Review

RNA Editors and DNA Mutators: Cancer Heterogeneity Through Sequence Diversification

Rafail Nikolaos Tasakis [1, 2], F. Nina Papavasiliou [1, *], Rita Shaknovich [3, 4, *]

1. Division of Immune Diversity, German Cancer Research Center (DKFZ), INF 280, 69120 Heidelberg, Germany; E-Mails: r.tasakis@dkfz-heidelberg.de, n.papavasiliou@dkfz-heidelberg.de
2. Faculty of Biosciences, Uni Heidelberg, 69120 Heidelberg, Germany
3. Department of Medicine, Division of Hematology and Medical Oncology, Weill Cornell Medicine, New York, NY, 10021, USA; E- Mail: rshaknovich@gmail.com
4. GRAIL, Inc., 1525 O'Brien Drive, Menlo Park, California, 94025, USA

* **Correspondences:** F. Nina Papavasiliou and Rita Shaknovich; E-Mails: n.papavasiliou@dkfz-heidelberg.de, rshaknovich@gmail.com

Academic Editor: Kakoli Das

Special Issue: Genetic Heterogeneity in Cancer

Abstract

Cancer development and progression is strongly associated with somatic mutations. From oncogenic hits that initiate the primary tumor formation to metastasis, the tumor mutational burden (TMB) plays a prominent role in the disease progression for the vast majority of cancer types. Not only are heterogeneous mutational loads or genetic heterogeneity causal to transcriptomic and proteomic discrepancies and to phenotypic diversity between individuals, they are also between tumor cells. But in addition to mutations, a new set of alterations at the level of RNA (termed "epitranscriptomic") can also cause informational heterogeneity. Chief among epitranscriptomic modifiers are the RNA editing enzymes (ADAR and AID/APOBEC family deaminases), which catalyze A-to-I and C-to-U base changes within RNA transcripts. Because all of these enzymes also retain the ability

to mutate DNA, they can be generally considered as central drivers of inter-tumor heterogeneity either at the DNA level or at the RNA level. Here, we connect genome/transcriptome diversification with tumorigenesis and suggest specific roles for these DNA/RNA deaminases in the context of tumorigenesis both through genome level events (mutations) and through transcriptome level events as these converge into a common sequence diversification strategy.

Keywords

RNA editing; DNA mutation; base deamination; ADAR; AID/APOBEC; sequence diversification; cancer immunotherapy; TMB; MSI

1. Introduction

Cancer is indisputably the most diverse group of diseases and is one of the most challenging for today's biomedical research. Cancer may arise due to different extrinsic and/or intrinsic causes that drive transformation through heterologous mechanisms often associated with a variety of genomic events, which can then lead to an even more diverse clinical presentation and course [1, 2]. Cancer development often follows the same pathogenesis: starting from a single genetic event and leading to acquisition of additional genetic and epigenetic abnormalities, clonal evolution and plasticity (events that confer remarkable biological advantage over normal differentiated cellular elements, together with the ability to evolve, develop drug resistance and escape immune surveillance) [3, 4]. One of the first attempts to explain cancer development, with regards to somatic (DNA) mutations, was the two-hit hypothesis proposed by Alfred Knudson in 1971, which posits that cancer development requires at least two mutations: one in a proto-oncogene, which then turns into an oncogene, and one in a tumor-suppressor gene [5]. In other words, to become cancerous, a cell must lose control over cell division but must also escape death. What we have learned in the decades since, is that cancer cells incur well beyond two mutations: in fact, the mutational load in most cancers is high and this is generally thought to speed up the evolution of the tumor [6, 7].

