workflow simplifies and reduces the overall workflow time while providing highly uniform coverage across repetitive or uneven genome regions. For sensitive applications such as human WGS, de novo assembly of microbial genomes, or tumour–normal variant calling, Illumina DNA PCR-Free delivers uniform coverage, and high-accuracy data.

Libraries sequencing and NGS coverage: Libraries will be sequenced using the NovaSeq 6000 system (Illumina) by generating 150 bp Paired End reads.

Next-generation sequencing (NGS) coverage describes the average number of reads that align to, or "cover," known reference bases. Sequencing coverage requirements vary by application, as noted below. At higher levels of coverage, each base is covered by a greater number of aligned sequences reads, so base calls can be made with a higher degree of confidence.

For WGS applied to population and rare disease studies an average human genome coverage of 20X per sample will be obtained.

For WGS applied to Cancer studies an average human genome coverage of 50X per sample will be obtained.

For WGS applied to Metagenomic studies the coverage per sample will be determined considering the type and complexity of the metagenome under study.

For WGS applied to Plant studies the coverage per sample will be determined considering the plant genome size and ploidy level.

Within this call the NF for Genomics will provide:

WGS for Population and medical studies (coverage 20X) for a maximum number of 3000 samples. Projects with a sample size ranging from a minimum of 1000 to a maximum of 3000 samples will be accepted.

WGS for Rare diseases studies (coverage 20X) for a maximum number of 100 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 100 samples will be accepted.

WGS for Cancer studies (coverage 50X) for a maximum number of 100 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 100 samples will be accepted.

WGS for Metagenomics studies (coverage determined considering type and complexity of the metagenome) for a maximum number of 200 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 200 samples will be accepted.

WGS for Plants studies (coverage determined considering the plant genome size and ploidy level) for a maximum number of 20 samples. Projects with a sample size ranging from a minimum of 5 to a maximum of 20 samples will be accepted.

Technical requirements: gDNA samples should be provided in low-bind full-skirted PCR plates (i.e. Eppendorf twin.tec PCR plates) sealed using peelable adhesive PCR films. Samples should be ordered column-wise without leaving empty positions. gDNA samples should have a concentration of at least 15 ng/µl in 50 ul of nuclease-free ultrapure water. gDNA samples should be quantified by using a fluorometer (i.e Qubit/ Glomax). Quality of gDNA should be evaluated by Agilent Tape Station/Bioanalyzer, samples should have a DIN≥6 (DNA Integrity Number≥6). Purity of the gDNA samples should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files

- QC report

- Mapping metrics (if reference genome/transcriptome is available)

For projects in which the bioinformatic analysis of WGS data will be requested as a combined service from the NF for Data Handling and Analysis, the following files will be delivered to the Users:

- Alignment (BAM files)

- Mapping metrics

- Identified variants (VCF files)

- Analysis QC reports

Access modality available: Access to NF services.

Services available in combination with the NF for Data Handling and Analysis: SID: NF62.001 - Bioinformatic analysis.

SID: G-002 – Whole Exome Sequencing (WES)

Services description:

Whole Exome Sequencing (WES) is a targeted sequencing approach that focuses on sequencing the protein-coding regions of the genome, known as the exome. The exome comprises the exons, which are the coding regions of genes, and represents only a small fraction (about 1-2%) of the entire genome. Despite this, the exome contains the majority of known disease-causing mutations, making WES a cost-effective alternative to Whole Genome Sequencing (WGS) for many applications. WES is widely used in both clinical and research settings. In clinical genetics, it is employed for diagnosing genetic disorders, identifying causative mutations, and understanding the genetic basis of rare diseases. In research, WES is valuable for studying the genetics of complex traits and diseases.

Here's an overview of the Whole Exome Sequencing process:

DNA Extraction: Like Whole Genome Sequencing, the process begins with the extraction of genomic DNA from the biological sample of interest, such as cells or tissues (this task must be undertaken by the Users).

Library Preparation: The extracted DNA is then fragmented into smaller pieces. Adapters are added to these fragments to allow the subsequent sequencing process.

Exome Capture: The next step involves selectively capturing and enriching the DNA fragments corresponding to the exonic regions. This is typically done using target-specific probes or baits designed to hybridize with and capture the exonic sequences.

Sequencing: The prepared exome library is subjected to high-throughput sequencing technologies, such as next-generation sequencing (NGS). The sequencing generates short DNA sequences, or reads, from the exonic regions.

Data Analysis: The generated sequencing data can be analysed bioinformatically to align the reads to a reference genome and identify genetic variants within the exonic regions. This includes single nucleotide variants (SNVs), insertions, deletions, and other structural variants.

Bioinformatic analysis of WES data can be provided as a combined service by the NF for Data Handling and Analysis. SID: NF62.001

Library preparation protocol:

The NF for Genomics uses the Twist Comprehensive Exome Panel for Whole Exome analysis; this panel offers coverage of greater than 99% of protein coding genes. The panel's superior performance provides the optimal exome sequencing solution, while focusing on the most accurate curated subset-CCDS database. It has been created to include expanded content of RefSeq and GENCODE databases. The panel targets in total 36.8 Mb with a design size of only 41.2 Mb covering more than 99% of protein coding genes.

DOC-001065_Protocol_UniversalAdapters-EnzFrag-REV5-SINGLES.pdf (twistbioscience.com)

Protocol_NGS_TargetEnrichmentFastHybridizationProtocol_2FEB22_Rev4.0.pdf (twistbioscience.com)

Libraries sequencing and NGS coverage: Libraries will be sequenced using the NovaSeq 6000 system (Illumina) by generating 150 bp Paired End reads.

For WES an average exome coverage of 50X per sample will be obtained.

Within this call the NF for Genomics will provide:

Whole Exome Sequencing for Rare diseases studies (coverage 50X) for a maximum number of 500 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 500 samples will be accepted.

Whole Exome Sequencing for Cancer studies (coverage 50X) for a maximum number of 700 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 700 samples will be accepted.

Technical requirements: gDNA samples should be provided in low-bind full-skirted PCR plates (i.e. Eppendorf twin.tec PCR plates) sealed using peelable adhesive PCR films. Samples should be ordered column-wise without leaving empty positions. gDNA samples should have a concentration of at least 5 ng/µl in 40 ul of nuclease-free ultrapure water. gDNA samples should be quantified by using a fluorometer (i.e Qubit/ Glomax). Quality of gDNA should be evaluated by Agilent Tape Station/Bioanalyzer, samples should have a DIN≥6 (DNA Integrity Number≥6). Purity of the gDNA samples should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files

- QC report

- Mapping metrics (if reference genome/transcriptome is available).

For projects in which the bioinformatic analysis of WGS data will be requested as a combined service from the NF for Data Handling and Analysis, the following files will be delivered to the Users:

- Alignment (BAM files)

- Mapping metrics

- Identified variants (VCF files)

- Analysis QC reports

Access modality available: Access to NF services.

Services available in combination with the NF for Data Handling and Analysis: SID: NF62.001 - Bioinformatic analysis.

SID: G-003 – Amplicon sequencing for microbiome analysis (16S-ITS)

Services description:

Microbiome analysis using 16S and ITS amplicon sequencing is a widely used technique to study the composition and diversity of microbial communities, particularly bacteria and fungi. The 16S ribosomal RNA (rRNA) gene is a molecular marker found in the genomes of bacteria and archaea, and its variable regions are commonly used for taxonomic classification, while ITS is used to profile fungal communities. Microbiome analysis using 16S and ITS amplicon sequencing is valuable in a range of fields, including environmental science, human health, agriculture, and more. It provides a cost-effective way to characterize microbial communities and understand their roles in various ecosystems or host-associated environments.

Here's an overview of the process:

Sample Collection: Microbiome analysis typically begins with the collection of samples from the environment of interest, such as soil, water, or biological samples like feces, saliva, or skin swabs (this task must be undertaken by the Users).

DNA Extraction: The next step involves extracting DNA from the collected samples. This DNA will contain the 16S rRNA gene from the microbial organisms present in the sample (this task must be undertaken by the Users).

PCR Amplification: Polymerase Chain Reaction (PCR) is used to selectively amplify the variable regions of the 16S rRNA gene. Primers designed to bind to conserved regions flanking the variable regions are used in this process. The choice of primers can influence the taxonomic resolution and coverage of the analysis.

Library Preparation: The PCR-amplified DNA is then converted into a sequencing library. Adapters are added during the amplification step to enable high-throughput sequencing.

High-Throughput Sequencing: The prepared library is subjected to high-throughput sequencing, commonly using next-generation sequencing (NGS) platforms. This step generates short DNA sequences (300bp) from the variable regions of the 16S rRNA and ITS genes.

Data analysis can consist of different steps depending on the biological question:

Taxonomic classification analysis: The reads are clustered into operational taxonomic units (OTUs) or more recently, into amplicon sequence variants (ASVs). Taxonomic classification is then performed by comparing the sequences to reference databases, allowing researchers to identify the types of bacteria present in the sample.

Diversity and Community Analysis: Various metrics, such as alpha diversity (diversity within a sample) and beta diversity (diversity between samples), are calculated to assess the microbial community structure. These analyses provide insights into the richness, evenness, and composition of the microbial communities.

Library preparation protocol:

Libraries will be prepared by following the protocol:

QIAseq 16S/ITS Region Panels (qiagen.com)

HB-2547-004_HB_QIAseq 16S_ITS Panel_1121_WW (4).pdf

QIAseq 16S/ITS Panels are used to perform a robust profiling of bacterial and fungal communities. The panels have been developed for sequencing 16S rRNA and ITS regions on Illumina platforms. QIAseq 16S/ITS Panels use "phased primers" to increase the quality of reads and base calling, and eliminate the need for PhiX spike-in. In addition, QIAseq 16S/ITS Panels incorporate low-bioburden reagents to decrease background contamination.

