November 25, 2013

On mathematics and brain cancer

A new approach for discovery from data opens a new way of studying cancer

Searching for mathematical patterns hidden in brain cancer data, a team of University of Utah bioengineers, geneticists and mathematicians found that differences in the lengths of genes that are produced in abundant numbers in a cell reflect metabolic differences between healthy and tumor cells.

This work, published in the journal *PLoS One* on November 25, 2013, opens a new way of studying cancer.

"We show for the first time that information about the functioning of a cell can be deduced from the lengths of genes, independent of the sequences of the genes," said Orly Alter, USTAR associate professor of bioengineering and human genetics at the Scientific Computing and Imaging Institute, and the team leader. Genes are stretches of DNA of different sequences and lengths. Cells produce RNA copies from genes as part of the process of making proteins, the primary working molecules of the cell.

By using the patterns found in the data, Alter and her team – mathematics and bioengineering undergraduate student Nicolas M. Bertagnolli, alumnus Justin A. Drake, and NIH postdoctoral fellow of human genetics Jason M. Tennessen – were able to identify the probability distributions, that is the formulas that chart the diversity of gene lengths in healthy and tumor cells.

This demonstrates a new approach for discovery from data.

"Natural processes are often described by mathematical formulas that chart diversity," said Alter. "Usually, these formulas are derived by making prior assumptions about the data. Here we were able to get the formulas for the diversity of gene lengths directly from the data."

Surprisingly, the formulas that the team found resemble the formula in physics that describes a pendulum. This suggests that the diversity of gene lengths is the result of evolutionary forces of selection that affect gene length in a manner similar to the way in which gravity acts upon a pendulum. The farther you lift up a pendulum, the larger is the force that swings it back to its resting position. Similarly, because of these evolutionary forces, the length of a gene is most likely to be a particular length, with the likelihood decreasing the farther you move from this length in either direction.

By using these formulas, the team found that genes overproduced in brain tumor cells are likely to be significantly shorter than genes overproduced in healthy brain cells. Healthy cells overproduce short as well as long genes. The short genes are involved in making proteins and in generating energy by consuming oxygen. The long genes are involved in brain activity and in drawing energy from glucose. Tumor cells overproduce the same short genes as healthy cells, but not the long ones.

The team concluded that a cell's metabolism – such as energy generation – can be regulated by controlling the production of RNA copies of long genes.

To make this discovery, Alter and her team combined data from different sources, including glioblastoma multiforme (GBM) data from The Cancer Genome Atlas (TCGA), a national effort to accelerate cure for cancer.

"It may very well be that the data needed to cure cancer are already published. TCGA's brain cancer data, for example, were published back in 2008," said Tennessen. "The bottleneck to discovery is in the analysis of the data. What made our discovery possible is our new mathematical approach, going from data patterns to formulas."

GBM – the most common brain cancer in adults – causes over 10,000 deaths each year in the US alone, often within one year of diagnosis. Despite extensive studies with the latest biotechnological tools and the collection of huge amounts of data, the basic biology of this brain cancer is poorly understood, and a cure remains unknown.

The insights into GBM formation and growth that are presented in this study may someday lead to personalize treatments for patients.

By merging mathematics, biology and medicine, Alter and her team ultimately hope to bring physicians one step closer to being able to predict and control the progression of cancer as readily as NASA engineers plot the trajectories of spacecraft today.

Funding for Alter's research comes from National Science Foundation (NSF) CAREER Award DMS-0847173, National Human Genome Research Institute (NHGRI) R01 Grant HG-004302 and the Utah Science, Technology and Research (USTAR) initiative. Tennessen's research is funded by National Institute of General Medical Sciences (NIGMS) K99/R00 Pathway to Independence Award GM-101341.

To read the November 25, 2013 *PLoS One* article go to:
https://doi.org/10.1371/journal.pone.0078913