Tumors originate from cancer stem cells (CSCs) or progenitor cells, and tumor formation follows a clonal evolution model [8]. Mutations arise relatively frequently and are generally thought to be a byproduct of cellular proliferation processes through errors in DNA replication and transcription [9]. They can also arise by the targeted action of enzymes that insert non-canonical bases into the genome (e.g. deaminases, reviewed below; Figure 1). In both cases, mutations are corrected by DNA repair enzymes [10], but when they are excessive in number they can also be "fixed" as an error into the genome and get inherited onto future generations [11]. Mutations that give a growth advantage to a specific cell will be positively selected [12]; the cell will then be called a CSC and the mutations that initiated CSC formation will be termed "driver" mutations [13]. CSCs share to a great extent characteristics of "regular" stem cells, such as self-renewal, multipotency and tumor-initiating capacity [14]. The last two aforementioned characteristics, which reflect the cellular plasticity of CSCs, can be enhanced by a set of complementary mutations (often termed "passenger" mutations). Chronologically, passenger mutations occur after driver mutations and,

even though they may have not been the events that initiate tumorigenesis, they do provide additional sequence heterogeneity [15] powering intra-tumor clonal evolution, which follows the principles of the Darwinian evolution concept [16]. These additional mutations will eventually allow the CSC to further adapt to different environments, empowering metastasis (Figure 1) [17-19].
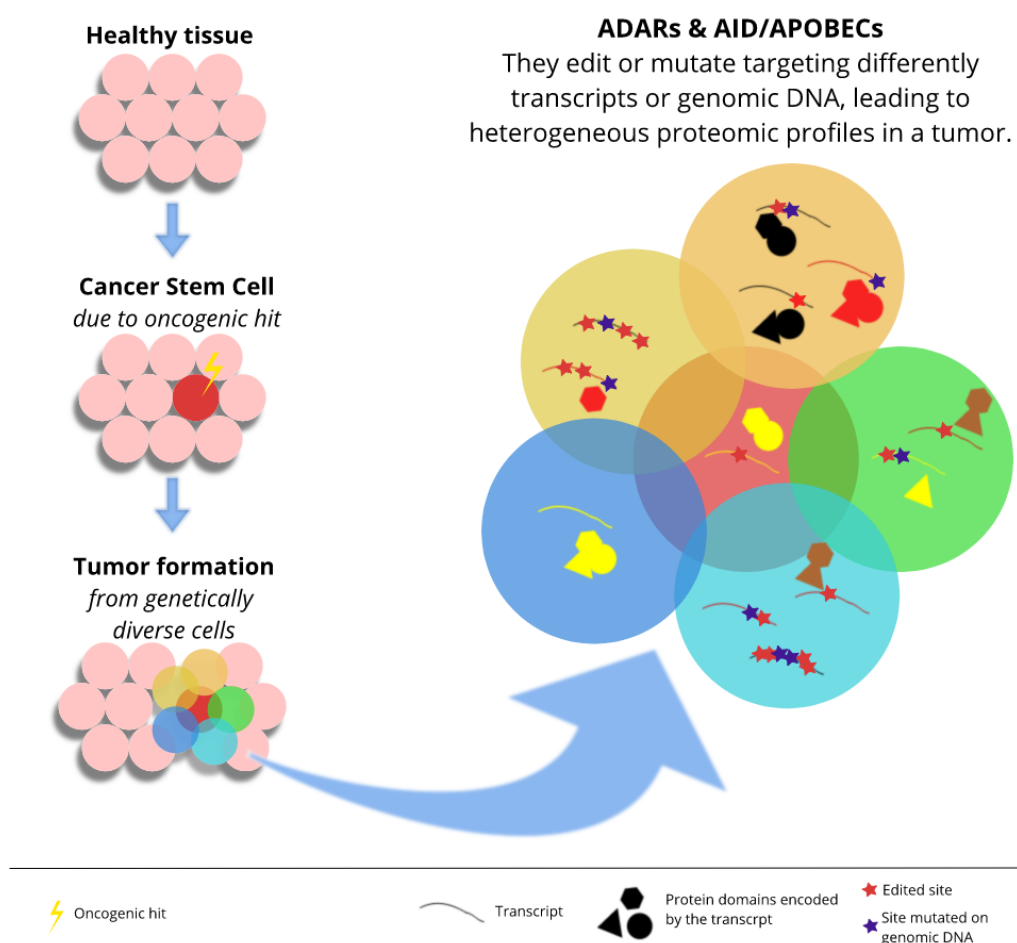

Figure 1 RNA editors and DNA mutators (ADARs & AID/APOBECs) mediate intra-tumor heterogeneity. The clonal evolution model of cancer development maintains that a cell receives an oncogenic hit and evolves into a cancer stem cell that is responsible for forming a primary tumor. Aside from the initial hit, additional hits (e.g. through targeted base changes catalyzed by enzymes of the RNA/DNA deaminase family) continually empower phenotypic diversity.