QIAseq 16S/ITS Panels can be configured to target different 16S variable regions and/or ITS according to different pools of primers.

Libraries sequencing and NGS coverage: Libraries will be sequenced using either the MiSeq or the NextSeq 2000 system (Illumina) by generating 300 bp Paired End reads.

The sequencing throughput per sample will be determined considering the type and complexity of the microbiome under study.

Within this call the NF for Genomics will provide:

<u>Microbiome analysis 16S-ITS</u> for a maximum number of 1200 samples. Projects with a sample size ranging from a minimum of 100 to a maximum of 1200 samples will be accepted.

> **Technical requirements:** DNA samples should be provided in 96 full-skirted PCR plate. 20 µl of High-quality DNA having a concentration ranging from 0,1-1 ng/ul should be provided. Quality of DNA should be evaluated by Agilent Tape Station/Bioanalyzer, samples should have a DIN≥6 (DNA Integrity Number≥6). Purity of the DNA samples should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). DNA samples should be quantified by using a fluorometer (i.e Qubit/ Glomax).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files

- QC report

Access modality available: Access to NF services.

SID: G-005 – mRNA sequencing from standard and low input

Services description:

mRNA sequencing is a powerful molecular biology technique used to analyse the transcriptome of a biological sample. The transcriptome refers to the complete set of RNA molecules, in particular messenger RNA (mRNA), in a cell or tissue. mRNA sequencing is widely used in genomics research, functional genomics, and clinical studies to understand gene expression patterns, identify novel transcripts, and investigate how gene expression varies under different conditions.

Here's a brief description of the mRNA sequencing process:

<u>Isolation of RNA</u>: The first step involves extracting RNA from the biological sample, such as cells or tissues. This can be done using various methods to ensure the preservation of the RNA molecules (this step must be taken by Users).

<u>cDNA synthesis</u>: Complementary DNA (cDNA) is synthesized from the fragmented RNA using reverse transcription. This step converts RNA into a complementary DNA strand, creating a library of cDNA molecules.

<u>Library preparation</u>: The cDNA library is then prepared for sequencing. Adapters are added to the cDNA fragments, allowing them to be sequenced efficiently.

<u>Sequencing</u>: The prepared library is subjected to high-throughput sequencing techniques, such as next-generation sequencing (NGS). This step generates short sequences, or reads, from the cDNA fragments.

<u>Data analysis</u>: The generated sequencing data is then analysed bioinformatically. The reads are mapped to a reference genome or transcriptome to identify the gene expression levels and discover novel transcripts.

Bioinformatic analysis of mRNAseq data can be provided as a combined service by the NF for Data Handling and Analysis. SID: NF62.001

Library preparation protocol:

mRNA sequencing from standard input

Libraries will be prepared by following the protocol:

[Illumina Stranded mRNA Prep, Ligation Reference Guide (1000000124518)](#)

Illumina Stranded mRNA Prep enables precise measurement of strand orientation, uniform coverage, and high-confidence discovery of features such as novel isoforms, gene fusions, and allele-specific expression. Illumina Stranded mRNA Prep is optimized to provide good polyA capture efficiency and coverage uniformity. It minimizes the required sequencing depth for accurate, unbiased detection of the coding transcriptome.

mRNA sequencing from low input

Libraries will be prepared by following the protocol:

[SMART-Seq® v4 PLUS Kit User Manual (takarabio.com)](#)

This kit uses oligo(dT) priming to generate high-quality, full-length cDNA directly from multiple intact cells or total RNA. In addition to the cDNA synthesis kit, the SSv4 PLUS kit also includes a library preparation kit and a single-use unique dual index (UDI) plate to generate Illumina-compatible sequencing libraries.

The SMART (Switching Mechanism at 5' End of RNA Template) technology employed by the SMART-Seq v4 kit provides full-length transcript information, enabling analysis of transcript isoforms, gene fusions, point mutations, etc. Additionally, it incorporates locked nucleic acid (LNA) technology in the SMART-Seq v4 Oligo for more efficient template switching. This allows for the identification of higher numbers of genes relative to other methods, high reproducibility, even gene-body coverage, and an accurate representation of GC-rich transcripts.

Libraries sequencing and NGS coverage: Libraries will be sequenced using either the the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating 100 bp Paired End reads.

On average 40 million reads pairs (20 million clusters 100bp PE) per sample will be generated for species with reference genomes and 80 million reads pairs (40 million clusters 100bp PE) will be generated for species without reference genomes.

Within this call the NF for Genomics will provide:

mRNAseq for studies on diseases (humans, animal models, organoids) for a maximum number of 300 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 300 samples will be accepted.

mRNAseq for differentiation studies (humans, animal models, organoids) for a maximum number of 200 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 200 samples will be accepted.

mRNAseq in plants studies for a maximum number of 100 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 100 samples will be accepted.

Technical requirements:

mRNA sequencing from standard input: totalRNA samples should be provided in LoBind full skirted PCR plate (i.e. Eppendorf twin.tec PCR plates) properly sealed using peelable adhesive PCR films. Samples should be ordered column-wise without empty wells. The totalRNA amount should be between 500ng-1000ng in a volume range of 20-50ul in nuclease-free ultrapure water. totalRNA should be DNase treated and the RIN≥7 (RNA Integrity Number≥7), the quality of totalRNAs should be evaluated by Agilent Tape Station/Bioanalyzer. Purity of the totalRNAs should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). totalRNA samples should be quantified by using of a fluorometer (i.e Qubit/Glomax).

mRNA sequencing from low input: RNA samples should be provided in LoBind full skirted PCR plate (i.e. Eppendorf twin.tec PCR plates) properly sealed using peelable adhesive PCR films. Samples should be ordered column-wise without empty wells. The RNA amount should be between 50pg-10ng in a volume range of 15-20 ul in nuclease-free ultrapure water. Total RNA should be DNase treated and the RIN≥7 (RNA Integrity Number≥7), the quality of totalRNAs should be evaluated by Agilent Tape Station/Bioanalyzer. Purity of the totalRNAs should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). totalRNA samples should be quantified by using of a fluorometer (i.e Qubit/Glomax).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report

For projects in which the bioinformatic analysis of WGS data will be requested as a combined service from the NF for Data Handling and Analysis, the following files will be delivered to the Users:

- Transcriptome alignment (BAM files)
- Mapping metrics
- Raw and normalized expression matrix
- Differential expression analysis
- Analysis QC reports

Access modality available: Access to NF services.

Services available in combination with the NF for Data Handling and Analysis: SID: NF62.001 - Bioinformatic analysis

SID: G-006 – totalRNA sequencing from standard input

Services description:

Total RNA sequencing is a powerful and widely used molecular biology technique that aims to analyse and quantify the entire transcriptome of a biological sample. The transcriptome represents the complete set of RNA molecules, including messenger RNA (mRNA), and non-coding RNAs, present in a cell or tissue. Total RNA sequencing provides a comprehensive view of the transcriptome, allowing researchers to gain insights into gene expression patterns, identify novel transcripts, and understand the regulatory mechanisms underlying various biological processes.

Here is a step-by-step description of the Total RNA sequencing process:

RNA Extraction: Total RNA is isolated from the biological sample of interest, such as cells or tissues. This extraction process is crucial to obtain a representative snapshot of the RNA present in the sample (this step must be taken by Users).

Library Preparation: After ribosomal RNA depletion, the RNA is converted into a complementary DNA (cDNA) library through a process called reverse transcription. This step involves the use of reverse transcriptase to synthesize cDNA from the RNA template.

Fragmentation and Adaptor Ligation: The cDNA is then fragmented, and sequencing adaptors are added at the ends of the fragments. Adaptors contain sequences necessary for the subsequent steps of the sequencing process.

Library Amplification: The prepared library is amplified using polymerase chain reaction (PCR) to generate sufficient material for sequencing.

Sequencing: The amplified cDNA library is then subjected to high-throughput sequencing, with NGS sequencing platforms.

Data Analysis: The raw sequencing data is processed through bioinformatics pipelines to align the reads to a reference transcriptome or assemble them into transcripts *de novo*. This step helps identify and quantify gene expression levels, alternative splicing events, and novel transcripts.

Bioinformatic analysis of totalRNAseq data can be provided as a combined service by the NF for Data Handling and Analysis. SID: NF62.001

Library preparation protocol:

Libraries will be prepared by following the protocol:

[Illumina Stranded Total RNA Prep Ligation with Ribo-Zero Plus Reference Guide](#)

Illumina Total RNA Prep with Ribo-Zero Plus supports a broad range of RNA inputs. It's compatible with various sample types, including formalin-fixed paraffin-embedded (FFPE) and other low-quality samples. The included Ribo-Zero Plus or Ribo-Zero Plus Microbiome removes abundant RNA from multiple species, including human, mouse, rat, bacteria, and epidemiology samples or complex microbial samples, including stool samples, for meta-transcriptomic studies.

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating 100 bp Paired End reads.

On average 80 million reads pairs (40 million clusters 100bp PE) per sample will be generated for species with reference genomes and 160 million reads pairs (80 million clusters 100bp PE) will be generated for species without reference genomes or for meta-transcriptomics studies.

Within this call the NF for Genomics will provide:

totRNAseq for studies on diseases (humans, animal models, organoids) for a maximum number of 200 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 200 samples will be accepted.

totRNAseq for differentiation studies (humans, animal models, organoids) for a maximum number of 200 samples. Projects with a sample size ranging from a minimum of 50 to a maximum of 200 samples will be accepted.

totRNAseq in plants studies for a maximum number of 100 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 100 samples will be accepted.

totRNAseq in infection studies (host/pathogen interaction, in vitro and ex-vivo) for a maximum number of 100 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 100 samples will be accepted.

