ATTACHMENT 3: CONTENT OF THE SERVICES

NEXT GENERATION SEQUENCING SERVICES for project HERCULES

We are requesting tenders for Next Generation Sequencing services for whole and targeted sequencing of DNA and RNA samples derived from fresh frozen human cancer and normal tissues and cells. Additionally, sequencing services for a pilot study on free circulating tumor DNA (ctDNA) is requested starting from plasma samples. The study is part of the ongoing HERCULES (http://www.project-hercules.eu/) EU-project coordinated by Professor Sampsa Hautaniemi (University of Helsinki) focusing on finding solutions to drug resistance in high-grade serous ovarian cancer. We have already completed sequencing of some pilot samples. The existing sequence data have been produced with Illumina HiSeq4000 (RNAseq, methylome sequencing, exome sequencing), HiSeqX10 (WGS) and Illumina MiSeq/NexSeq500 (targeted/ amplicon sequencing). We have also developed a bioinformatics infrastructure (Anduril, http://csbi.ltdk.helsinki.fi/site/#anduril) and implemented a comprehensive tool box to manage and analyze Illumina short read (single and paired end reads of 50-300 bp). The data provided have to be comparable to our existing data and compatible with our existing bioinformatics pipelines in Anduril. Because our major focus is on heterogenous cancer samples, very high data quality is needed (for example low error rate even in low frequence variants and chromosomal level aberrations).

Tenders are accepted also for subsections of the project if the tenderer is unable to provide all the required services, and separate providers can be selected for the 3 sections described below.

Section 1) RNA sequencing

RNA SAMPLES FOR STRANDED totalRNA seq (including long non-coding and coding RNAs):
- estimated **35 samples** with high amount of data (min 60 M reads per sample, PE 75 bp, or other setting yielding similar coverage)
- estimated **12 samples** with low amount of data (30 M reads, PE75 bp, or other setting yielding similar coverage)

We deliver total RNA in original concentration extracted from human cancer or normal tissue or cells with Qiagen AllPrep DNA/RNA kit. RNA is stored in 80 C degrees and delivered in dry ice.
Service provider has to run accurate RNA quantification and qualification (Bioanalyzer or comparable method) again regardless of QC run in our laboratory.

RNA SAMPLES FROM CELL LINES (characterization) using polyA-RNA quantification (for example 30 M reads, SE 75 bp)
Final **number of samples is estimated to be 30** by the end of 2016. Samples will be sent and need to be sequenced in small batches (5-15 samples at a time) optimal for one lane.

Please give 2-3 price examples for a single lane (depending on the number of samples and sequencing settings).

Section 2) DNA sequencing

DNA SAMPLES FOR WHOLE GENOME SEQUENCING
Estimated altogether **55 samples**, all extracted from fresh frozen tumor tissue or cells. DNA extracted using Qiagen AllPrep DNA/RNA kit and quantified with Qubit and or NanoDrop will be delivered in original concentration. Service provider has to run accurate DNA quantification and qualification (Qubit or comparable, sizing with gel or comparable methods) again regardless of QC run in our laboratory.

In the sequencing Phase I all the samples will be sequenced with 30x average coverage (PE 100-150 bp)

After evaluation of tumor purity, sample heterogeneity and other key parameters another round (Phase II) of sequencing will be performed for a subset of these samples (**estimated number 30-35**) to yield altogether 70x, 140x, or 200x coverage. Average coverage is estimated to be 140x.
To minimize duplication rate, for example in Illumina HiSeqX10 platform, a separate library has to be prepared at least for each 70x sequencing coverage, i.e. 70x = 1 library, 140x = 2 libraries, 200x = 3 libraries. The design should aim at duplication rate max 15% for tumor samples (70x) and max 10% for low coverage (30x) tissue samples and for cell lines.

After the round II libraries have been successfully prepared, the remaining DNA will be used for Targeted methylation sequencing, such as Agilent SureSelectMethylSeq. Min 30x coverage at target. If using Illumina HiSeq4000 or comparable platforms the service provider is responsible for optimal pooling with normal DNA samples to avoid known biases caused by unusual base distribution in sequencing of bisulphite converted DNA. The estimated number **of samples for methylome sequencing is 40**.