However, intra-tumor mutational heterogeneity, is not always high [20]. Furthermore, a high mutation load does always translate to heterogeneity; several tumor types harbor predominant mutations within a small set of oncogenes or tumor suppressors [21]. And yet, in such cases, poorly or homogeneously mutated tumors often behave as if they are informationally heterogeneous: they are functionally diverse, and can relapse just as efficiently as their heterogeneous counterparts. For example, excluding hypermutated samples, colon and rectum cancers were found to have considerably similar patterns of genomic alteration - even though they are phenotypically and functionally different entities [22]. Glioblastomas are another example of a

tumor entity that is poorly mutated but quick to relapse suggesting highly heterogeneous disease [23].

Emerging evidence suggests that such mutationally "silent", and yet clonally evolving, tumors are often highly modified at the epitranscriptomic (mRNA) level. Epitranscriptomic modifications and specifically those that are catalyzed by the adenosine or cytidine deaminases (the "RNA editing" family [24, 25]), are already known to impart sequence heterogeneity at the level of the transcript and have been presumed to increase the informational diversity of otherwise genetically similar cell types [26, 27]. RNA editing is mediated by members of the same family of enzymes (polynucleotide deaminases) that were already characterized as prolific DNA mutators. These enzymes come in two flavors: those that belong to the AID/APOBEC family that induce Cytosine-to-Uracil (C-to-U) nucleotide changes [28], and those that belong to the ADAR family that induce Adenosine-to-Inosine (A-to-I) nucleotide changes (where inosine (I) is decoded by the cellular expression machinery as a guanosine (G)) [29]. While initially these were thought to be strictly selective with regard to their nucleic acid substrate (e.g. with AID being a DNA mutator [30] and APOBEC1 or ADAR1 functioning exclusively as RNA editors [31-32]) this view has been revised in light of well-defined contexts where an RNA editor like APOBEC1 or like ADAR1 can function quite capably as a DNA mutator [31, 32] and vice versa, as is the case for APOBEC3A [33].

This review aims to highlight the APOBEC/ADAR DNA/RNA family of deaminases as they contribute to tumor evolution through sequence diversification either by RNA editing or by DNA mutation.

2. Defining the natural roles of AID/APOBECs and ADARs

2.1 The AID/APOBEC Family of DNA/RNA Deaminases

As mentioned, there are two distinct types of RNA editing by deamination: C-to-U editing and A-to-I (where I is recognized as a G) [34]. C-to-U editing is catalyzed by the AID/APOBEC family. APOBEC1, the founding member [35], was initially identified based on its ability to catalyze editing on apolipoprotein B (ApoB); it has been named after this activity (APOlipoprotein-B-mRNA Editing enzyme Catalytic polypeptide-1). ApoB comes in two protein forms encoded by the same genetic locus [36]. In the liver, ApoB exists in its long form (ApoB-100, 4563 amino acid residues), while in the small intestine it is truncated into its short form (ApoB-48). The protein size difference was traced to a single culprit, the APOBEC1 deaminase, which is primarily expressed in small intestine and deaminates a C in the position 6666 (in the triplet CAA) of the ApoB mRNA resulting in stop codon formation (UAA), and in significant truncation of ApoB-100 [37]. The editing event occurs within the exonic region of ApoB, and its biological impact is striking: whereas ApoB-100, expressed in liver, acts as a ligand to the low-density lipoprotein receptor (LDL-R), ApoB-48 (produced through editing in small intestine), lacks the ligand domain, and is involved in the formation and secretion of chylomicrons [38]. However, RNA editing events in exonic regions are rare. Transcriptome-wide screening studies have demonstrated that C-to-U editing preferentially occurs within AU-rich 3'UTRs [39], which suggests that editing plays a prominent role in transcript processing.