> **Technical requirements:** totalRNA samples should be provided in LoBind full skirted PCR plate (i.e. Eppendorf twin.tec PCR plates) properly sealed using peelable adhesive PCR films. Samples should be ordered column-wise without empty wells. The totalRNA amount should be between 500ng-1000ng in a volume range of 20-50ul in nuclease-free ultrapure water. totalRNA should be DNase treated and the RIN≥4 (RNA Integrity Number≥4), the quality of totalRNAs should be evaluated by Agilent Tape Station/Bioanalyzer. Purity of the totalRNAs should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). totalRNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report

For projects in which the bioinformatic analysis of WGS data will be requested as a combined service from the NF for Data Handling and Analysis, the following files will be delivered to the Users:

- Transcriptome alignment (BAM files)
- Mapping metrics
- Raw and normalized expression matrix
- Differential expression analysis
- Analysis QC reports

Access modality available: Access to NF services.

Services available in combination with the NF for Data Handling and Analysis: SID: NF62.001 - Bioinformatic analysis

SID: G-007 – smallRNA sequencing

Services description:

Small RNA sequencing is a specialized technique designed to analyse and profile small RNA molecules present in a biological sample. Small RNAs are short RNA molecules, typically ranging from 18 to 30 nucleotides in length, and they play essential roles in various cellular processes, including gene regulation, RNA silencing, and post-transcriptional control. Small RNA sequencing is widely used to study the expression profiles of miRNAs and other small RNAs, providing valuable insights into their roles in gene regulation, development, and disease.

Here is a step-by-step description of small RNA sequencing:

RNA Extraction: Like total RNA sequencing, the process begins with the extraction of RNA from the biological sample. However, in small RNA sequencing, specific methods can be employed to selectively enrich small RNA molecules (this step must be taken by Users).

Size Selection: The extracted RNA can be subjected to size selection to isolate the small RNA fraction, but also totalRNA can be used as input.

Adapter Ligation: After smRNAs poly-adenilation and retro transcription adapters are ligated to both ends of resulting cDNAs and serve as primers during the library preparation steps.

Library Amplification: The cDNA library is then amplified using PCR. This step adds the necessary sequences for subsequent sequencing steps.

Size Selection and Purification: The amplified library undergoes a size selection step to remove unwanted fragments and purify the small RNA-containing fraction.

Sequencing: The purified small RNA library is subjected to high-throughput sequencing, using NGS platforms.

Data Analysis: Bioinformatics analysis is performed on the sequencing data to identify and quantify different types of small RNAs, such as microRNAs (miRNAs), small interfering RNAs (siRNAs), and piwi-interacting RNAs (piRNAs). This involves aligning the sequences to a properly annotated reference genome or small RNA databases.

Bioinformatic analysis of smallRNAseq data can be provided as a combined service by the NF for Data Handling and Analysis. SID: NF62.001

Library preparation protocol:

Libraries will be prepared by following the protocol:

SMARTer® smRNA-Seq Kit for Illumina® User Manual (takarabio.com)

The SMARTer smRNA-Seq Kit is used to generate small RNA-seq libraries for sequencing on Illumina platforms. This kit works directly with total RNA or enriched small RNA (including microRNA). It incorporates features of the SMARTer Stranded RNA-Seq kits, including the SMART (Switching Mechanism at the 5' end of RNA Template) technology, and primers that

include locked nucleic acids (LNAs). This kit allows Users to analyze a wide range of smRNA species and generate sequencing libraries of considerable complexity. Illumina adapters and index sequences are incorporated in a ligation-free manner during library amplification, ensuring that diverse smRNA species are represented with minimal bias.

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating 100 bp Paired End reads.

On average 20 million reads pairs (10 million clusters 50bp PE) per sample will be generated.

Within this call the NF for Genomics will provide:

smallRNAseq in humans, animal models, organoids samples for a maximum number of 60 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 60 samples will be accepted.

Technical requirements: totalRNA samples should be provided in LoBind full skirted PCR plate (i.e. Eppendorf twin.tec PCR plates) properly sealed using peelable adhesive PCR films. Samples should be ordered column-wise. The totalRNA amount should be between 100ng-1000ng in a volume range of 20-50ul in nuclease-free ultrapure water. RNA should be DNase treated and the RIN≥7(RNA Integrity Number≥7), the quality of totalRNAs should be evaluated by Agilent Tape Station/Bioanalyzer. Purity of the totalRNAs should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). totalRNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report

For projects in which the bioinformatic analysis of WGS data will be requested as a combined service from the NF for Data Handling and Analysis, the following files will be delivered to the Users:

- Transcriptome alignment (BAM files)
- Mapping metrics
- Raw and normalized expression matrix
- Differential expression analysis
- Analysis QC reports

Access modality available: Access to NF services.

Services available in combination with the NF for Data Handling and Analysis: SID: NF62.001 - Bioinformatic analysis

SID: G-008/G-012 – Single-cell 3'RNAsequencing or Single-cell gene Expression Flex (10X Genomics)

Services description:

Single-cell 3'RNAsequencing

Single-cell 3' RNA sequencing with 10x Genomics technology is a powerful method for studying gene expression at the single-cell level. It enables the profiling of thousands to tens of thousands of single cells in parallel. It captures heterogeneity within cell populations, allowing the identification of rare cell types and subpopulations. It reveals differences in gene expression between individual cells, providing a more nuanced understanding of cellular diversity.

Here's an overview of the process:

Cell Capture: Single cells are isolated into individual droplets using microfluidics technology. Each droplet contains a bead with a unique barcode.

Barcoding and cDNA Synthesis: Within each droplet, the cell's RNA is captured at the 3' end and is barcoded using both a barcode associated to each cell and unique molecular identifiers (UMIs) that will allow transcripts counting and normalization. This step ensures that each transcript originating from the same cell gets the same cell barcode. cDNA (complementary DNA) is synthesized from the barcoded RNA templates.

Library Preparation: The cDNA is then amplified, and Illumina sequencing adapters are added, preparing the library for sequencing.

Pooling and Sequencing: All the individually barcoded and amplified cDNAs from the different cells are pooled together. The pooled library is sequenced on a NGS platform.

Data Analysis: After sequencing, bioinformatics tools are used to demultiplex and assign the reads to their respective cells based on the unique cellular barcodes. The resulting data provides a transcriptomic profile for each individual cell, allowing the analysis of gene expression at a single-cell resolution (data analysis is the responsibility of Users).

Single-cell gene Expression Flex:

The Fixed RNA Profiling assay, also called the Single Cell Gene Expression Flex assay by 10X Genomics, is a way to prepare single cell RNA-seq libraries from formaldehyde-fixed cells and tissues as well as FFPE tissue blocks. This process allows researchers to lock in the biological state of their samples at the time of fixation and store the fixed cell suspension for at least 6 months at -80°C.

Because the fixation process is expected to lead to a certain amount of RNA degradation, the fixed RNA assay uses pairs of sequence-specific probes to bind transcripts for approximately 18,000 genes for human and about 19,000 genes for mouse. These probe pairs bind adjacent sites on their target transcripts, are ligated together, and then become the substrate for library construction.

Sample multiplexing allows Users to label discrete cell populations with a sample-specific barcode so that these samples can be pooled into a single GEM droplet reaction, reducing the number of reagent units needed for the experiment and potentially lowering the overall cost. Multiplexing in the Flex assay is built in by the inclusion of sample barcodes in the transcript-

detecting probes. The 64-reaction multiplex kit, that will be used by the NF for Genomics to provide this service, includes four units of gel beads and 16 sets of probes with unique barcodes, allowing to pool up to 16 samples and to recover and analyse potentially 128.000 cells per GEM reaction.

This type of analysis can be performed only using the Chromium X platform.

Library preparation protocols:

In the Single-cell 3' RNA sequencing approach depending on the number of cells the User wishes to analyse; the samples will be processed either using the Chromium controller platform (which allows to recover and analyse a maximum of 10,000 cells per single sample) or using the Chromium X platform (which allows to recover and analyse a maximum of 20,000 cells per single sample). Should the User wish to multiplex several samples together, the Cell Plex protocol will be used to barcode cells from different samples before multiplexing them. In the case of cells multiplexed with the CellPlex protocol the maximum cell recovery that can be achieved with the Chromium controller platform will be 30,000 cells per pool and with the Chromium X platform will be 60,000 cells per pool.

Here below the list of protocols that will be used by NF staff to perform Single-cell 3'RNAsequencing or Single-cell gene Expression Flex experiments:

CG000315_ChromiumNextGEMSingleCell3-_GeneExpression_v3.1_DualIndex__RevE.pdf (10xgenomics.com)

https://cdn.10xgenomics.com/image/upload/v1689110007/support-documents/CG000416_Chromium_NextGEM_SingleCell3-_HT_v3.1_GeneExp_RevD.pdf

https://cdn.10xgenomics.com/image/upload/v1694503892/support-documents/CG000527_ChromiumFixedRNAProfiling_MultiplexedSamples_UserGuide_Rev_E.pdf

CG000391_DemonstratedProtocol_CellLabelingwithCellMultiplexingOligo_RevB.pdf (10xgenomics.com)

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating reads length as required by the Single-cell 3'RNAsequencing or the Single-cell gene Expression Flex protocol (10x Genomics).

On average 50.000 reads per cell will be generated for sequencing GEX libraries obtained with the Single-cell 3'RNA sequencing protocol and at least 10.000 reads per cell for sequencing GEX libraries obtained with the Single-cell gene Expression Flex protocol.

Within this call the NF for Genomics will provide:

Single-cell 3'RNAsequencing for studies on diseases (humans, animal models, organoids) for a maximum number of 32 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 32 samples will be accepted.

Single-cell 3'RNAsequencing for differentiation studies (humans, animal models, organoids) for a maximum number of 32 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 32 samples will be accepted.

Single-cell gene Expression Flex (humans, animal models, organoids) (16-plex) for a maximum number of 64 samples. Projects with a sample size ranging from a minimum of 32 to a maximum of 64 samples will be accepted.