Section3) ctDNA sequencing

ctDNA DEEP SEQUENCING
We have started collecting plasma samples for a ctDNA pilot. **Piloting with 3 samples** will be performed. Service provider should offer:
- DNA extraction from plasma (freshly repaired for ctDNA sequencing).
- Targeting assay which includes genes and gene regions commonly mutated in ovarian cancer (particularly in High Grade Serous Ovarian Cancer).
- Ultra-deep sequencing with clinically applicable data quality.

Please, give **prices separately for the pilot and for the optional larger study (20-100 samples**) (fill in attachment 1).

Steps in the project
1. Agreement with adjusted sample numbers
2. Sending of samples
3. Sample quality and quantity analysis by the service provider
4. Reporting of QC results
5. For samples passing QC evaluation the library preparation will be started for Phase I. If possible we send new samples for those not passing QC and not accepted by us.
6. Library QC and reporting of possible issues
7. Redoing of low quality libraries & QC

8. Sequencing (Phase I, RNA and WGS with 30x)
9. Evaluation of data yield and quality by the service provider. If the data fulfills the agreed requirements it will be delivered in a password protected hard drive or using a fast secure internet connection (if feasible considering the amount of data) (Phase I)
10. We run preliminary analysis for the Phase I data, including, overall quality, mappability, duplication rate, estimation of variant calling error rate, copy number status, tumor purity and tumor heterogeneity.
11. We decide for which samples and how much additional data is needed in Phase II sequencing.
12. Repeat steps 5-9 for Phase II
10. After we have accepted the data the final invoice can be activated (30 days from data delivery)

Other obligatory information
Service provider has to report to us immediately about any delays, issues or changes during the project.
If samples pass the QC criteria provided by the service provider but the library construction fails, these libraries will be done max twice without extra cost.
If samples and sequencing libraries pass the QC criteria provided by the service provider but the sequencing does not fulfill the agreed requirements, additional sequencing will be done without extra cost (if DNA or RNA available).
If samples and sequencing libraries pass the QC criteria provided by the service provider but the sequencing fails, new libraries and sequencing will be done ones without extra cost.
The remaining samples need to be stored properly without extra cost for max 6 months and returned to us when requested. Data need to be stored min. 3 months after the project has been accepted. No bioinformatics needed. Alternative experimental settings can be suggested in the tender in addition / instead. Suggestions to negotiate about technical terms is accepted.

The following details in the tender are obligatory (fill in attachments 1 & 2)
- Detailed description of the sequencing design.
- Amount and quality of the data guaranteed.
- State any limitations such as minimum number of samples in a batch.
- Days from receiving the samples to delivering the data in Phase I and Phase II separately. RNA, WGS and methylation data may have different throughput times.
- General sample requirements.

- Prices separately for DNA 1) sample and library quality control and preparation and 2) for sequencing. This is important because the final number of samples and amount of data needed depends on the quality steps described above.
- All possible additional costs for data delivery, returning of samples, invoicing, payment delays etc.

Payment terms (but not more stringent that described below)
- Payment term 30 days net from the date in invoice
- Invoicing not accepted before service provider has started sample processing.
- Max 50% of the phase cost can be invoiced before 50% of each phase has been completed.
- Phases I and II need to be invoiced separately according to the final actualized project cost.
- Last part of the payment need to be at least 25% of the phase cost and can be invoiced not earlier than 30 days after final approval of the data.

Deadline for tenders: 6th of June 2016 by 12.00 noon (Eastern European time)
We aim at signing the agreements and starting the sequencing project (sending the 1st batch of samples) as soon as possible in June.

Send the tender by email to: hy-kirjaamo@helsinki.fi
Use title 'Tender for HERCULES NGS services' and attach the documents in pdf-format.

Futher information:
Questions by email (with title 'HERCULES NGS services') by 31.5.2016 to:
Tiia Pelkonen (tiia.pelkonen@helsinki.fi), Project Manager
on behalf of Dr Rainer Lehtonen, Senior researcher (on vacation until 6th of June)

University of Helsinki,
Systems Biology of Drug Resistance in Cancer research group,
Horizon2020 project HERCULES
http://www.project-hercules.eu/
http://research.med.helsinki.fi/gsb/hautaniemi/default.html