While APOBEC1 is capable of editing RNA on its own when ectopically expressed, it normally functions within an editing complex that contains one of several partner proteins (like RNA Binding Motif 47 (RBM47) and APOBEC1 Complementation Factor (A1CF) and probably others as well)

whose function is to guide the editase to specific target transcripts [40]. Consequently, *Rbm47-*deficient mice fail to edit a number of transcripts (including the *ApoB* transcript) *in vivo*, while elimination of A1CF ablates editing of a different set, and double deficiency does not lead to complete loss of editing [41, 42]. These RNA binding proteins might have an additional function, which is to selectively place APOBEC1 onto mRNA (vs DNA). Specifically, APOBEC1 without its cofactors is a capable DNA mutator in *E.coli* (substantially more potent in that regard than AID [43] (see below)), and can also mutate DNA in cancer cells (for example mutational signatures of APOBEC1 correlate with advanced esophageal adenocarcinoma [31]). Therefore APOBEC1, within the nucleus [44], can act both on DNA (to induce DNA mutation) and on RNA (to mediate RNA editing) - but the circumstances under which APOBEC1 functions as an RNA editor vs a DNA mutator are not clear, and could include the lack of the RBP co-factor. In that regard, it is worth noting that mutations that disrupt the function of the Rbm47 co-factor are intimately related to cancer progression [45].

AID (the product of the *Aicda* gene) was the second member of the AID/APOBEC family to be characterized. AID is a key player in adaptive immunity: it was first identified by virtue of its expression in B cells undergoing Ig class switch recombination of the antibody (Ig) locus [46]. It was subsequently discovered that it is also a key mediator of diversification (through somatic hypermutation) of antibody genes [46]. In addition to Ig genes, AID has other genome-wide targets [47]. AID has demonstrable activity as a single stranded DNA mutator: *in vitro*, it deaminates ssDNA [48]. At switch loci, ssDNA becomes accessible during transcription-mediated R-loop formation [48]; at the coding portion of the Ig, which is the recombined variable region, ssDNA is also thought to become accessible during transcription, though evidence for that is not nearly as compelling as for the (stable) R-loops of the switch region [49, 50]. After deamination has taken place, the U is either "fixed" (not repaired) within DNA (as an A:T base pair leading to a transition mutation as compared to the original G:C base pair) or further processed into a DNA break by components of the base excision repair (BER) pathway, leading to class switch recombination or chromosomal translocations [51].

Whereas strong experimental evidence suggests that AID is a DNA mutator, it is still capable of binding RNA [52, 53]. Indeed, it is possible that under the right circumstances (perhaps together with an unknown co-factor) it can also act as an RNA modification enzyme [54] - it is worth remembering that APOBEC1, as well as other APOBECs, are quite capable mutators in the absence of their co-factors. Whereas AID is not a catalytic component of an RNA editase in the context of *ex vivo* stimulated B cells [55], compelling evidence that it cannot act as an RNA editor in other contexts - perhaps even within GC B cells that undergo SHM - does not currently exist. It is interesting to recall that mutational loads for the recombined V region used to be reported in the low percentage digits when detected through reverse transcription and cDNA amplification (e.g. 2-4% mutation accumulation in the recombined VDJ region of Ramos cells [56]); these are currently reported through genomic DNA amplification with rates that are 10 to 100-times lower (in Ramos cells, rates are given as roughly 2×10^{-3} / bp [57]). While this discrepancy might be in part due to different reporting schemes, it is not impossible to imagine that it could also be evidence of RNA modification by AID within the context of GC B cells; this remains to be determined.

The APOBEC3 family was discovered soon after AID was characterized. The human genome encodes 7 APOBEC3s (3A-3D, 3F- 3H), which have important roles in anti-retroviral immunity by virtue of their ability to deaminate viral DNA intermediates and hinder retroviral replication [58,

