Methods for molecular profiling of tumor samples by next generation sequencing.

Protocol and Ethical issues.

Our molecular laboratory is an accredited Italian Society of Pathology reference centre for RAS testing and the organiser in Italy for the ESP Colon External Quality Assessment Scheme. After obtaining the patient's consent, oncologists and the primary pathologists from outside institutions record the clinical and pathological data (including the original pathology report) on a dedicated website. Then, the corresponding tissue sample is express-mailed to our central laboratory. Upon receipt of each sample, a representative H&E stained slide is reviewed by a pathologist and the area with the highest density of neoplastic cells is marked, annotating the percentage of neoplastic cells.

Since RAS mutational analysis is the standard of care in diagnostic workup of patients with CRC, and our analysis did not interfere anyhow with the patient management, the need for ethic committee's approval was not necessary for this study, in accordance with medical ethical guidelines of the Università degli Studi di Napoli Federico II and in accordance with general authorisation to process personal data for scientific research purposes from 'The Italian Data Protection Authority', All samples and clinical data used in this study have been irreversibly anonymized.

Depending on the complexity of histology and on the density of the tumour, DNA was extracted using the QIAamp DNA Mini Kit (Qiagen, Crawley, West Sussex, UK) from two (resection specimens) or three (biopsy specimens) 10 μm-thick serial sections. An additional section (biopsy specimens only) was stained by H&E to confirm tumour cell percentage. DNA was extracted from cell lines and clinical tissue samples using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. DNA was suspended in 30 μL of molecular biology water. DNA quantity and quality were assessed using the Qubit photometer (Life Technologies) and the Qubit dsDNA HS (High Sensitivity) Assay Kit according to the manufacturer's instructions.

According to the manufacturer's protocols, 10 ng of DNA for each sample was used for library preparation with the Ion AmpliSeq Library 96LV Kit 2.0 (Life Technologies) and the Colon and Lung Cancer Panel (Life Technologies). This panel gives 90 amplicons covering 504 mutational hotspot regions in 22 genes (AKT1, ALK, BRAF, CTNNB1, DDR2, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, KRAS, MAP2K1, MET, NOTCH1, NRAS, PIK3CA, PTEN, SMAD4, STK11, TP53), with performance of at least 500× sequence coverage for eight samples on one Ion 316 chip. For samples yielding less than 10 ng DNA input, additional cycling conditions were used for library preparation as recommended by

the manufacturer. Each library was barcoded with the Ion Xpress Barcode Adapters 1–16 Kit (Life Technologies). Barcoded libraries were combined to a final concentration of 100 pM. Template preparation by emulsion PCR (emPCR) was performed on the Ion OneTouch 2 system (Life Technologies). Library quality control was performed using the Ion Sphere Quality Control Kit according to the manufacturer's instructions, ensuring that 10–30% of template positive Ion Sphere particles (ISP) were targeted in the emPCR reaction. Sequencing primer and polymerase were added to the final enriched ISPs prior to loading onto 316 (100 Mb output) chips. Sequencing was carried out on the PGM (Life Technologies). Data analysis was carried out with Torrent Suite Software V.3.2 (Life Technologies). After alignment to the hg19 human reference genome, the Variant Caller plug-in was applied using the Colon and Lung hotspot file as a reference (downloaded from Ion Community, http://www.ioncommunity.lifetechnologies.com, last accessed 1 September 2015). The Ion Reporter suite (Life Technologies) was used to filter polymorphic variants. In addition, all nucleotide variations with less than a 5% variant frequency were masked. All detected variants were manually reviewed with the Integrative Genomics Viewer (IGV V.2.1, Broad Institute, Cambridge, Massachusetts, USA) or with Genome Brower web app.

Performance parameters

In all cases analyzed, a 100 pM DNA library was obtained; only in 24 cases, the library preparation procedure was repeated, after an initial failure. While most cases yielded a DNA input > 10 ng, eight samples did not satisfied this request. However, even for these cases an increase in the number of amplification cycles enabled to get an adequate library. An average of 3.9 million of the total 6.3 million addressable wells in the Ion 316 chip were consistently loaded with ISPs, and 3.2 million (92%) of these particles contained library templates. After subtraction of multiple-templated beads and poor quality sequence reads, an average of 2.7 million reads were obtained. Samples averaged 193,000 mapped sequence reads (range, 10,331 to 1,010,971) with a mean read length was 115 bp. Multiplex PCR mediated target capture was very effective, as an average of 93.5% of the sequence reads mapped to targeted gene regions. The distribution of reads across the 90 amplicons was consistent across samples and there was an average of 1930 reads per amplicon (range, 102 to 10982).