2 Methods

2.1 Introduction

Sample ID: [redacted] was analysed for the clinical diagnosis of [redacted]. Clinical grade sequencing (ISO 15189 accredited) was used to generate whole genome sequence (WGS) data at the Swiss Multi-Omic Center (SMOC) (SMOC). We handled this sensitive clinical data according to established SPHN/BioMedIT guidelines on the sciCORE platform the SwissPedHealth `momic` tenant (SPHN/BioMedIT). The full set of SwissPedHealth pipeline documentation can be found at our pipedev docs homepage. The analysis design documents for this specific report can be found at https://swisspedhealth-pipelinedev.github.io/docs/pages/design_dna_snvindel_v1.html. The pipeline used here is `Design DNA SNV INDEL v1` and concerns WGS germline short variant discovery (SNVs + Indels) and interpretation.

2.2 Sequencing and processing

We used Illumina whole genome sequencing (WGS) with TruSeq DNA PCR-free library protocol. Sequencing was performed on NovaSeq 6000, for paired-end 150 bp reads with indexing. We mapped to the human reference genome GRCh38/hg38 assembly (link to ref docs) ensuring high accuracy and comprehensive coverage for genomic analysis (figure 2.2). The `Design DNA SNV INDEL v1` protocol allows for the discovery and interpretation of germline short variants (SNVs and INDELs) from whole genome sequencing data. This protocol includes GATK best practices to process raw sequencing data into a joint cohort gVCF, followed by variant qualification and annotation for clinical and research applications. Key stages include quality control, genome alignment, variant calling, and rigorous filtering to identify disease-associated variants, ensuring that both known and novel genetic determinants are

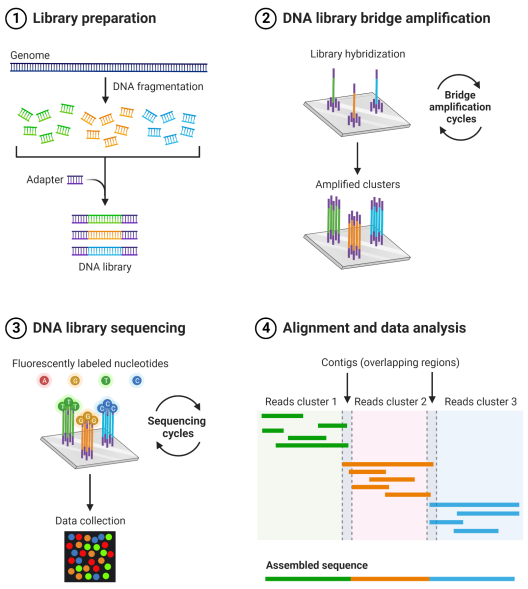

Figure 1: Illumina WGS data generation.