P36

The preliminary results from the 100,000 genomes project: the genomic landscape of colorectal cancer

Cross, W.C.H.[1]; Caravagna, G.[2]; Houseman, J.[1]; Sosinsky, A.[3]; Murugaesu, N.[3]; Chubb, D.[2]; Cornish, A.[2]; Wedge, D.[4]; Houlston, R.[2]; Sottoriva, A.[2]; Graham, T.[1]; Tomlinson, I.[5]

[1]Barts Cancer Institute, London, United Kingdom; [2]ICR, London, United Kingdom; [3]Genomics England, London, United Kingdom; [4]Big Data Institute, Oxford University, Oxford, United Kingdom; [5]Institute of Cancer and Genomic Sciences, Birmingham, United Kingdom

Purpose of Study: The remit of the 100,000 Genomes England Project is to develop a new genomics service for the NHS, while facilitating the research of a large number of human conditions, including more than 20 types of cancer. Here, we present the preliminary analysis of 1000 colorectal cancer genomes which represents the most comprehensive dataset of it's kind in the world.

Methods: We performed bioinformatic processing and analysis using a wide variety of tools: the Illumina ISAAC workflow was used in combination with Strelka, Starling and Sequenza to obtain somatic and germline mutational calls. The Ensembl variant effect predictor (VEP) was used to score and annotate mutations and we referred to the TCGA, COSMIC and other sources to identify known driver mutations and compare our results to those published so far.

Summary of Results: The cohort comprises 1077 patients with 400 detailed clinical annotations and at least 120X whole genome sequencing of each primary and normal pairs. Hierarchical clustering of the mutations present in 264 known driver genes revealed a multitude of differing genotypes, including clusters dominated by several rarer wnt signalling disruptions and SMAD4, FBXW7 and PIK3CA, and several genes with as yet unknown etiology in colorectal cancer. Notably, although the canonical genes APC, KRAS and TP53 were the most frequently mutated, only a minor fraction of cancers possessed mutations in all three. Analysis of the karyotypes underlying these genomes revealed the presence of a core set of copy number aberrations. Karyotypes also fit into several subtypes including mostly diploid, those containing chromothripsis, and those with genome doubling which represents around 40% of non-MSI CRCs.

Conclusions: Colorectal cancers can be driven by hugely varied drivers and karyotypic states and in these preliminary results we point to several new features that are only made obvious through en masse analysis of whole genomes.