mRNA TRANSCRIPTOMICS PROTOCOL
NIGERIAN INSTITUTE OF MEDICAL RESEARCH
CENTRAL RESEARCH LABORATORY

Sample collection and RNA extraction

Approximately 15 hair follicles at the vertex area of the scalp shall be collected per participant. The sampled hair follicles were preserved by RNA later solution (Qiagen, ON, Canada) for future batch analysis.

Total RNA will be extracted for both RNA-Seq and quantitative PCR (qPCR) analyses using an RNeasy Micro Kit (Qiagen). Collected samples will be homogenized using a sonicator (ThermoFisher, ON, Canada) in a homogenization buffer prior to the extraction and purification steps with the HiBind RNA Spin Columns. The kit provides both homogenization buffer and the spin columns. The concentration and purity of the RNA extracts will be determined using a NanoDrop spectrophotometer (ThermoFisher) and as for RNA integrity number (RIN) on a Bioanalyzer (Agilent Technologies, ON, Canada). RNA samples with RIN >7, a quantity of 10 ng – 1 ug, and RNA integrity: $OD_{260/280}$ = 1.8~ 2.0, $OD_{260/230} \geq 2$ will be used for sequencing (Zhang, *et al*., 2017).

Sequencing using MGIEasy RNA Library Prep

Ribosomal RNA (rRNA) depleted total RNA libraries will be prepared using the MGIEasy RNA Library Prep (MGI, China) following manufacturer's instruction. All resulting libraries will be checked for size and concentration using Bioanalyzer (Agilent Technologies) and Qubit (Thermo Fisher Scientific, ON, Canada), respectively. Sequencing runs will be performed on a MGI DNBSEG G50 instrument (MGI, China) with the DNBSEQ-G50 (SE50/SE100) sequencing kit. Sample barcodes will be de-multiplexed and raw data would be deposited to the short reads archive (SRA) or Gene Expression Omnibus (GEO).

RNA-Seq data preprocessing and gene differential expression (DE) analysis

The raw sequencing reads from all the samples will be aligned to the human genome (release version hg19) using TopHat (Trapnell et al. 2009). Prior to read counting, software samtools (Li

et al. 2009) will be used to sort the bam files resulting from the sequence alignment by the name of the genome features. Reads will be counted for all the genomic features using HTseq-count (Anders et al. 2015) with the annotation file corresponding to the reference genome version. With the rest settings being default, feature type, counting mode and stranded were set to "gene," "intersection-nonempty," and "reverse," respectively. The raw read counts will then be filtered and normalized using the voom method (Law et al. 2014) through the R package limma (Ritchie et al. 2015).

Statistical Analysis

The R package limma (Ritchie et al. 2015) will be used for gene differential expression (DE) analysis. Statistical analysis will be based on linear model fitting with empirical Bayesian test. The empirical Bayesian test p values will be corrected using FDR (false discovery rate) with an alpha value set to 0.05. According to the scope of this study, genome features under the categories of protein coding and miRNA (microRNA) will be extracted for further analysis.

Quantitative PCR

Statistical significantly differential expressed genes from RNA-Seq analysis will be confirmed using qPCR (Quantstudio 3, ThermoFisher-USA. A subset of general house keeping genes will be used. All primers will be designed using the Primer-BLAST tool from NCBI (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and synthesized at the oligo lab of the Nigerian Institute of Medical Research.

REFRENCES

Anders, S. , Pyl P. T., and Huber W.. 2015. HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169. [PMC free article] [PubMed] [Google Scholar]

Law, C. W. , Chen Y., Shi W., and Smyth G. K.. 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 15:R29. [PMC free article] [PubMed] [Google Scholar]

Li, H. , Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., et al.; 1000 Genome Project Data Processing Subgroup . 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. [PMC free article] [PubMed] [Google Scholar]

Ritchie, M. E. , Phipson B., Wu D., Hu Y., Law C. W., Shi W., et al. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43:e47. [PMC free article] [PubMed] [Google Scholar]

Trapnell, C. , Pachter L., and Salzberg S. L.. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105–1111. [PMC free article] [PubMed] [Google Scholar]

Zhang J, Wallace SJ, Shiu MY, Smith I, Rhind SG, Langlois VS. Human hair follicle transcriptome profiling: a minimally invasive tool to assess molecular adaptations upon low-volume, high-intensity interval training. *Physiol Rep*. 2017;5(23):e13534. doi:10.14814/phy2.13534