Technical requirements:

Single-cell 3'RNAseq: Frozen Cells should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). If performing standard assay (no multiplexing of samples), provide an input of minimum 500k frozen cells per sample. If multiplexing with the CellPlex is required, provide an input of minimum 2 million frozen cells per sample. Cells should be prepared following the protocol:

CG00053_Handbook_CellPreparation_SingleCellProtocols_Rev_D.pdf (10xgenomics.com)

Cell viability before freezing should be ≥80%.

Only samples that will have a cell viability ≥70% and a low level of cellular debris after thawing, will be processed.

RNA Flex: Fixed cells should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). Cells fixation should be performed following the protocols listed below:

CG000553_DemonstratedProtocol_TissueFixationDissociation_RevB.pdf (10xgenomics.com)

CG000478_DemonstratedProtocol_Cell_NucleiFixation_Chromium_FixedRNA_Profiling_RevD.pdf (10xgenomics.com)

Cell viability before fixation should be ≥80%.

Two aliquots containing 500k to 2M fixed cells are required per sample.

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report
- Mapping metrics

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-009/012/014 – Single-cell Immune profiling-V(D)J (10X Genomics)

Services description:

Single-cell immune profiling with 10x Genomics technology is a powerful method for 5' RNA sequencing at the single-cell level and to profile at the same time the T-cells and/or B-cells receptors at single cell level by sequencing the V(D)J regions.

Here's an overview of the process:

<u>Cell Capture</u>: Single cells are isolated into individual droplets using microfluidics technology. Each droplet contains a bead with a unique barcode.

<u>Barcoding and cDNA Synthesis</u>: Within each droplet, the cell's RNA is captured at the 5' end and is barcoded using both a barcode associated to each cell and unique molecular identifiers (UMIs) that will allow transcripts counting and normalization. This step ensures that each transcript originating from the same cell gets the same cell barcode. cDNA (complementary DNA) is synthesized from the barcoded RNA templates.

<u>Library Preparation</u>: Three libraries can be potentially obtained from the cDNA synthesized from the barcoded RNA templates. The GEX library, for Gene Expression profiling, resulting from cDNA amplification and addition of Illumina sequencing adapters. Libraries for V(D)J Gene Profiling deriving from targeted cDNA amplification of the V(D)J regions specific of T cell receptors (TCRs) and B cell receptors (BCRs).

<u>Sequencing</u>: The pooled libraries are sequenced on a high-throughput NGS platform.

<u>Data Analysis</u>: After sequencing, bioinformatics tools are used to demultiplex and assign the reads to their respective cells based on the unique barcodes and to identify different T and B cell clones, their abundance, and diversity (data analysis is the responsibility of Users).

Library preparation protocol:

Depending on the number of cells the User wishes to analyse, the samples will be processed either using the Chromium controller platform (which allows to recover and analyse a maximum of 10,000 cells per single sample) or using the Chromium X platform (which allows to recover and analyse a maximum of 20,000 cells per single sample).

Libraries will be prepared by following the protocols:

CG000331_ChromiumNextGEMSingleCell5-v2_UserGuide_RevE.pdf (10xgenomics.com)

CG000330_ChromiumNextGEMSingleCell5_v2_CellSurfaceProtein_UserGuide_RevF.pdf (10xgenomics.com)

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating reads length as required by the Single-cell Immune profiling-V(D)J protocol (10x Genomics).

On average 50.000 reads per cell will be generated for sequencing GEX libraries and at least 5.000 reads per cell for sequencing TCR and BCR libraries obtained with the Single-cell Immune profiling-V(D)J protocol.

Within this call the NF for Genomics will provide:

<u>Single-cell Immune profiling-VDJ for studies on diseases (humans, animal models)</u> for a maximum number of 64 samples. Projects with a sample size ranging from a minimum of 16 to a maximum of 64 samples will be accepted.

Technical requirements: Frozen Cells should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). Provide an input of minimum 500k frozen cells per sample. Cells should be prepared following the protocol:

CG00053_Handbook_CellPreparation_SingleCellProtocols_Rev_D.pdf (10xgenomics.com)

Cell viability before freezing should be ≥80%.

Only samples that will have a cell viability ≥70% and a low level of cellular debris after thawing will be processed.

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report
- Mapping metrics

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: 010/012/013 – Single-cell multiome ATAC + Gene expression (10X Genomics)

Services description:

Single-cell multiome ATAC + Gene Expression is a cutting-edge technology that enables the simultaneous profiling of chromatin accessibility and gene expression at the single-cell level, providing a comprehensive view of the molecular landscape within individual cells. This technology is widely used in various biological research areas, including understanding cellular diversity in tissues, identifying cell types and states, deciphering regulatory networks, and gaining insights into how chromatin accessibility relates to gene expression at the single-cell level.

Here's an overview of the key steps:

Single-Cell Resolution and Multiome analysis: The technology captures information at the level of individual cells/nuclei, combining two crucial molecular profiles – ATAC-seq (Assay for Transposase-Accessible Chromatin with high-throughput sequencing) for chromatin accessibility and RNA-seq (or gene expression profiling) for understanding the transcriptional activity of each cell/nucleus.

ATAC-seq: This part of the technology focuses on assessing the accessibility of chromatin, providing insights into the regions of the genome that are open and accessible for transcription factors and other regulatory elements.

RNA-seq: Concurrently, the gene expression analysis captures the messenger RNA transcripts present in each nucleus, shedding light on the active genes and their expression levels.

The experimental workflow for Single-cell multiome ATAC + Gene Expression using the 10X Genomics platform typically involves several key steps:

Cell Isolation and Nuclei Preparation: Begin by isolating the cells of interest. This could be from tissues or cell cultures. Prepare a single-cell suspension to ensure individual cells can be processed independently (this step must be taken by Users). Extract the nuclei from the cells, as the chromatin accessibility information is primarily derived from the open regions in the nuclei.

Transposase Reaction (ATAC-seq): Add a transposase enzyme to the nuclei suspension. The transposase fragments the chromatin and adds sequencing adapters simultaneously. These sequencing adapters are essential for downstream library preparation.

Chromium Gel Bead-in-Emulsion (GEM) Formation: The isolated nuclei are encapsulated into Gel Bead-in-Emulsion (GEM) droplets using the 10X Genomics Chromium platforms. Each GEM contains a unique barcode that will be associated with both the chromatin accessibility and gene expression data from a specific nucleus.

Reverse Transcription: In the same GEMs, reverse transcription is performed to convert nuclear RNA into complementary DNA (cDNA).

Library Preparation: Following the transposase reaction and reverse transcription, PCR amplification is performed to create the final sequencing libraries (ATAC and GEX libraries). Sequencing adapters are added during this step to allow sequencing on NGS platforms.

High-Throughput Sequencing: Sequence the prepared libraries using high-throughput sequencing platforms. The sequencing reads contain information about both chromatin accessibility and gene expression in individual cells/nuclei.

Data Analysis: Employ specialized bioinformatics tools and pipelines to process and analyze the sequencing data. Demultiplex the data based on cell barcodes, align the reads to the reference genome, and perform separate analyses for chromatin accessibility and gene expression (data analysis is the responsibility of Users).

Library preparation protocol:

Libraries will be prepared by following the protocol:

CG000338_ChromiumNextGEM_Multiome_ATAC_GEX_User_Guide_RevF.pdf (10xgenomics.com)

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating reads length as required by the Single-cell multiome ATAC + Gene expression protocol (10x Genomics).

On average 50.000 reads per cell/nucleus will be generated for sequencing GEX libraries and 25.000 reads per cell/nucleus for sequencing ATACseq libraries obtained with the Single-cell multiome ATAC + Gene expression protocol.

Within this call the NF for Genomics will provide:

Single-cell multiome ATAC + Gene expression for studies on diseases (humans, animal models, organoids) for a maximum number of 16 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 16 samples will be accepted.

<u>Single-cell multiome ATAC + Gene expression for differentiation studies (humans, animal models, organoids)</u> for a maximum number of 16 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 16 samples will be accepted.

Technical requirements: Frozen cells should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). Provide an input of minimum 500k frozen cells. Cells should be prepared following the protocol:

CG00053_Handbook_CellPreparation_SingleCellProtocols_Rev_D.pdf (10xgenomics.com)

Cell viability before freezing should be ≥80%.

Nuclei will be prepared by NF staff following the protocol:

CG000365_DemonstratedProtocol_NucleiIsolation_ATAC_GEX_Sequencing_RevC.pdf (10xgenomics.com)

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report
- Mapping metrics

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-011/013 – Single-cell ATAC sequencing (10X Genomics)

Services description:

Single-cell ATAC sequencing with the 10X Genomics platform involves profiling the chromatin accessibility of individual cells/nuclei at a high resolution. The assay, commonly known as scATAC-seq (Single-cell Assay for Transposase-Accessible Chromatin sequencing), utilizes the 10X Genomics Chromium system to barcode and index individual nuclei, it provides valuable insights into the epigenomic landscape of individual cells, allowing researchers to study cellular heterogeneity and regulatory processes at high resolution.

Here's a general overview of the protocol:

<u>Cell Isolation and Nuclei Extraction</u>: Start by isolating the cells of interest. This could be from tissues or cell cultures (this step must be taken by Users). Extract the nuclei from the cells, as the chromatin accessibility information is primarily derived from the open regions in the nuclei.

<u>Transposase Reaction</u>: Add a transposase enzyme to the nuclei suspension. The transposase simultaneously fragments the chromatin and adds sequencing adapters. The sequencing adapters are important for later steps in library preparation.

<u>Chromium Gel Bead-in-Emulsion (GEM) Formation</u>: The isolated nuclei are then encapsulated into Gel Bead-in-Emulsion (GEM) droplets using the 10X Genomics Chromium platforms. Each GEM contains a unique barcode that will be associated with the chromatin accessibility data from a specific nucleus.

