Distinguishing the handful of somatic mutations expected to initiate and maintain cancer growth, so-called driver mutations, from mutations that play no role in cancer development, passenger mutations, remains a major hurdle for understanding the mechanisms of cancer and the design of more effective treatments. Recognizing this, National Cancer Institute’s “Provocative Questions” Project (1) specifically highlights the urgent need to better discriminate between driver and passenger events as a key research priority. In the past, many studies took a genecentric approach to the problem (2-7), identifying potential driver mutations as those that occur in genes mutated in a high percentage of the tumor samples. A pathway analysis typically follows to add functional context to the mutated genes. Unfortunately, this approach is limited to a small subset of genes and inherently disregards gene mutations occurring in a low percentage of tumor samples. To truly uncover the significance of somatic cancer genomics, we will need to embrace the highly complex mutation landscape originating from distinct DNA damage and repair processes.

The study by Nik-Zainal et al. was landmark for this very reason (8). The authors sequenced the genomes of 21 breast cancers and created a catalogue of all the mutations in the genomes of the 21 cancer genomes and identified distinct patterns of mutations in breast cancer. They were able to detect five mutational signatures of which three signatures have never before been described. These signatures help guide our expanding understanding of DNA damage and repair mechanisms. Importantly, their findings challenged classical theories of cancer evolution (9). Instead of a gradual accumulation of genetic events, the authors showed that point mutations can occur at somatic hypermutation hotspots and result in a catastrophic mutational event. The authors call this “kataegis” (from the Greek for thunderstorm): although never described before, kataegis was remarkably common occurring, to some extent, in the genomes of 13 of the 21 breast cancers. Within areas of kataegis, one of the more commonly seen cancer somatic mutation signature is an overrepresentation of C-to-T and C-to-G at the TpCpX dinucleotide. One potential mechanism for the increased localized hypermutation could be mediated by the action of APOBEC family of proteins. The APOBEC1 protein shows a context specific preference for C residues preceded by a dT and is involved in the deamination of cytosines to uracil, which can be either read through or create an abasic site through base excision repair (10,11). A recent study further validated the role APOBEC proteins play in breast cancer (12). Tumors that express high levels of APOBEC3B have twice as many mutations as those that express low levels and are more likely to have mutations in TP53. APOBEC3B-catalysed deamination provides a chronic source of DNA damage in breast cancers that could select p53 inactivation and explain how some tumours evolve rapidly and manifest heterogeneity.

Beyond the mutational signatures identified, a moderate degree of strand bias was detectable for C>A/G>T transitions across the 21 breast cancer genomes and is present in almost all cases. This bias was characterized by fewer G>T mutations on transcribed than untranscribed
strands. A strand bias was also observed for T>G/A>C mutations with fewer T>G mutations on transcribed than untranscribed strands. Others have also observed a similar strand bias in breast cancer (13). The authors propose that this may be due to transcription-coupled DNA repair (TCR). TCR is implicated in the removal of bulky DNA adducts which are normally formed from exposure to an endogenous or exogenous genotoxic insult (14). What the genotoxic insult is in breast cancer is as yet underdetermined. Similar observations are seen in UV-light-associated skin cancers, where C>T and CC>TT transitions are common and occur at dipyrimidines, reflecting the formation of pyrimidine dimers following exposure of DNA to UV light (15) and also show transcriptional strand bias due to the action of TCR on pyrimidine dimers. Beyond recognizing mutation strand bias, the study showed that gene expression was inversely correlated with specific types of substitutions. That similar results were seen in both lung cancer and melanoma (16-18) suggest that mutational processes characterized by both transcriptional strand bias and expression-related mutation prevalence may be operative in many cancers types.

Another insight from the study was the surprisingly similar mutational pattern seen in breast cancers associated with BRCA1/2 germline mutations. In an unsupervised hierarchical clustering analysis, the cancers with BRCA1/2 mutations grouped together. This corroborates previous observations of distinct somatic allelic imbalance profiles separating BRCA1/2-related breast cancers from sporadic ones (19). BRCA1/2-related cancers exhibited a mutational signature representative of a broad distribution of mutations rather than a predominance of C>T mutations at XpCpG seen in the other breast cancers. BRCA1/2-mutant cancers were associated with more and larger indels flanked by regions of microhomology. Overlapping microhomology is often considered a signature of nonhomologous end-joining (NHEJ) DNA double-strand break repair. The authors speculate that because BRCA1/2 are involved in homologous recombination (HR)-based DSB repair (20), the elevated frequency of microhomology-mediated indels in BRCA1/2 mutant cancers presumably reflects usage of alternative methods of DSB repair in these cancers. Such a mechanism would explain the observation that somatic allelic imbalance frequencies in BRCA1/2-related breast cancers and their surrounding microenvironment are significantly higher than those of sporadic counterparts (19).

Large-scale efforts such as the Collaborative Oncological Gene-environment Study (COGS) recently uncovered new genetic susceptibility loci for breast, ovarian and prostate cancers (21-25). These newly identified susceptibility loci explain an increasing proportion of the familial risk of these cancers and emphasizes the need to fully understand genetic susceptibility related to tumor heterogeneity and pleiotropy. Future studies should seek to better characterize the functional impact of identified cancer susceptibility loci and it’s interactions with environmental and/or lifestyle factors.

With advances in computational methods, it is highly likely that many further cryptic mutational signatures will be extracted once more cancers have been analyzed. It is imperative that we continue to chip away at our understanding of the biological basis of these mutational signatures. We have seen how an early understanding of exogenous genotoxic insults in lung cancer and melanoma paved the way for successful public health preventative efforts in reducing the prevalence of both these cancers. That we see mutational signatures implicating similar DNA damage and repair processes in breast cancer suggest an urgent need to better understand environmental/exogenous as well as endogenous processes underlying somatic mutagenesis across the spectrum of human cancer if we are to ever outpace and outsmart the tempest that is cancer. If, in fact, the kataegis event is universal in carcinogenesis, then that doyen of lung cancer prevention Ki Hong’s concept of prevention through reverse migration might seem a reasonable roadmap (26,27).

**Acknowledgements**

J.N. is the Ambrose Monell Foundation Cancer Genomic Medicine Clinical Fellow at the Cleveland Clinic Genomic Medicine Institute. C.E is the Sondra J. and Stephen R. Hardis Chair of Cancer Genomic Medicine at the Cleveland Clinic and is an American Cancer Society Clinical Research Professor, generously funded in part, by the F.M. Kirby Foundation. E.N. was funded, in part, by 2T32GM007250 and 5T32GM008613 and is funded, in part, by 1F30CA168151-01A1.

**Disclosure:** No author had any financial or personal relationships that could inappropriately influence or bias this work.

**References**

guide/rfa-files/RFA-CA-12-017.html


