Math Technique De-Clutters Cancer-Cell Data, Revealing Tumor Evolution, Treatment Leads

Cold Spring Harbor Laboratory

In our daily lives, clutter is something that gets in our way, something that makes it harder for us to accomplish things. For doctors and scientists trying to parse mountains of raw biological data, clutter is more than a nuisance; it can stand in the way of figuring out how best to treat someone who is very sick.

Using increasingly cheap and rapid methods to read the billions of "letters" that comprise human genomes – including the genomes of individual cells sampled from cancerous tumors -- scientists are generating far more data than they can easily interpret.

Today, two scientists from Cold Spring Harbor Laboratory (CSHL) publish a mathematical method of simplifying and interpreting genome data bearing evidence of mutations, such as those that characterize specific cancers. Not only is the technique highly accurate; it has immediate utility in efforts to parse tumor cells, in order to determine a patient's prognosis and the best approach to treatment.

CSHL Assistant Professor Alexander Krasnitz, who developed the new technique jointly with American Cancer Society Professor Michael Wigler, explains that it reduces the burden of interpretation by identifying what he and Wigler call COREs, an acronym for "cores of recurrent events."

Consider the example of a cancerous breast tumor. Central to the CORE concept is what Krasnitz and Wigler refer to as "intervals." An example of an interval would be a segment of DNA that is missing in the genetic sequence of one or more cells sampled from the tumor. Tumor cells are often missing DNA that should normally be present; or conversely, they often have genome intervals in which the normal DNA sequence is amplified – it appears in multiple copies. Such deletions and amplifications are called copy-number variations, or CNVs.

"In cancer," says Krasnitz, "we find intervals in the genome that are hit again and again. You might see this in many cells coming from a single patient's tumor; or you may see these repeating patterns in cells sampled from many patients with a similar cancer type."

In either case, if you superimpose the location of each "hit" – whether a deletion or an amplification of DNA -- against a map of the full human genome, "you end up with these wobbly pile-ups, stacks of 'hits' at the same locations in the genome."

Due to the vagaries of collecting genome data and a certain amount of small-scale variation in the precise boundaries of the deleted or amplified DNA intervals, the

stacks don't line up straight; as Krasnitz says, they look "wobbly." This makes them very hard to accurately interpret.

The CORE method he and Wigler describe in a paper appearing in Proceedings of the National Academy of Sciences "is a mathematical way of cleaning up this mess and untangling these stacks of data, which often overlap." When data from 100 cells from a single tumor are analyzed, for example, and the mathematical algorithm devised by Krasnitz and Wigler is applied, the regularity of the stacks is revealed, and the rich structure of the data emerges.

In the example of analyzing 100 cells from one tumor, the net result is that populations and subpopulations of cancer cells can be distinguished; and if the cancer has already become metastatic, CORE will be useful in discerning the relations among cancer cell subpopulations in various parts of the body. Such analysis is a potentially valuable guide to prognosis and can also help to make important treatment decisions.

For more information, visit Cold Spring Harbor Laboratory [1].

Source URL (retrieved on *09/18/2014 - 10:46am*):
http://www.mdtmag.com/news/2013/06/math-technique-de-clutters-cancer-cell-data-revealing-tumor-evolution-treatment-leads

Links:
[1] http://www.cshl.org