59]. Specifically, upon retroviral infection and after the viral ssRNA is reverse transcribed, APOBEC3G (the founding member of the APOBEC3 family in humans) was shown to deaminate cytosines (C) to uracils (U) on the cDNA. This can lead both to viral cDNA degradation through UDG and base excision repair engagement, and to viral cDNA hypermutation upon integration as the host's repair machinery replaces Us with Ts [60]. Whereas the vast majority of anti-viral work in the APOBEC3 context has been done through overexpression and viral production in artificial settings (such as in HEK293T cells), it is abundantly clear that most proteins of the APOBEC3 family have the ability to deaminate ssDNA and are proper DNA mutators. Consequently, most APOBEC3 family members are kept away from genomic DNA and are sequestered in the cytoplasm [60]. Exceptions to this rule are APOBEC3A and -3B, which can be found in the nucleus and are correlated with a particular type of genomic mutation termed "kataegis" [61] that is prevalent in a number of tumors. It is possible that kataegis mutations are the result of the requirement for these enzymes to access the genome for the purposes of RNA editing (with mutation being the unfortunate side effect). Indeed, recent work has suggested that APOBEC3A and -3B can function as RNA editors in human monocytes and macrophages [33]. Whether they require co-factors to transition from DNA mutation to RNA editing within the nucleus is currently unknown.

All AID/APOBEC family enzymes are evolutionarily related, with AID and APOBEC2 being the oldest: both are present in the jawed vertebrates. AID plays a central role in the mechanism of antibody diversification. In contrast, APOBEC2 appears to lack the ability to deaminate; yet, its deletion leads to specific functional outcomes in muscle where it is most highly expressed, which suggests it plays a yet to be defined role [62, 63]. APOBEC1 appears later on (in lizards [28]), likely as gene duplication of the AID locus. Finally, the APOBEC3 family arose and expanded in the placental mammals [64]. APOBEC4, the last member of the family identified to date is primarily expressed in testis, seems enzymatically inactive and has yet to be assigned a biological role [65]. Together, these close evolutionary relationships together with the ability of many members of this family of enzymes to toggle between RNA editing and DNA mutation, suggests a previously unacknowledged flexibility of function, which is likely dependent on context, such as co-factors or nuclear/cytoplasmic localization [66]. The roles of AID/APOBECs, as well as of ADARs (discussed in section 2.2), are summarized in Table 1.

Table 1 The physiological roles of ADAR and AID/APOBEC family members. ADARs and AID/APOBECs play important roles in a variety of biological processes, in which diversification is required. Although not all of them demonstrate deamination activity, their expression is necessary for healthy individuals. Here we summarize the physiological roles of ADARs and AID/APOBECs. Further details are discussed in the text (sections 2.1 and 2.2).

Enzymes	Physiological function & facts
ADAR1 & ADAR2	• Deaminates Adenosine to Inosine (A-to-I) in double-stranded RNA (dsRNA) [67] • ADAR1 is ubiquitously expressed, while ADAR2 is expressed only in brain. [67,70] • ADAR1 has two isoforms: p110 (exclusively nuclear) and p150 (mostly cytoplasmic – interferon inducible) [25]. • Involved in host response to viral infections or retrotransposon mobilization [68-69] • A-to-I RNA editing is robust in mammalian RNAs [72].
ADAR3	• Currently considered enzymatically inactive [67].
AID	• Deaminates Cytosine to Uracil (C-to-U) in single stranded DNA (ssDNA) [48] • In B-Cells it is responsible for Ig class switch recombination (CSR) and somatic hypermutation (SHM), leading to antibody diversification [46].
APOBEC1	• Deaminates Cytosine to Uracil (C-to-U) in single-stranded RNA (ssRNA) together with a cofactor (A1CF, RBM47, or others) [40]. • In small intestine, it leads to ApoB truncation due to stop codon formation in exonic region of the *ApoB* transcript [37]. • It edits (pre-) mRNAs primarily in non-coding regions (mainly 3'UTR) and especially in AU-rich regions [39].
APOBEC2	• There is no proven deamination activity for this enzyme. • Expressed highly in muscle with a probable role in muscle differentiation [62-63].
APOBEC3s	• These deaminate Cytosine to Uracil (C-to-U) in ssDNA [58]. • 7 APOBEC3s (3A-3D, 3F-3H) are encoded in the human genome [58]. • Most family members have some role in anti-retroviral immunity, as they deaminate viral nucleic acid intermediates [59].
APOBEC4	• Expressed in testis; currently considered as enzymatically inactive [65]

