Department of Genetics

University of Delhi South Campus

New Delhi- 110021

Tender Enquiry for Human Whole Exome Sequencing

Two component sealed tender is invited from national companies with experience and expertise in **whole exome sequencing, for sequencing 200-500 human exomes with and without data analysis** as per specifications below. The sequencing and data analysis cost should be mentioned separately in the offer. Quotations with technical details may be sent to the undersigned latest by **5 pm on September 6, 2016.**

1) **Whole exome sequencing:**

- Whole exome sequencing should be performed using the Illumina HiSeq 2000/2500/4000/X Ten/X Five platforms
- Exome enrichment should be performed using **Agilent SureSelect Human All Exon V5**
- Each sample should be sequenced to at least 140-fold (140X) mean target depth with 101-bp/150-bp paired-end sequencing method
- For each sample >98% target regions (**Agilent SureSelect Human All Exon V5)** should be captured and sequenced with minimum 1-X coverage, and >96% of the target region should be captured and sequenced with minimum 10-X coverage. The target depth calculation should be based on per base and using well established tools like bedtools
- Minimum 7.0 GB clean (after removal of adapters, short reads and low quality sequence/reads removed-that is phred quality score (Q score) <20) sequence data per sample should be delivered.
- PCR duplicates in the cleaned data should be <2% and the PCR duplication rate should be calculated based using Picard and QualiMap
- >95% of total data should be >Q30 Phred score
- Summary of complete sequencing runs should be provided.

Bioinformatic analysis of exome data

1) **Read quality check and data filtering:** Detailed quality report before and after trimming should be provided. It should include base quality and sequence quality score distribution details, average base content and GC distribution in the reads, PCR amplification details, check for over-represented sequences, adapter trimming details, read length details (percentage of read length distribution), parameters used for trimming and names and details of software used and other details of trimming/removal of low quality sequence/reads etc.
2) **Read alignment:** The paired end reads should be aligned to reference genomeGRCh37/hg19. Alignment should be done using standard tools such as

BWA.Details of parameters used for alignment andsoftware details, read alignment statisticsand quality metrics for each sample should be provided.

3) **Details of target depth and raw depth:** Percentage of target regions captured in each sample and details target region covered by the probe but of not captured and coverage depth details should be provided.
4) **Variant calling:** Variant calling (SNPs and INDELs) should be performed using standard procedures of GATK or SAMtools. Summary reports of variants called from each sample should be. provided. Number of SNPs,INDELs, homozygous and heterozygous variants and transition to transversion ratio, read depth and quality distribution of identified variants in each sample should be provided. Variant calling should be done individually and together (all samples together to produce single VCF file for all samples). Variant filtration should be done using standard procedures and parameters and details of procedures and parameters used for the filtering.
5) **Variant annotation:** Complete annotation of the variants obtained above using contemporary tools and their summary statistics should be provided. Also the names of software and parameters used for the annotation should be provided.

Sample Details: 200-500 samples will be sent to the firm.

Data delivery deadline: Sample DNA QC report should be submitted within 5 days after receiving the samples. If not, it will be considered as QC passed and liability for quality and timely delivery of data will lie with the vendor. Raw FastQ file, Clean FastQ file and other quality and parameter files mentioned above (at 1) should be delivered within 25 days after sample QC completed. BAM file, filtered and unfiltered VCF files (individual sample's VCF and all samples together) and other parameters and quality files mentioned above (at 2,3,4,5) should be delivered within maximum 45 days after sample qc check is completed.

Other mandatory conditions

1) Vendor should submit evidence of prior experience in human whole exome/genome sequencing
2) Vendor should have carried out minimum 100 human whole exome sequencing in the last one year.
3) Complete list of clients/institutes wherein whole exome/genome/clinical genome sequencing projects in last one year successfully completed should be submitted

Penalty clause: If data quality does not fulfil the above detailed criteria, the data will not be accepted and payment will not be made and order will be cancelled. If timely delivery of data is not done, penalty will be applied both in the case of raw data and analysed data. In case of late delivery of raw and/analysed data, 1% of total cost will be deducted for delay of every day and delay of more than 25 days will lead to the cancellation of the order.

- **Vendor does not follow the above mentioned detailed criteria, the institute/Lab has the full right to cancel the order at any time. In that case no payment would be made.**

- **The selection of the service provider for the above mentioned service will be based on evidence of expertise in the field and technical specifications mentioned in the tender and lowest quoted price.**
- **The qualified vendor should provide a bank guaranty of 20% of the total cost (refundable after completion of work) and needs to sign an agreement. If the awarded work is not completed in time with above mentioned conditions the order will stand cancelled and bank guaranty will not be returned. If the vendor cancels the order because of their own internal problems the bank guaranty will not be returned.**

- **Technical and commercial bids should be submitted separately**

- Per sample cost for the human whole exome sequencing for 200 to 300 and 301 to 500 samples with and without bioinformatic analysis may be provided.

Quotations may be sent to

Prof B.K. Thelma

University of Delhi South Campus
Department of Genetics
Benito Juarez Road, DhaulaKuan
New Delhi- 110021
Ph: + 91 -1124118201

On or before September 06, 2016

Thelma BK
Aug 17, 2016