

Rebellious enzymes: mutation showers in cancer induced by APOBEC ‘kataegisters’

In 2012, cancer genomics saw the addition of a new phrase to its lexicon: kataegis, Greek for thunderstorm and an apt description for the discovery of localized regions of hypermutation that were revealed when numerous breast cancer genomes were sequenced in a landmark study in *Cell*. Michael Stratton from the Wellcome Trust Sanger Institute, UK, and fellow co-authors of the study speculated that the APOBEC family of enzymes could be responsible. The first attempt to test this theory was led by Youri Pavlov and Artem Lada from the University of Nebraska Medical Center, USA, and colleagues, who demonstrated that the APOBEC PmCDA1 could indeed cause kataegis in a yeast model, as [published in a Biology Direct study](#). Here, Pavlov and Lada discuss how this story unfolded, explaining why APOBECs became the prime suspect in the case of kataegis.

Cancer is a genetic disorder, meaning that mutations, either somatic or germline, underlie the ‘Emperor of Maladies’ etiology. Nucleotide substitutions in DNA, chromosomal rearrangements and copy number changes spontaneously occurring in predisposed individuals or induced by numerous intrinsic and environmental mutagens such as oxidative stress, tobacco smoke, and UV light, lead to the activation of oncogenes and inactivation of tumor suppressors, therefore promoting malignancy. [This textbook scenario](#) has been recognized, experimentally scrutinized and confirmed for decades. Research was mainly focused on the driver mutations that ultimately lead to cancer, while what happens to the rest of the genome during tumor development has been elusive. During the last decade, a burst in so-called ‘next-generation’ sequencing technologies allowed reading of the whole genomes of solid tumors and liquid malignancies belonging to different types and stages of cancer, giving birth to the exciting, new field of cancer genomics.

One of the most striking findings from these whole-genome sequencing efforts was that cancer genomes are highly enriched with mutations of different kinds. Numbers of sequence nucleotide polymorphisms (SNPs) in some cancers are in the orders of tens of thousands. It still has to be determined what proportion of these mutations is silent (neutral), and what is beneficial and detrimental (of course, from the tumor’s selfish ‘point of view’). Nevertheless, one important finding is standing solid: various mutations can be classified to the ‘families’ based on their mutation signatures ([Nature Reviews Clinical Oncology. 2013, 10, 545](#)). A mutation signature is the type of base substitution (e.g. C:G to T:A) in the context of neighboring nucleotide sequence (the bases upstream and/or downstream of the mutation). Signatures help scientists to understand the nature of the mutagenic process that operated during tumor evolution. One of the most indisputable signatures that was found in the breast and other cancers is characterized by the C:G to T:A or C:G to G:C substitutions that were predominantly found in the 5'-TC sequence motifs. These mutations were frequently found in clusters, i.e. some regions of the genome were highly enriched

with SNPs of this type. The clusters are reminiscent of mutation rains poured over the genomes, motivating Michael Stratton and colleagues to call this phenomenon [kataegis](#), from the Greek word meaning thunderstorm.

It has been immediately recognized that the most likely suspects in producing mutations of this kind are editing cytosine deaminases. These enzymes, collectively called APOBECs, deaminate cytosine in the single-stranded DNA, producing uracil. Replication through the uracil leads to the insertion of A, therefore causing the C to T substitution. Alternatively, REV1 translesion DNA polymerase can insert C in front of an abasic site that is produced as an intermediate of uracil repair, thus leading to C:G to G:C mutations. Deaminases possess inherent sequence specificity. For example, Activation Induced Deaminase (AID) prefers to deaminate in the 5'-WRC motifs (W – A or T, R – A or G), and APOBEC3G on the last cytosine in the 5'-CCC motif. Two other APOBEC3 enzymes – APOBEC3A and APOBEC3B – prefer 5'-TC sequences. Another prominent (and not well understood) feature of APOBEC enzymes is their ability to act in a processive fashion, i.e. able to catalyze multiple deaminations in one substrate binding event ([see the movie from the Myron Goodman lab](#)). Deaminase-like mutation signatures in many cancers have also been reported by the Harris and Gordenin groups, and APOBEC3B is frequently upregulated in breast and ovarian tumors ([Nature Genetics comment by Larry Loeb](#)). Altogether, based on the sequence specificity and processive properties, it was hypothesized that APOBEC deaminases are the driving force of kataegis in cancer.

