MedicalXpress

Overexpression of repetitive DNA sequences discovered in common tumor cells

January 13 2011

Massachusetts General Hospital (MGH) Cancer Center researchers have discovered a previously unknown feature of common tumor cells – massive overexpression of certain DNA sequences that do not code for proteins. These DNA sequences – called satellite repeats – have been studied for their role in chromosomal structure but previously were not suspected of having a role in cancer. The report will appear in the journal *Science* and is receiving early online release.

"Satellite repeats make up a large part of our genome but had been thought to be inactive," explains David Ting, MD, of the MGH Cancer Center, co-lead author of the *Science* paper. "We found that these regions are, in fact, very active in cancer but not in normal tissue. The findings may give us a novel cancer biomarker, as well as new insights into how cancers behave."

Because previously available tools for analyzing the transcription of DNA into RNA were designed to focus on sequences that are eventually translated into proteins, they excluded segments present in multiple-repeat copies that do not produce proteins. Among these stretches of DNA are satellite repeats, repetitive sequences often found near the centers or the tips of chromosomes. Significant expression of satellite repeats had been seen previously only in embryonic tissues or embryonic stem cells.

The current study was designed to give a more comprehensive picture of the transcriptome – the full range of RNA molecules – of primary

tumors. Using an advanced digital gene expression analysis system called single molecule next-generation sequencing, the MGH team first studied samples from a mouse model of pancreatic cancer and were surprised to find that satellite DNA was expressed at levels more than 100 times what would be expected in normal tissues. Greatly increased satellite expression was also found in mouse colon and lung tumors, and all the tested samples were epithelial cancers, the most common type of solid tumor.

Analysis of human tumor samples produced similar results, with powerful overexpression of two satellites called HSATII and ALR in the majority of cancers studied, including tumors of the pancreas, lung, prostate. Ting notes that finding increased satellite expression in lower-grade tumors suggested that overexpression begins early in tumor development, which has implications for early detection.

"Cancer diagnoses are increasingly being made on the basis of fine-needle biopsies, which yield small numbers of cells that must be correctly identified as malignant," he explains. "In a few of the analyzed samples, our team demonstrated that pancreatic cancer cells were correctly identified based on satellite RNA expression, which was appreciably higher than in nonmalignant cells. If confirmed in large prospective clinical trials, satellite RNA expression may provide a new and highly specific biomarker relevant to multiple types of epithelial cancers."

Daniel Haber, MD, PhD, director of the MGH Cancer Center and senior author of the *Science* paper, says, "What is most remarkable is how such a dramatic abnormality was only revealed because of new powerful sequencing technologies that allow us to study a type of RNA that was previously discarded. Our hope is that this abnormality will serve as an important biomarker in cancer diagnosis and that it will also shed light on common mechanisms by which cancer develops."

Provided by Massachusetts General Hospital