FEATURE

Mutational Signatures in Human Cancer

Written by Brian Hoyle

To date, 10 000 to 20 000 cancer genomes have been fully or partially sequenced. Knowledge of the genetic composition of this array of cancer genomes has (1) allowed the identification of 600 to 700 genes whose mutation triggers the change of normal cell to a cancer cell, (2) implicated biological processes that are important in cancer development, and (3) identified drug targets. Accordingly, cancer genome sequences could prove crucial in predicting therapy outcome and response and could allow the development of assays that detect sequence fragments that leak out of cancer cells. It is hoped that these fragments will prove to be diagnostic hallmarks.

In his discussion on the mutational signatures of human cancer, Michael Stratton, PhD, Wellcome Trust Sanger Institute, Hinxton, United Kingdom, noted that the analysis of cancer genomes is increasing the understanding of mutational processes that occur in carcinogenesis. It is generally accepted that cancers arise because of somatic mutations—mutations that arise in a cell and are passed on to future generations of cells. Yet, despite this awareness, little is known of what causes the mutations that drive the cancer transformation.

Replication of the some 600 million base pairs of DNA during cell division commonly introduces errors. Cells are also under a daily onslaught of exposure to irradiation and mutagens among other things that can cause genetic changes. As a defense, cells have evolved a repair mechanism that identifies the genetic error and restores the correct sequence. But the process is not perfect, especially when genetic changes are occurring at a higher-than-random rate because of exposure to sunlight or tobacco smoke, for example.

Studies that have examined mutations in the *TP53* gene—one of the most frequent sites of mutation in human cancers—have revealed 6 mutations due to substitution of one nucleoside base by another: C (cytosine) to T (thymine), C to A (adenine), C to G (guanine), T to A, T to C, and T to G. The pattern of these mutations differs in various cancers. For example, examination of skin cancer tumors has revealed the predominance of the C-to-T mutation consistent with exposure to ultraviolet light, while samples of lung cancers display more C-to-A mutations consistent with the known effect of tobacco carcinogens. Put another way, the mutagen exposure leaves a molecular signature of the particular cancer type.

But this picture is not as simple as a regular pattern of mutations. The accumulation of mutations in a late-stage cancer cell does not occur at a constant rate with time. Rather, some mutations can occur throughout the lifetime of the cell, with 1 of the 6 possible base substitutions predominating, while other patterns of base substitution can occur more intensely at certain times in the cell life (eg, exposure to tobacco smoke), with different mutations occurring at different times. Thus, the final cancer genome is a snapshot of the final product of the numerous temporally occurring mutations that does not resemble the patterns of the contributing mutations.

Cancer genome sequences represent the final genome snapshot. The challenge is to dissect the genome information to identify the hidden molecular signatures during the cell lifetime. This is important, since some mutations could be more instrumental in driving the transition of a normal cell to a cancer cell.

To this end, Prof Stratton and colleagues have identified the bases immediately before and after the mutated base in each of the 6 known base substitutions in cancer cells. Because there are 4 possible bases before and after a mutated base, there are 96 different base patterns that can be present. These 96 patterns form the basis of what is termed the molecular signature of cancer cells. Analysis of sequences from >7000 cancer cases involving 30 cancer types representing about 5 million somatic mutations revealed the pattern of molecular signatures as well as the contribution of each signature to the overall mutations in each cancer genome.

Peer-Reviewed Highlights From the

European Society for Medical Oncology Annual Meeting

September 26–30, 2014
Madrid, Spain

The many mutational signatures (originally 26, since expanded to about 30) and the markedly diverse patterns of the type and frequency of mutated bases indicate a lot of underlying mutational processes, more than the researchers were expecting to find. The molecular signatures can be assigned to respective cancer types.

Some signatures are ubiquitous in the variety of cancers examined, while others are more type specific. Similarly, some cancers feature only a few signature types, while others have more mutational signatures, with different cancers presenting differing signature patterns.

The patterns are just mathematical constructs. What is important is to determine the underlying mutational processes that can generate varied mutations among the cancers. Research has identified a number of functions associated with the mutational signatures and, in the case of signature 2, has found evidence suggesting that the basis of the mutation may be viral infection, retrotransposon movement in the genome, or inflammation. In this scenario, the cancer arises due to collateral damage inflicted to the genome from 1 or more of these 3 causes.

Mutations can occur throughout the genome, but localized hypermutation is also present. Dr Stratton and colleagues coined the term *kataegis* (the ancient Greek word for "thunderstorm") to describe the clustering of mutations. The locations of the hypermutations differ among cancers. A common feature is that these regions tend concentrate in 1 or several of the 6 possible base substitutions.

In summary, there are >20 genomewide mutational signatures in the cancer types. Mutations affecting cytosine deamination are common in many cancer types. Regions of hypermutation are also common in many cancers. Work proceeds to decode the mutational processes that generate the mutational signatures. This understanding is crucial. By interpreting the molecular details of how cancers arise, strategies of better treatment will be devised. The ultimate goal is to prevent the transformation of a normal cell to a cancer cell.