A018 Implementation of Next-Generation Sequencing Bioinformatics Pipeline for Solid Tumor Gene Panel

Ethan P.M. LaRochelle, Donald C. Green, Sophie J. Deharvengt, Kimberly N. Winnick, and Gregory J. Tsongalis
Dartmouth Hitchcock Medical Center, Lebanon, NH

Introduction: Genomic analysis of Formalin-fixed, Paraffin-embedded tissue samples collected from solid tumors are used to personalize treatments. The TruSight Tumor 170 Panel (Illumina) (TST170) detects a variety of clinically actionable variant types. The sample preparation and sequencing are identical for clinical and research specimens, but due to cost constraints the bioinformatics workflow differs. Open-source tools have been used to programmatically analyze sequencing data and further automate variant calling. **Methods:** After sequencing paired DNA and RNA libraries, clinical and research FASTQ files are analyzed using two different workflows. Clinical FASTQ files are uploaded to the Clinical Genomic Workspace (PierianDx), whereas research samples are processed on a local server. Both analysis channels utilize the TST170 Local App v1.0.1 (Illumina) which is a Docker-based application used to perform alignment and variant calling. The variant calls for the research samples are filtered using a customized Python script based on read quality, population prevalence, and variant location and consequence. An additional script is used to analyze all reported variants and combined with run-specific quality information. This script is run monthly for bulk historical analysis to generate a visual interactive dashboard for both clinical and research samples. **Results:** Over the past 8 months 400 clinical samples from 376 patients, and an additional 57 research samples have been processed using the described workflow. The mean DNA and RNA quantification after library preparation is 18.4 +/- 4.4 ng/mcL and 22.5 +/- 6.3 ng/mcL, respectively. The interactive dashboard provides a visual breakdown of reported variant types for each tumor type. Of the six most clinically tested tumor types the distribution of substitutions (SNVs), copy number variants (CNVs) and RNA Fusions is tabulated as: lung (81%, 4%, 2%), colon (77%, 3%, 1%), glioma (66%, 14%, 6%), prostate (59%, 4%, 23%), breast (65%, 24%, 1%) and melanoma (92%, 1%, 1%). **Conclusions:** Open source tools can be used to efficiently extract information from next-generation sequencing runs, providing insight on run quality as well as clinical trends such as tumor type, variant type, involved genes or fusion call sequences. An initial application of this dashboard shows a higher percentage of CNVs in breast and glioma compared to other tumor types, and a high percentage of fusions in prostate tumors.