**Opis wymagań dotyczących sekwencjonowania oraz bioinformatycznych analiz danych uzyskanych z sekwencjonowania całego eksomu (WES, ang. whole-exome sequencing) dla 100 próbek: 50 pacjentów z ADHD i 50 zdrowych kontroli**

|  |  |  |  |
| --- | --- | --- | --- |
| **L.p.** | Nazwa parametru lub funkcja pomiarowa | **Wymagana odpowiedź** | **Odpowiedź****Wykonawcy** |
|  | Próbki DNA izolowane zestawem Oragene saliva kit Firmy Genotek zwieszone w buforze TE, o parametrach minimalnych 6ug gDNA (min stężenie 50ng/ul i objętości 120-150ul) Czystość: OD260/280 w zakresie 1.7 - 2.0.  | Tak |  |
|  | Wykonanie oczyszczania próbek w zależności od potrzeb. | Tak |  |
|  | Przygotowanie bibliotek eksomowych (WES) z wykorzystaniem zestawu Twist Human Core Exome (Twist Bioscience) lub SureSelect Human All Exon V7 (Agilent) | Tak |  |
|  | Wykonanie sekwencjonowania z wykorzystaniem urządzenia Illumina Novaseq 6000 | Tak |  |
|  | Całkowita ilość danych - minimum 8Gb na próbkę (+/- 10%). | Tak |  |
|  | Długość sparowanych odczytów – 2x150 bp (PE150) (Paired-end) (minimum 2 x 100 bp) | Tak |  |
|  | Dokładność odczytu Q30 (maksymalnie 1 błąd odczytu na 1000pz) dla co najmniej 70% sekwencji |  |  |
|  | Średnie pokrycie 50x na próbkę ( on target) | Tak |  |
|  | Jakość surowych odczytów genomu w formacie FASTQ musi zostać potwierdzona za pomocą raportu z programu fastQC.  | Tak |  |
|  | Odczyty mają być poddane obróbce wstępnej za pomocą programu Illumina bcl2fastq. Wycinanie sekwencji adaptorowych za pomocą programu Skewer Usuwanie duplikatów za pomocą samtools lub Cutadapt. Obowiązują najnowsze wersje programów. | Tak |  |
|  | Uliniowienie odczytów oraz identyfikacja wariantów (“variant calling, indel calling”) wykonane przy pomocy systemu oprogramowania Burrows-Wheeler Aligner (BWA-mem) lub ABRA z wykorzystaniem genomu referencyjnego w wersji hg19 lub nowszej.Wykrywanie wariantów wraz z ich potencjalnymi konsekwencjami (hgncGene, affectedTranscript, Consequence, BaseChange, AAchange, codonChange). Annotacje wariantów do baz danych: Ensembl 92, RefSeq Curated (20180710), CCDSr22, dbSNP151, GnomAD 2.1 | Tak |  |
|  | Dodatkowa analiza bioinformatyczna obejmuje:1) filtrowanie pod kątem rzadkich wariantów (przy użyciu gnomAD) oraz wariantów skracających białka (STOP codon, frameshift)2) filtrowanie wariantów na podstawie dostarczonej listy genów3) porównania statystyczne (np. dokładny test Fishera) w celu ustalenia, czy istnieją różnice w częstości występowania wariantów wytypowanych genów z punku 1 i 2 pomiędzy grupą pacjentów z ADHD w porównaniu do zdrowej grupy kontrolnej.4) przygotowanie materiałów i tabel z przeprowadzonych analiz do publikacji | Tak |  |
|  | Wyniki mają zostać dostarczone w postaci plików \*.FASTQ (surowe odczyty), \*. FASTQ.GZ (usunięte adaptery), \*.BAM/.BAI (uliniowione odczyty), \*.VCF (variant calling oddzielnie dla SNP oraz INDEL), każdemu \*.VCF towarzyszy \*.TSV (annotacje wariantów do baz danych). | Tak |  |
|  | Czas wykonania usługi sekwencjonowania eksomów nie może przekroczyć 14 tygodni od daty dostarczenia próbek.  | Tak |  |
|  | Punkty: A. Cena: 20 punktowB. Wykonanie analiz na aparatach Illumina NovaSeq6000. 40 punktówC. Certyfikaty potwierdzające doświadczenie i umiejętnosć sekwencjonowania np.: CAP, CLIA, ISO lub EMQ. 40 punktow | Tak |  |