Library Preparation: Following the transposase reaction and nuclei encapsulation in GEM beads, PCR amplification is performed to create the final sequencing libraries. for accurate quantification and removal of PCR duplicates during data analysis.

High-Throughput Sequencing: Sequencing is performed on the prepared libraries using high-throughput sequencing platforms. The sequencing reads contain information about the open chromatin regions in individual cells/nuclei.

Data Analysis: Use specialized bioinformatics tools and pipelines to process and analyze the sequencing data. Demultiplex the data based on cell/nucleus barcodes, align the reads to the reference genome, and identify accessible chromatin regions in each cell (data analysis is the responsibility of Users).

Library preparation protocol:

Libraries will be prepared by following the protocol:

CG000496_Chromium_NextGEM_SingleCell_ATAC_ReagentKits_v2_UserGuide_RevB.pdf (10xgenomics.com)

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating reads length as required by the Single-cell ATAC sequencing protocol (10x Genomics).

On average 25.000 reads per cell/nucleus will be generated for sequencing ATACseq libraries obtained with the Single-cell multiome ATAC + Gene expression protocol.

Within this call the NF for Genomics will provide:

Single-cell ATAC sequencing for studies on diseases (humans, animal models, organoids) for a maximum number of 16 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 16 samples will be accepted.

Single-cell ATAC sequencing for differentiation studies (humans, animal models, organoids) for a maximum number of 16 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 16 samples will be accepted.

Technical requirements: Frozen cells should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). Provide an input of minimum 500k frozen cells. Cells should be prepared following the protocol:

CG00053_Handbook_CellPreparation_SingleCellProtocols_Rev_D.pdf (10xgenomics.com)

Cell viability before freezing should be ≥80%.

Nuclei will be prepared by NF staff following the protocol:

CG000365_DemonstratedProtocol_NucleiIsolation_ATAC_GEX_Sequencing_RevC.pdf (10xgenomics.com)

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- ▪ FASTQ files

- ▪ QC report

- ▪ Mapping metrics

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-015/016 – Visium Spatial gene expression from Fresh-Frozen or FFPE tissues (10X Genomics)

Services description:

The Visium Spatial Gene Expression solution (10X Genomics) allows the spatial profiling of gene expression within intact tissue sections. The protocol allows for the analysis of gene expression while preserving the spatial context of cells within a tissue sample.

Here is a general overview of the Visium Spatial Gene Expression protocol for Fresh-Frozen (FF) or Formalin Fixed Paraffin Embedded (FFPE) tissues:

Fresh-Frozen Tissues:

Tissue Sectioning: Obtain a fresh-frozen tissue sample included in OCT. Section the tissue into thin slices, typically in the range of 10-20 µm, using a cryostat (this step must be taken by the Users).

Tissue Mounting: Place the tissue sections onto a Visium Spatial Gene Expression slide (this step must be taken by the Users, Visium slides will be provided by the NF). Ensure proper mounting and orientation to maintain the spatial information.

Tissue Fixation: Fix the tissue sections on the slide to preserve the spatial structure and prevent RNA degradation.

Permeabilization: Permeabilize the fixed tissue sections to allow the probes to access the cellular RNA.

Reverse Transcription: Capture the spatially barcoded mRNA using reverse transcription, converting RNA into cDNA.

Library Preparation: Amplify and tag the cDNA with sample-specific barcodes during library preparation. Add sequencing adapters for subsequent high-throughput sequencing.

High-Throughput Sequencing: Sequence the prepared libraries using high-throughput NGS platforms.

Data Analysis: Use specialized bioinformatics tools to analyze the sequencing data. Align reads to the reference genome and quantify gene expression while retaining spatial information.

Spatial Mapping: Map the gene expression data back to the spatial locations on the tissue section using the spatial barcodes and dedicates softwares. (data analysis is the responsibility of Users)

FFPE (Formalin-Fixed Paraffin-Embedded) Tissues:

Tissue Sectioning: Obtain an FFPE tissue sample. Section the tissue into thin slices, typically around 10 μm, using a microtome (this step must be taken by the Users).

Tissue Mounting: Place the tissue sections onto a histological slide (this step must be taken by the Users).

Tissue Mounting with Visium Cytassist: Place the tissue sections onto a Visium Spatial Gene Expression slide using the CytAssist instrument.

Tissue Deparaffinization: Remove the paraffin from the FFPE tissue sections using a deparaffinization process. This step is crucial to expose the RNA for downstream processing.

Rehydration: Rehydrate the tissue sections to restore their natural state.

Antigen Retrieval: Perform antigen retrieval to improve the accessibility of RNA for subsequent steps. This step is critical for FFPE tissues, where formalin fixation can crosslink and modify nucleic acids.

Library Preparation: Amplify and tag the spatially barcoded oligonucleotide probes that have hybridized the mRNAs.

High-Throughput Sequencing: Sequence the prepared libraries using high-throughput NGS platforms.

Data Analysis: Utilize specialized bioinformatics tools to analyze the sequencing data. Align reads to the probe panel and quantify gene expression while retaining spatial information.

Spatial Mapping: Map the gene expression data back to the spatial locations on the tissue section using the spatial barcodes. (data analysis is the responsibility of Users)

Library preparation protocol:

Histological slides containing FFPE tissues will be processed by using the Visium CytAssist instrument, both libraries from FF and FFPE tissue will be prepared following the protocol:

CG000495_VisiumCytAssist_GeneExpressionUserGuide_RevE (10xgenomics.com)

Libraries sequencing and NGS coverage:

Libraries will be sequenced using either the NextSeq 2000 system or the NovaSeq 6000 system (Illumina) by generating reads length as required by the Visium Spatial gene expression from FF or FFPE tissues protocol (10x Genomics).

On average 50.000 reads per spot will be generated for sequencing libraries obtained with the Visium Spatial gene expression from FF tissue protocol and 25.000 reads per spot will be generated for sequencing libraries obtained with the Visium Spatial gene expression from FFPE tissue protocol.

Within this call the NF for Genomics will provide:

Visium Spatial gene expression from FreshFrozen or FFPE tissues for studies on diseases (humans, animal models, organoids) for a maximum number of 24 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 24 samples will be accepted.

<u>Visium Spatial gene expression from FreshFrozen or FFPE tissues for differentiation studies (humans, animal models, organoids) </u>for a maximum number of 24 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 24 samples will be accepted.

Technical requirements:

For Fresh Frozen Visium: Tissues, previously frozen and embedded in OCT according to the protocol:

CG000240_Demonstrated_Protocol_VisiumSpatialProtocols_TissuePreparationGuide_RevE.pdf (10xgenomics.com)

should be placed by the User on the Visium slides/capture areas, that will be provided by the NF for Genomics. For correctly placing the tissue slices on the Visium slides/areas and storing them please follow the same protocol (section 2.3). For each sample/tissue the quality of RNA extracted from the tissue should be evaluated following the "RNA quality assessment section" of the same protocol and by using the Tape Station/Bioanalyzer. Only samples having a RIN≥7 will be processed.

For FFPE Visium: Tissues should be placed by the User on Fisherbrand superfrost plus microscope slides according to the protocol:

CG000518_Demonstrated_Protocol_VisiumCytAssistSpatialProtocolsFFPE_TissuePreparationGuide_RevC.pdf (10xgenomics.com)

For each sample/tissue the quality of RNA extracted from the tissue should be evaluated following the protocol

CG000408_Demonstrated_Protocol_VisiumSpatialProtocolsFFPE_TissuePreparationGuide_RevE.pdf (10xgenomics.com)

The quality of RNAs should be evaluated by using the Agilent Tape Station/Bioanalyzer. Only samples having a DV200≥30 will be processed.

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report
- Mapping metrics

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-019 – Nanopore gDNA sequencing (long reads or ultra long reads)

Services description:

Nanopore sequencing is a next-generation sequencing technology that uses nanopores to directly sequence DNA molecules and to analyse DNA bases modifications. This method is known for producing long reads, and in some cases, ultra-long reads, making it valuable for various genomic applications.

Here's a general overview of the Nanopore gDNA (genomic DNA) sequencing protocol for long reads or ultra-long reads:

DNA Extraction: Start with the extraction of high-quality genomic DNA from the biological sample of interest. Different DNA extraction methods can be used based on the sample type.

Library Preparation steps:

Fragmentation: Fragment the genomic DNA to the desired size. For long reads, the fragmentation is typically minimal, while for ultra-long reads, larger fragments may be preserved so no fragmentation will be performed.

Repair Ends: Repair the DNA ends to ensure they have a consistent structure suitable for sequencing.

Adapter Ligation: Attach sequencing adapters to the DNA fragments. These adapters contain nanopore-specific sequences that facilitate the capture and reading of the DNA sequence.

Loading the Sequencing Device: Prepare the nanopore sequencing device according to the manufacturer's instructions. This may involve priming the flow cell, loading the prepared library, and initiating the sequencing run.

Sequencing: In nanopore sequencing, a single DNA strand passes through a nanopore, and as the DNA moves through the pore, the changes in electrical current are measured. The sequencing instrument records these changes in current, allowing for the identification of individual bases in the DNA sequence.

Data Collection: During the sequencing run, data is continuously collected in the form of raw electrical signals. The raw signals are then base-called to convert the electrical signal data into the corresponding DNA sequence.

Data Analysis: Process the base-called data through bioinformatics pipelines to correct errors, filter out low-quality reads, and assemble the long or ultra-long reads (data analysis is the responsibility of Users).

