Genomic Instability Revealed by Ultra-deep Sequencing of a Human Ultraconserved Region

De Grassi A[1], Segala C[1], Volorio S[2], Bertario L[3], Radice P[2,4], Bernard L[1], Ciccarelli FD[1]

[1] Department of Experimental Oncology, European Institute of Oncology, IFOM-IEO Campus, Milano
[2] IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, IFOM-IEO Campus, Milano
[3] Hereditary Colorectal Tumor Registry, Milano
[4] Department of Experimental Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan

Motivation

Genomic instability (GI) is the result of spontaneous DNA changes that accumulate in a cell population. It is a common trait of cancer genomes and plays a pivotal role in promoting cancerogenesis in several hereditary tumors. Current methods to detect GI require the presence of already established neoplastic cells because they only detect mutations present in the most of the clonal cell population. This returned picture is a “static snapshot” of the cancer genome, where only the tip of the iceberg is captured (i.e. clonal mutations). The detection of rare mutations in addition to clonal mutations would offer a “dynamic” view of the mutational landscape and be instrumental to clarify many controversial aspects of cancer genetics. In principle, next-generation sequencing technologies could furnish a valid approach to increase the sensitivity in detecting GI, as they rely on the amplification and sequencing of distinct DNA filaments. Because sensitivity of these methods increases with coverage, rare mutations should become detectable by performing an ultra-deep re-sequencing of the same DNA region. The obvious drawback is connected with specificity: at deep coverage, low frequency substitutions are an indistinguishable mixture of technical errors and true mutations, which makes hard to distinguish true signal from noise. A possible solution to weigh technical errors is to use internal controls, i.e. genomic elements that do not accumulate mutations. Ultraconserved regions (UCRs) of the human genome constitute a possible repository of hypomutable segments. The working hypothesis is that, by re-sequencing an UCR together with its non-conserved flanking regions, true mutations should prevail over sequencing errors in the flanking segments but not in the UCR core. A difference of mutability between the two regions would be considered an evidence of GI because the mutation rate of healthy human genome is so low that sequencing errors would overcome true mutations along the entire region.

Methods

We carefully selected a ~1,500 bp genomic region containing an UCR and amplified it from the genome of patients affected by Hereditary Non-Polyposis Colon Cancer (HNPCC). This is an autosomal dominant condition due to germline heterozygous mutations in mismatch repair genes and is typically associated with GI. We amplified the region from three different tissues of the patients: neoplastic colon mucosa; non-neoplastic colon mucosa and peripheral blood. As a negative control we used the peripheral blood of healthy donors. We pooled the DNA from different individuals in equimolar ratio to obtain four tissue-specific samples. Each sample was sequenced on both sides using a dedicated 70x75 mm PicoTiterPlate run on the GSFLX Sequencer. This protocol led to the sequencing of more than 80 million single bases per run, corresponding to a coverage of ~40,000 reads per sample. We divided bases differing from the reference sequence into high and low frequency mutations, above and below the detection power of the sequencing technology (0.1%), respectively. Starting from experimental data, we applied a model of cell grow that reproduced the progressive accumulation of mutations during cancer clonal expansion. By comparing the observed number of high frequency mutations to the number expected by the model we could estimate the mutation rate, which is the number of mutations introduced per nucleotide per cycle, in the HNPCC genome. Low frequency mutations are instead an indistinguishable mix of mutations and errors. We screened these mutations along the entire region to check for a significant tendency of bases flanking the UCR to preferentially accumulate mutations with respect to the UCR core.

Results

We exploited the frozen status of an UCR to increase sensitivity and specificity of ultra-deep se-

quencing and hence quantify cancer-associated GI. The obtained results offered several insights into cancer genetics. We could estimate that the mutation rate of a random not constrained region of the HNPCC genome does not exceed 6x10-6. Assuming that replication errors hit random positions of the human genome, less than one out of 60 human genes will acquire one somatic mutation at each cell division. Of those, only mutations that are either neutral or beneficial for the cancer cell will be kept and eventually fixed into the population. We observed a tendency of the region flanking the UCR to accumulate significantly more mutations than the UCR core in both neoplastic and non-neoplastic HNPCC tissues. No difference between the two regions was detectable in cells from healthy donors. This is the first clear-cut evidence of an intrinsic GI of individuals with heterozygous mutations in the mismatch repair genes. Our analysis suggests a predisposition of such individuals to acquire the second hit necessary for cancer development, and constitutes the proof of principle for the development of a more sensitive molecular marker of GI.

Contact : anna.degrassi@ifom-ieo-campus.it