

Helgenomsekvensering (WGS) – fördjupad teknisk information

Den aktuella analysen innebär sekvensering av hela det mänskliga genomet (helgenomsekvensering, WGS), och har mycket god teknisk förmåga att detektera såväl små genetiska varianter (SNVs och mindre indels) som större kopietalsavvikelser (CNVs, deletioner och duplikationer). Då kunskapen om icke-kodande genetisk variation är begränsad så kommer klinisk tolkning för s.k. singelprover, d.v.s när endast symptomatisk patient finns tillgänglig, främst vara begränsad till kodande sekvens inom angiven genlista (se individuellt labsvar), och redan välbeskrivna patogena varianter i icke-kodande sekvens. Vid utredning där det bereds tillgång till föräldraprover (s.k. trio) utförs en bredare analys, vilken inkluderar samtliga kodande gener. Se nedan för en mer detaljerad beskrivning av det bioinformatiska flödet vid WGS.

Varianter vilka tolkas som benigna, sannolikt benigna eller av oklar klinisk signifikans rapporteras ej, såtillvida det inte motiveras av möjlighet till uppföljande klinisk analys. Framtida kunskap kan komma att modifiera aktuell tolkning, varför det är möjligt att med förnyad remiss begära omanalys av framtagna data.

Metoden detekterar även trinukleotidexpansioner (t.ex. Fragilt-X, Dystrofia myotonika, Huntington, C9ORF72-relaterad ALS/FTD, samt flera spinocerebellära ataxier) men i dagsläget är sensitiviteten för den bioinformatiska algoritmen inte fastställd. Vid stark misstanke om en specifik expansionsjukdom rekommenderas därför riktad analys. Analysen innefattar för tillfället inte förändringar i mitokondriellt DNA eller komplexa strukturella varianter.

Syftet med WGS-analysen är att upptäcka orsaken till ett tillstånd hos patienten, inte att upptäcka bärarstatus av recessiva sjukdomar. Bärarstatus av en given sjukdom kan förfrågas i specifika situationer, t.ex. vid konsanguinitet, eller om partner till patienten har en molekylärgenetiskt bekräftad sjukdom/bärarskap.

Nomenklatur i labsvar följer HGVS nomenklatur guidelines (<http://varnomen.hgvs.org>) och variantklassificering följer riktlinjer från ACMG (Genetics in Medicine, 2015, PMID: 25741868).

Bioinformatiskt flöde vid WGS

Preparering av sekvensbibliotek för helgenomsekvensering (WGS) görs med TruSeq DNA PCR-Free (Illumina) och sekvensering utförs med NovaSeq 6000 (Illumina). Basbestämning görs med bcl2fastq (Illumina) och efterföljande sekvensinpassning och variantbestämning utförs med Sentieons analysflöde för DNAscope med genomversion GRCh38 som referenssekvens (https://support.sentieon.com/manual/DNAscope_usage/dnascope). Avseende SNV- och Indel-varianter bedöms endast varianter lokaliserade i eller inom 30 baspar från exon (enligt senaste Ensembl-version) samt introniska/intrageniska varianter tidigare rapporterade som patogena eller sannolikt patogena i databasen ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar>). Eftersökning av kliniskt relevanta kopietalsvarianter (CNVs) och genotypning för detektion av t.ex. uniparentell disomi (UPD) utförs däremot för hela genomet.

Sekvensvarianter (SNVs/Indels) annoteras med bl.a. populationsfrekvens och proteinskadlighet med verktyget Variant Effect Predictor (VEP, <https://www.ensembl.org/info/docs/tools/vep/index.html>) och rankas för deras sannolikhet att vara patogena med Genmod (en flerstegsprocess där all information som krävs för att beräkna rankscore, t.ex. populationsfrekvenser, proteinpåverkan och nedärvningsmönster, tillfogas VCF-filen, <https://github.com/moonso/genmod>). Manuell variantbedömning görs primärt i Scout (<https://github.com/Clinical-Genomics/scout/>, SciLifeLab).

CNV-varianter eftersöks med hjälp av tre olika variantbestämmare; CNVnator (PMID: 21324876), TIDDIT (PMID: 28781756) och Manta (PMID: 26647377). CNVnator använder sig av läsdjupet över kromosomer för variantbestämning, specialiserat för att hitta stora kopietalsvarianter. TIDDIT och Manta använder sig av diskordanta parade läsningar i kombination med split-read-läsningar för variantbestämning, och har därför bättre känslighet vid små avvikelser samt möjliggör detektion av kopieneutrala strukturella varianter. Detekterade CNV-varianter annoteras och rankas i ett annoteringsflöde bestående av verktygen VEP, AnnotSV (PMID: 29669011), Pre-score (Perlskript), Genmod och Compound finder (Perlskript). Varianttolkning och manuell annotering utförs i Scout med stöd av Alamut Visual (<https://www.interactive-biosoftware.com/alamut-visual>) och IGV (<https://software.broadinstitute.org/software/igv>) samt lokalt utvecklade visualiseringsverktyg (bl.a. WGS-gSNP, WGS-POD och Gens).