Library preparation protocol:

Libraries will be prepared by following the protocols:

Ligation Sequencing Kit V14 (SQK-LSK114) for long reads

[Ligation sequencing DNA V14 (SQK-LSK114)](#)

DNA ends are repaired and dA-tailed using the NEBNext End Repair/dA-tailing module before the sequencing adapters, supplied in the kit, are ligated onto the prepared ends. The kit is optimised to achieve sequencing accuracies of over 99% (Q20+) with high output on the latest nanopore Flowcells R10.4.1. The Ligation Sequencing Kit V14 is compatible with upstream processes such as target enrichment by sequence capture, whole genome amplification, and size selection (for enrichment of specified fragment lengths). PCR- and WGA-free workflows remove amplification bias and retain base modification information, which can be analysed using bioinformatic tools supported by Oxford Nanopore.

Ultra-Long DNA Sequencing Kit V14 (SQK-ULK114) for ultra long reads

[Ultra-Long DNA Sequencing Kit V14 (SQK-ULK114)](#)

The Ultra-Long DNA Sequencing Kit V14 offers a means of preparing ultra-high molecular weight (uHMW) DNA for sequencing, which has shown to give N50s >50 kb and reads up to 4+ Mb. The kit is based on transposase chemistry: the transposase simultaneously cleaves template molecules and attaches tags to the cleaved ends. Rapid sequencing adapters are then added to the tagged ends. The last step is an overnight elution of the DNA library. This kit has been updated to use the newest Kit 14 chemistry which includes improved modal raw read sequencing accuracies with higher output on the latest nanopore Flowcells R10.4.1.

Libraries sequencing and NGS coverage:

For nanopore gDNA sequencing applied to Human studies a coverage of 20X per sample will be obtained.

For nanopore gDNA sequencing applied to Animal studies the coverage per sample will be determined considering the animal genome size.

For nanopore gDNA sequencing applied to Plant studies the coverage per sample will be determined considering the plant genome size and ploidy level.

Within this call the NF for Genomics will provide:

nanopore gDNA sequencing (for Human samples) for a maximum number of 32 samples. Projects with a sample size ranging from a minimum of 8 to a maximum of 32 samples will be accepted.

nanopore gDNA sequencing (for Plants samples) for a maximum number of 12 samples. Projects with a sample size ranging from a minimum of 5 to a maximum of 12 samples will be accepted.

nanopore gDNA sequencing (for Animals samples) for a maximum number of 10 samples. Projects with a sample size ranging from a minimum of 5 to a maximum of 10 samples will be accepted.

Technical requirements:

For long-reads generation: gDNA Samples should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). At least 1 µg in 50 µl of gDNA should be provided (DNA quantity evaluated by fluorimeter; i.e. Qubit dsDNA BR Assay Kit); gDNA should be of high molecular weight (DIN≥6 obtained from Tape Station Agilent/Bioanalyzer), and clear of contaminant. Purity of the gDNA should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). gDNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax).

For ultra long-reads generation: At least 3 pellets of 6 million each of frozen cells per sample should be provided. Cell pellets should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- Pod5 and FASTQ files
- QC report

▪ Mapping metrics (if reference genome/transcriptome is available)

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-020 – Nanopore small gDNA sequencing (long reads)

Services description:

Nanopore sequencing of small gDNAs will be performed with the native barcoding kit that refers to a set of reagents and protocols designed to enable the simultaneous sequencing of multiple samples by adding unique barcodes to each sample before sequencing. This approach is particularly useful for studying small bacteria genomes (gDNA) as it allows for high-throughput sequencing and analysis of multiple samples using a single flowcell.

This approach provides a cost-effective and efficient way to sequence multiple small bacteria genomes simultaneously, making it a valuable tool in microbiome studies, environmental monitoring, and other applications in microbial genomics.

Here is a step-by-step description of the process:

DNA Extraction: Begin with the extraction of genomic DNA (gDNA) from small bacteria samples using standard DNA extraction methods (this step must be taken by Users).

Library Preparation: The native barcoding kit includes reagents to tag each gDNA sample with a unique barcode. This is typically done by ligating barcode adapters to the gDNA fragments.

Barcoding: Introduce sample-specific barcodes during library preparation. These barcodes serve as unique molecular identifiers, allowing for the identification and demultiplexing of the sequences from different samples in the subsequent analysis.

Sequencing: Load the barcoded libraries onto a nanopore sequencing platform and Flowcell.

Sequencing: The nanopore sequencer reads individual DNA strands as they pass through a nanopore, generating long reads with real-time sequencing information.

Base Calling and Analysis: The raw electrical signals generated by the nanopore sequencer are translated into DNA base calls through base-calling algorithms.

Data Analysis: Use bioinformatics tools to analyse the sequencing data, including aligning the reads to a reference genome, variant calling, and identifying unique features of the small bacteria genomes (data analysis is the responsibility of Users).

Library preparation protocol:

Libraries will be prepared by following the protocol:

Ligation sequencing gDNA - Native Barcoding Kit 96 V14 (SQK-NBD114.96)

The Native Barcoding Kit 96 V14 is a kit providing 96 unique barcodes to enable PCR-free multiplexing of dsDNA samples such as gDNA and amplicons. The library preparation method is similar to the Ligation Sequencing Kit protocol; gDNA is repaired and dA-tailed using the NEBNext End Repair/dA-tailing module, and then a unique dT-tailed barcode adapter is ligated on the dA-tailed template. Barcoded samples are then pooled together. Each barcode

adapter also has a cohesive end and this is used as a hook to ligate to the supplied sequencing adapter. The kit is optimised to achieve modal sequencing accuracies of over 99% (Q20+) with high output on the latest nanopore Flowcells R10.4.1.

Libraries sequencing and NGS coverage:

For nanopore smallgDNA sequencing applied to Bacteria genomes the coverage per sample will be determined considering the Bacteria genome size and the number of bacteria genomes loaded in a single flowcell.

Within this call the NF for Genomics will provide:

nanopore small gDNA seq (for Bacteria samples) for a maximum of 24 batches/pools of samples considering that every batch/pool of samples can be composed of a maximum of 96 samples (max batch 96).

Technical requirements: gDNA Samples should be provided in Low Bind full skirted PCR plates (i.e. Eppendorf twin.tec PCR plates). At least 1 µg in 15 µl of gDNA should be provided; gDNA should be of high molecular weight (DIN≥6 obtained from Agilent Tape Station/Bionalyzer), and clear of contaminant. Purity of the gDNA should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). gDNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- Pod5 and FASTQ files
- QC report
- Mapping metrics (if reference genome/transcriptome is available)

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-021 – Nanopore Direct RNA Sequencing

Services description:

Direct RNA Sequencing with Nanopore technology is a cutting-edge method for sequencing RNA molecules without the need for conversion to complementary DNA (cDNA) as required in traditional RNA sequencing methods. The protocol involves the direct sequencing of RNA strands through nanopores, allowing for the real-time detection of RNA sequences and RNA bases modifications. Direct RNA Sequencing with Nanopore technology offers the advantage of studying RNA molecules in their native state, providing valuable insights into RNA processing, alternative splicing, and modifications without the need for reverse transcription. It is particularly valuable for capturing the full complexity of the transcriptome.

Here's a general overview of the Direct RNA Sequencing protocol using Nanopore technology:

RNA Extraction: Begin by isolating high-quality RNA from the biological sample of interest. This could be total RNA or specific RNA fractions, depending on the experimental goals (this step must be taken by Users).

Library Preparation steps:

Adapter Ligation: Attach sequencing adapters to the RNA fragments. These adapters contain nanopore-specific sequences necessary for capturing and sequencing the RNA.

Loading the Sequencing Device: Prime the nanopore sequencing device according to the manufacturer's instructions. Load the prepared RNA library onto the flow cell, initiating the sequencing run.

Sequencing: As RNA strands pass through the nanopore, changes in electrical current are measured in real-time. The changes in current are characteristic of the specific RNA bases passing through the nanopore, allowing for base-calling and the reconstruction of the RNA sequence.

Data Collection: Raw electrical signals are continuously recorded during the sequencing run. Base-calling algorithms are applied to convert the electrical signals into the corresponding RNA sequence.

Data Analysis: Process the base-called data through bioinformatics pipelines to correct errors and filter out low-quality reads. Align the reads to a reference genome or transcriptome, if available, to identify known transcripts (data analysis is the responsibility of Users).

Library preparation protocol:

Libraries will be prepared by following the protocol:

 Direct RNA sequencing (SQK-RNA004)

The Direct RNA Sequencing Kit (SQK-RNA004) is used to prepare and sequence native RNA without conversion to cDNA. Inputs include poly(A)-tailed RNA or total RNA, such as eukaryotic mRNA and viral RNA.

This kit upgrade includes increased sequencing output and improved modal raw read accuracy on the new RNA flow cell (FLO-MIN004RA and FLO-PRO004RA).

Libraries sequencing and NGS coverage:

For nanopore Direct RNA Sequencing applied to Human samples each sample will be sequenced in one PromethION Flowcell.

Within this call the NF for Genomics will provide:

nanopore Direct RNA Sequencing (for Human samples) for a maximum number of 20 samples. Projects with a sample size ranging from a minimum of 10 to a maximum of 20 samples will be accepted.

> **Technical requirements:** RNA Samples should be provided in low bind tubes (i.e. 1,5 ml Eppendorf tubes). Three aliquots containing 1 µg of total RNA in 10ul per sample should be provided by the User. Total RNA should be DNase treated and the RIN≥8 (quality of RNAs evaluated by Agilent Tape Station/Bioanalyzer). Purity of the total RNAs should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). Total RNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax).

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- Pod5 and FASTQ files

- QC report

- Mapping metrics (if reference genome/transcriptome is available)

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-022 – Nanopore cDNA sequencing (bulk cDNA or single-cell cDNA from 10x Genomics protocol) (Human-Mouse)

Services description:

Nanopore cDNA sequencing is a powerful technique that combines the benefits of nanopore sequencing technology with the study of complementary DNA (cDNA), which represents the transcribed RNA in a biological sample. This method allows researchers to investigate gene expression, alternative splicing, and other aspects of RNA biology with long-read sequencing capabilities. Nanopore cDNA sequencing offers several advantages, including the ability to generate long reads that span entire transcripts. This makes it particularly valuable for studying complex transcriptomes, characterizing novel isoforms, and exploring the dynamics of gene expression in various biological contexts.