The idea of APOBECs being carcinogenic was originally proposed by Michael Neuberger and colleagues in early 2000s ([Trends in Biochemical Sciences. 2003, 28, 6, 305-312](#)), after the discovery that these proteins can edit DNA ([Molecular Cell. 2002, 10, 5, 1247-1253](#)) and, therefore, are mutators by definition. Under normal conditions, deaminases are involved in [adaptive \(AID\)](#) and [innate \(APOBEC3s\)](#) immunity, lipid metabolism (APOBEC1) and maybe even active DNA demethylation ([Annual Review of Genetics. 2012, 46, 419-441](#)) (see the [talk by Svend Petersen-Mahrt on AID and epigenetics](#)) both in development and in terminally differentiated cells. Extremely tight and complex (and thus, not surprisingly, poorly understood) regulation of APOBECs ensure that they edit cytosines in very specific sites such as immunoglobulin genes or viral cDNA. However, when the regulatory cordons fall, these housekeepers can rebel and edit DNA genome-wide.

To get a bulletproof confirmation that APOBECs directly cause kataegis, we developed a system combining baker's yeast and sea lamprey deaminase. This hyperactive enzyme is called PmCDA1 (for *Petromyzon marinus* cytidine deaminase) and is involved in the generation of antibody analogs in jawless fish ([Nature Immunology. 2007, 8, 647-656](#)). Like APOBEC3A and 3B it prefers 5'-TC sequences. Overexpression of PmCDA1 is highly mutagenic in haploid yeast – the standard genetic system for studying mutagenesis. We then decided to test PmCDA1 in the diploid yeast strain which is a more appropriate model because most human cells are diploid. In order to sensitize cells to deamination, in our strain we also disrupted a gene encoding uracil-DNA-glycosylase – a base excision repair enzyme catalyzing uracil removal. Lamprey deaminase was

extremely mutagenic in the diploid strain, where even the cells that were severely hit by deaminase and acquired multiple mutations can survive and produce viable mutant clones ([PLoS Genet. 2013, 9\(9\), e1003736](#)).

Whole-genome sequencing of the obtained diploid mutants revealed that lamprey deaminase induced hundreds to thousands of mutations per diploid 24 Mb yeast genome. In some mutant clones, up to 1 single nucleotide polymorphism (SNP) per 10 Kb has been observed. Strikingly, the distribution of mutations over the genome was far from random: many mutations were clustered. Most clusters resembled kataegistic micro-clusters found in tumor genomes. The clusters were non-random and some genomic regions have been found to be exceptionally prone to enzymatic deamination. For example, in one particular region of only ~50 bp on yeast chromosome X, mutations have been found in all clones analyzed, even though the selection was for mutations in a different genomic region - canavanine transporter gene *CAN1* (located on chromosome V). This is bad news for the cells where deaminases are out of control because it literally means that, under certain conditions, mutation frequency in the hypermutated cell subpopulation in a tiny genomic region equals 100 percent. Apparently, powerful APOBEC 'kataegister' binds to unprotected ssDNA and travels along it, catalyzing multiple deaminations in a processive manner. These results are [published in *Biology Direct*](#).

The Neuberger group also used yeast to study deaminase effects, and found kataegistic mutation clusters in haploid wild-type yeast ([eLife. 2013, 2, e00534](#)). Inactivation of uracil-DNA-glycosylase greatly reduced mutation clustering in haploids. This indicates that, depending on conditions such as ploidy and DNA repair proficiency of the cells, kataegis can be produced by various pathways.

The mechanisms of deaminase-induced kataegis are still far from being clear. Basically, all processes that have single-stranded DNA intermediates – replication, recombination, transcription and repair – can potentially result in kataegis in the presence of a deaminase. However, overexpression of active APOBECs is highly toxic in human cell lines (as reported by Matthew Weitzman's group ([EMBO Rep. 2011, 12\(5\), 444–450](#)) and later developed in the Harris lab ([Nature Reviews Cancer. 2013, 13, 220-221](#))) indicating that precise balancing of deaminase production and other factors are required in order to cause non-lethal genome-wide hypermutagenesis and kataegis. The majority of riots are non-productive. This is apparently also true in the case of APOBECs where only a small fraction of cells with unleashed deaminases and a fine-tuned environment survive and give rise to malignant clones. It is also possible that sudden overproduction of deaminases in tumor cells with genomes shaped by other mutagenic processes will effectively kill the tumor, unless the cells provide a special 'anti-venom' against APOBEC.

By kataegis, cancer cells acquire stretches of DNA, which are highly heterologous to DNA of normal cells. This could be used in the future to discriminate between the two cell types and selectively destroy DNA of cancer cells and kill them.