Here's a step-by-step description of the process:

RNA Extraction: Begin by extracting total RNA from the biological sample of interest. This can be done using standard RNA extraction protocols, ensuring the preservation of RNA integrity (this step must be taken by Users).

cDNA Synthesis: Reverse transcribe the RNA into complementary DNA (cDNA) using reverse transcription enzymes and primers. This step converts the RNA into a stable form of DNA that can be sequenced (this step can be taken by Users or asked to the NF).

Library Preparation: Prepare the cDNA for sequencing by adding sequencing adapters to the ends. This usually involves ligating adapters with unique barcodes to distinguish different samples or conditions.

Sequencing: Load the prepared cDNA libraries onto a nanopore sequencing ONT flowcell and platform. The nanopore sequencer reads individual cDNA strands as they pass through a nanopore, producing long-read sequences with real-time sequencing information.

Base Calling and Analysis: Translate the raw electrical signals generated by the nanopore sequencer into DNA base calls using base-calling algorithms.

Use bioinformatics tools to align the cDNA reads to a reference genome or transcriptome, identify gene isoforms, quantify gene expression levels, and analyze alternative splicing events (data analysis is the responsibility of Users).

Library preparation protocol:

For bulk cDNA sequencing libraries will be prepared by following the protocol:

cDNA-PCR Sequencing V14 (SQK-PCS114)

The PCR-cDNA Sequencing Kit is used to prepare cDNA for nanopore sequencing from an input of as low as 4 ng poly(A)+ RNA. When poly(A)+ enriched RNA is not available it is possible to use 200 ng of total RNA but additional optimisation may be required.

The protocol uses a strand switching method to select for full length transcripts, allowing the identification of splice variants, with the incorporation of unique molecular identifiers (UMI) during this step. Taking full-length poly(A)+ RNA, complementary strand synthesis and strand switching are performed using kit-supplied oligonucleotides. Double-stranded cDNA is then generated by PCR amplification using primers that contain 5' tags which facilitate the ligase-free attachment of Rapid Sequencing Adapters.

The PCR-cDNA Sequencing Kit also includes a new cDNA RT adapter and RT primer to prime cDNA synthesis from the end of a transcript to reduce overlaps during the reverse transcription step and to allow Users to measure polyA+ tail lengths.

For single-cell cDNA from 10x Genomics prep libraries will be prepared by following the protocol:

[cDNA-PCR Sequencing V14 (SQK-PCS114)](#)

A primer pair compatible with the PCR-cDNA Sequencing Kit and with a biotin on the primer annealing to the 3' end of the cDNA, is used to amplify 10 ng of cDNA generated with 10X Genomics Gene Expression protocols. The amplified cDNA is captured with streptavidin beads to enrich for full-length cDNA and remove RT artifact due to priming internal on the RNA. The captured cDNA is enriched by PCR amplification using primers that contain 5' tags which facilitate the ligase-free attachment of Rapid Adapter T (RAP T).

Libraries sequencing and NGS coverage:

For nanopore bulk cDNA or single-cell cDNA sequencing (from 10x Genomics protocols) applied to Human/mouse samples each sample will be sequenced in one PromethION Flowcell.

Within this call the NF for Genomics will provide:

nanopore bulk cDNA or single-cell cDNA Sequencing (for Human samples) for a maximum number of 20 samples. Projects with a sample size ranging from a minimum of 10 to a maximum of 20 samples will be accepted.

> **Technical requirements:**
>
> **For bulk cDNA sequencing**, provide either:
>
> - three aliquots containing 500 ng total RNA in 10ul or
>
> - three aliquots containing 200ng cDNA in 50ul.
>
> Total RNA should be DNase treated and the RIN≥8 (quality of RNAs evaluated by Agilent Tape Station/Bioanalyzer). Purity of the total RNAs should be assessed with a Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). Both total RNA and cDNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax).
>
> **For single-cell cDNA from 10x Genomics prep**, two aliquots containing 10 ng of cDNA in 25ul should be provided by the User. The cDNA should be generated by using either the
>
> CG000315CG000315_ChromiumNextGEMSingleCell3-GeneExpression_v3.1(DualIndex)_RevE.pdf
>
> Or
>
> CG000331_ChromiumNextGEMSingleCell5-v2_UserGuide_RevE.pdf
>
> cDNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax) and their quality evaluated by Agilent Tape Station/Bioanalyzer.

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- Pod5 and FASTQ files

- QC report

- Mapping metrics (if reference genome/transcriptome is available)

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-023 – Nanopore cell-free DNA sequencing (Human)

Services description:

Cell-free DNA (cfDNA) sequencing with Nanopore technology represents a revolutionary approach to interrogate the genetic information present in circulating DNA, often extracted from blood plasma. Unlike traditional sequencing methods, this protocol allows for the direct sequencing of cfDNA without the need for intermediate steps such as PCR amplification or conversion and allows the analysis in real time of the methylation status of circulating cfDNA. Its applications extend to clinical diagnostics, providing valuable information for personalized medicine and disease monitoring such as liquid biopsy for cancer detection, monitoring treatment response, and identifying minimal residual disease.

Here's a general overview of the Cell-free DNA Sequencing with Nanopore Protocol:

cfDNA Extraction: Begin by isolating cell-free DNA from a biological fluid, such as blood plasma, using optimized extraction methods (this step must be taken by Users).

Library Preparation: preparing the library for nanopore sequencing ligating sequencing adapters.

Loading the Sequencing Device: Load the prepared cfDNA library onto the Nanopore sequencing device to initiate the sequencing run.

Sequencing: As cfDNA molecules pass through the nanopore, real-time changes in electrical current enable the identification of DNA sequences and methylation status.

Data Collection: Record raw electrical signals continuously during the sequencing run, followed by super accurate base-calling to convert signals into DNA sequences.

Data Analysis: Process and analyze base-called data, aligning sequences to a reference genome or performing de novo assembly.

Variant Calling and Copy Number Analysis: Identify genetic variants and conduct copy number analysis to detect amplifications or deletions in cfDNA-covered genomic regions or evaluate the methylations status of the cfDNA fragments (data analysis is the responsibility of Users).

Library preparation protocol:

Libraries will be prepared by following the protocol:

[Human blood cell-free DNA (cfDNA)](#)

The library preparation method steps are: DNA ends are repaired and dA-tailed using the NEBNext End Repair/dA-tailing module before the sequencing adapters, supplied in the kit, are ligated onto the prepared ends. The kit is optimised to achieve sequencing accuracies of over 99% (Q20+) with high output on the latest nanopore Flowcells R10.4.1. The protocol uses a modified version of the long-reads protocol optimised to recover short DNA fragments.

Libraries sequencing and NGS coverage:

For nanopore cell-free DNA sequencing applied to Human samples each sample will be sequenced in one PromethION Flowcell.

Within this call the NF for Genomics will provide:

Nanopore cell-free DNA sequencing (Human) for a maximum number of 60 samples. Projects with a sample size ranging from a minimum of 20 to a maximum of 60 samples will be accepted.

Technical requirements: cfDNA samples should be provided in Low Bind full skirted PCR plates (i.e. Eppendorf twin.tec PCR plates). Provide an input of at least 30ng in 50 µl. DNA should be clear of contaminant. Purity of the DNA should be assessed with a Nanodrop 2000 Spectrophotometer (i.e. Nanodrop; 260/280≥1,8 and 260/230≥1,8). cfDNA samples should be quantified by using of a fluorometer (i.e Qubit/ Glomax) and their quality evaluated by Agilent Tape Station/Bioanalyzer

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- Pod5 and FASTQ files

- QC report

- Mapping metrics (if reference genome/transcriptome is available)

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-024 – Sequencing only with NovaSeq 6000 (Illumina)

Services description:

The NF for Genomics will provide sequencing of pools of libraries prepared by the Users with the NovaSeq6000 sequencing platform (Illumina).

The NovaSeq 6000 is a high-throughput sequencing platform developed by Illumina. It's designed to handle a wide range of applications and large-scale genomic projects.

NovaSeq 6000 is commonly used for: Whole Genome Sequencing (WGS), Exome Sequencing (WES), Transcriptome Analysis (bulk RNA-Seq or single-cell RNAseq), Metagenomics, Epigenomics studies, Population Genomics.

Libraries sequencing and NGS coverage:

For Sequencing only with NovaSeq 6000 (Illumina) the total amount of reads that will be generated per sample or the coverage per genome will be determined depending on the projects' needs and on the number of samples that will fit within the number of S4 Flowcells required.

Within this call the NF for Genomics will provide:

Sequencing only with NovaSeq 6000 (Illumina) for a total of 6 S4 Flowcells 300 cycles (generating 150bp PE reads) and for a total of 6 S4 Flowcells 200 cycles (generating 100bp PE reads). Projects that will require from a minimum of 1 to a maximum of 6 S4 Flowcells RUNs will be accepted.

Technical requirements: Libraries should be provided already pooled in a Low Bind tube (i.e. 1,5 ml Eppendorf tubes). The Pool molarity should be at least 5nM in a volume of 200ul in either nuclease-free water, Illumina RSB buffer, or 10mM Tris HCl pH 8.5. The pool should be quantified by a fluorometer (i.e Qubit/ Glomax) and its quality evaluated by Agilent Tape Station/Bioanalyzer. Only unique dual indexed libraries will be accepted.

At the following link a tool for calculating how to prepare the pool of libraries for sequencing is available:

Link to Pooling Calculator

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files

- QC report

Access modality available: Access to NF services (data analysis is the responsibility of Users).

SID: G-025 – Sequencing only with NextSeq 2000 (Illumina)

Services description:

The NF for Genomics will provide sequencing of pools of libraries prepared by the Users with the NextSeq2000 sequencing platform (Illumina).

The NextSeq 2000 is a next-generation sequencing platform developed by Illumina. It is designed to offer high-throughput sequencing with flexibility for various applications.

The NextSeq 2000 is commonly used for: Whole Genome Sequencing (WGS), RNA Sequencing (RNA-Seq and single-cell RNA-seq), Targeted Sequencing (amplicon sequencing), Exome Sequencing (WES), Metagenomics, ChIP-Seq and Epigenomics.

Libraries sequencing and NGS coverage:

For Sequencing only with NextSeq 2000 (Illumina) the total amount of reads that will be generated per sample or the coverage per genome will be determined depending on the projects' needs and on the number of samples that will fit within the number of P3 Flowcells required.

Within this call the NF for Genomics will provide:

Sequencing only with NextSeq 2000 (Illumina) for a total of 10 P3 Flowcells 300 cycles (generating 150bp PE reads), for a total of 10 P3 Flowcells 200 cycles (generating 100bp PE reads) and for a total of 10 P3 Flowcells 600 cycles (generating 300bp PE reads). Projects that will require from a minimum of 5 to a maximum of 10 P3 Flowcells RUNs will be accepted.

Technical requirements:

Libraries should be provided already pooled in a Low Bind tube (i.e. 1,5 ml Eppendorf tubes). The Pool molarity should be at least 5nM in a volume of 60ul in either nuclease-free water, Illumina RSB buffer, or 10mM Tris HCl pH 8.5. The pool should be quantified by a fluorometer (i.e Qubit/ Glomax) and its quality evaluated by Agilent Tape Station/Bioanalyzer. Only dual indexed libraries will be accepted.

At the following link a tool for calculating how to prepare the pool of libraries for sequencing is available:

Link to Pooling Calculator

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report

Access modality available: Access to NF services (data analysis is the responsibility of Users).

Services description:

The NF for Genomics will provide sequencing of pools of libraries prepared by the Users with the MiSeq sequencing platform (Illumina).

the MiSeq is a benchtop next-generation sequencing platform developed by Illumina. It's designed for smaller-scale sequencing projects and offers flexibility for various applications.

MiSeq is commonly used for: Targeted Sequencing (amplicon sequencing), Small Genomes Sequencing, Metagenomics, 16S rRNA Sequencing, Small RNA-Seq, Viral genomes Sequencing.

Libraries sequencing and NGS coverage:

For Sequencing only with MiSeq (Illumina) the total amount of reads that will be generated per sample or the coverage per genome will be determined depending on the projects' needs and on the number of samples that will fit within the number of Flowcells required.

Within this call the NF for Genomics will provide:

Sequencing only with MiSeq (Illumina) for a total of 30 MiSeq v3 (600 cycle) Flowcells (generating 300bp PE reads) Projects that will require from a minimum of 10 to a maximum of 30 MiSeq v3 (600 cycle) Flowcells RUNs will be accepted.

Technical requirements: Libraries should be provided already pooled in a Low Bind tube (i.e. 1,5 ml Eppendorf tubes). The Pool molarity should be at least 5nM in a volume of 60ul in either nuclease-free water, Illumina RSB buffer, or 10mM Tris HCl pH 8.5. The pool should be quantified by a fluorometer (i.e Qubit/ Glomax) and its quality evaluated by Agilent Tape Station/Bioanalyzer. Only dual indexed libraries will be accepted.

At the following link a tool for calculating how to prepare the pool of libraries for sequencing is available:

Link to Pooling Calculator

Results that will be delivered to the Users:

The NF for Genomics will deliver to the Users the following files for every sample sequenced:

- FASTQ files
- QC report

Access modality available: Access to NF services (data analysis is the responsibility of Users).

Appendix I: Summary of technical requirements

Service ID (SID)	Tube/Support type	N° of aliquots / sample	Min. Mass (ng)	Min. Conc. (ng/µl)	Min. Conc. (nM)	Min. Vol. (µl)	Max. Vol. (µl)	Min. DIN/RIN	Min. 260/280	Min. 260/280	DNAse treat.	Min. N° of Cells (M)	Min. Cell viability (%)
G-001 – Whole Genome Sequencing (WGS)	Low-bind full-skirted PCR plates			15		50		6	1.8	1.8			
G-002 – Whole Exome Sequencing (WES)	Low-bind full-skirted PCR plates			5		40		6	1.8	1.8			
G-003 – Amplicon sequencing for microbiome analysis (16S-ITS)	Low-bind full-skirted PCR plates			0.1		20		6	1.8	1.8			
G-005 - mRNA sequencing from standard input	Low-bind full-skirted PCR plates		500			20	50	7	1.8	1.8	x		
G-005 - mRNA sequencing from low input	Low-bind full-skirted PCR plates		0.05			15	20	7	1.8	1.8	x		
G-006 - totalRNA sequencing from standard input	Low-bind full-skirted PCR plates		500			20	50	4	1.8	1.8	x		
G-007 - smallRNA sequencing	Low-bind full-skirted PCR plates		100			20	50	7	1.8	1.8	x		
G-008/G-012 - Single-cell 3' RNA sequencing (10X Genomics)	Low-bind 1.5 ml tubes											0.5	80
G-008/G-012 - Single-cell 3' RNA sequencing with CellPlex (10X Genomics)	Low-bind 1.5 ml tubes											2	80
G-008/G-012 - Single-cell gene Expression Flex (10X Genomics)	Low-bind 1.5 ml tubes	2										0.5	80
G-009/012/014 - Single-cell Immune profiling-V(D)J (10X Genomics)	Low-bind 1.5 ml tubes											0.5	80
G-010/012/013 – Single-cell multiome ATAC + Gene expression (10X Genomics)	Low-bind 1.5 ml tubes											0.5	80
G-011/013 - Single-cell ATAC sequencing (10X Genomics)	Low-bind 1.5 ml tubes											0.5	80
G-015/016 – Visium Spatial gene expression from Fresh-Frozen tissues (10X Genomics)	Visium slides							7					
G-015/016 – Visium Spatial gene expression from FFPE tissues (10X Genomics)	Fisherbrand superfrost plus slides												
G-019 - Nanopore gDNA sequencing (long reads)	Low-bind 1.5 ml tubes		1000				50	6	1.8	1.8			
G-019 - Nanopore gDNA sequencing (ultra long reads)	Low-bind 1.5 ml tubes	3										6	
G-020 - Nanopore small gDNA sequencing (long reads)	Low-bind full-skirted PCR plates		1000				15	6	1.8	1.8			
G-021 - Nanopore Direct RNA Sequencing	Low-bind 1.5 ml tubes	3	1000				10	8	1.8	1.8	x		
G-022 - Nanopore cDNA sequencing (bulk cDNA) from RNA	Low-bind full-skirted PCR plates	3	500				10	8	1.8	1.8	x		
G-022 - Nanopore cDNA sequencing (bulk cDNA) from cDNA	Low-bind full-skirted PCR plates	3	200				50		1.8	1.8			
G-022 - Nanopore cDNA sequencing (single-cell cDNA from 10x Genomics protocol) (Human-Mouse)	Low-bind full-skirted PCR plates	2	10				25		1.8	1.8			
G-023 - Nanopore cell-free DNA sequencing (Human)	Low-bind full-skirted PCR plates		30				50		1.8	1.8			
G-024 - Sequencing only with NovaSeq 6000 (Illumina)	Low-bind 1.5 ml tubes				5		200						
G-025 - Sequencing only with NextSeq 2000 (Illumina)	Low-bind 1.5 ml tubes				5		60						
G-026 - Sequencing only with MiSeq (Illumina)	Low-bind 1.5 ml tubes				5		60						

ANNEX II: LETTER OF INSTITUTIONAL ENDORSEMENT TEMPLATE

(Print on paper bearing the official letterhead of the host Institution)

Endorsement letter of the host Institution

To whom it may concern:

I, the undersigned, ………….. (*name of legal representative or special attorney*), born in ………… (*city*) on …………..(*date*), as legal representative *(or special attorney, by means of special power of attorney identified by …………………………..)* and on behalf of ………………(*name of the host Institution*), legal residence in *(referred to the host Institution)* ……………….(*city*), address ………………….., regarding the project (*title*)…………………………………………………………….., presented by ………………………….(*Applicants's first name and surname*), as Principal Investigator on the call for Access to Human Technopole National Facilities…..(ID *of the call*),

Declare

- That the host Institution is among those eligible to participate in the call for Access as it belongs to the following eligible category: (select among University, IRCSS, Public Research Entities);

- That the Applicant, Dr ………… (*Applicant's first name and surname*) is an independent group leader (Principal Investigator) affiliated with a primary appointment at the host Institution and that they meet the eligibility criteria as indicated in the call;

- That the Applicant has not received funding for performing elsewhere, the aspects of the project for which they are seeking here support from or Access to Human Technopole National Facilities;

- That the services requested here cannot be performed by the Applicant at the host Institution, at a cost which is deemed affordable for them.

- That relevant authorisations, declarations and accreditation from the competent authority(ies) have been obtained in order to process samples and data through Human Technopole;

- That, if applicable, biological specimens have been obtained with the corresponding approval of the Bioethics Committee and appropriately signed 'informed consent', both for their collection and their use, including conservation, manipulation, derivation and processing to be carried out by Human Technopole National Facilities;

- That, if samples were obtained from subjects who signed an 'informed consent', said informed consent allows that sequencing data and results are included in secure controlled Access databases and accessed/ used by authorized third parties;

and is committed

-   To accept the terms and conditions to Access Human Technopole National Facilities as described in the National Facilities Access rules (link);
-   To sign the Access Agreement should the project be approved (link)

For the host Institution (Applicant legal entity/beneficiary):

Date …………………

Name and Title …………………. ; ………………….

Email and Signature of legal representative or delegated person

…………………. ; ………………………….