2.2 The ADAR Family of DNA/RNA Deaminases

The second broad class of polynucleotide deaminases catalyzes the most abundant type of editing (which is the conversion of A-to-I): these enzymes are termed Adenosine deaminases

acting on RNA (ADARs) [67]. There are three ADARs in mammals: ADAR1 and ADAR2, both of which have catalytic activity, and ADAR3, which is currently considered enzymatically inactive. ADAR1 is ubiquitously expressed and has two isoforms: p110, which resides exclusively in the nucleus, and the interferon-inducible p150, which is mostly cytoplasmic [25]. Both are involved in the host response to dsRNA which can arise either from exogenous insults, such as viral infections, or endogenous injury, like retrotransposon mobilization [68, 69].

ADAR2 appears to be functional only in the brain, or at least it is so in healthy individuals: it rose to prominence because of its ability to target the glutamate receptor B (gluR-B) pre-mRNA and within it edit an Arginine to a Glutamine within what is known as the Q/R region of a specific ion channel, decreasing its permeability to Ca^{2+} [70]. This example is probably one of the very few ones in mammals, where ADARs deaminate coding region mRNA to alter amino acid decoding and bring about a strong biological impact. In contrast, ADAR-mediated coding region editing seems to be routine in cephalopods, which use A-to-I editing to radically diversify their proteomes [71]. But overall, it is clear at least in mammals that ADAR mediated editing is most robust and prevalent in intronic regions of pre-mRNAs and in particular, mRNAs that carry Alu repeats [72].

ADARs do not require co-factors to edit RNA. A "cofactor" functionality is embedded within each ADAR by virtue of the dsRNA binding subdomain that they all contain. This subdomain is necessary and sufficient to target A-to-I editing within dsRNA, as for example the hairpins generated at intron-exon junctions, or the longer dsRNA regions generated by the folding of inverted Alu repeats. The modular aspect of ADAR structure - effectively, a deaminase domain tethered to a dsRNA binding subdomain - has found use in current synthetic biology approaches that utilize the catalytic subdomain in conjunction with other RNA binding domains, like cas13, together with guide RNAs to target editing to specific loci at will [73, 74], or with DNA binding domains, like cas9 and guide RNA, to target ADAR to ssDNA through RNA: DNA hybrid formation [32]. Indeed, recent work suggests that ADAR might also have a proper DNA mutator function within RNA: DNA hybrids, for instance, R-loops [32]. All in all, the scientific challenge now is to define the biological context where the dual roles of editors and mutators are understood and biological consequences are delineated.

3. The Contribution of DNA/RNA Polynucleotide Deaminases to Cancer Heterogeneity

The ability of AID/APOBECs and ADARs to target both DNA and RNA in their natural setting can also be exploited by tumors both in gain of function and in loss of function contexts. For example, most tumors overexpress specific AID/APOBECs and these have been correlated with specific gene level mutational spectra in many cancers [31, 61, 75, 76]. In particular, APOBEC3s have been correlated with the presence of tight C-to-T clusters in cancer genomes, which have been termed 'kataegis' mutations [75, 77]. Kataegis mutations were originally described in genetic analyses from 21 breast cancer genomes [61] and now form a specific signature that can be predictive of disease outcomes [78]. Intriguingly however, loss of APOBEC3 function can also lead to tumorigenesis, through a loss of the cells' ability to deaminate viral genomes or restrict retrotransposon intermediates leading to virally-driven transformation and cancer development [79].

Additionally, AID overexpression has been associated with a variety of lymphomas. As discussed, AID catalyzed deamination of the Ig locus can lead to class switch recombination (when

it is targeted to switch regions) to somatic hypermutation (when it is targeted to the rearranged variable region) and finally also to translocations (often between the Ig loci and oncogenes like myc) – [46, 51, 75]. The presence of such translocations in particular, due to AID overexpression, has been demonstrated to play a prominent role in multiple myeloma and several lymphomas [80]: Transgenic mice that overexpress AID developed T-cell lymphomas with AID mutational patterns, affecting the T-cell Receptor (TCR) gene, as well as the c-myc oncogene [76]. AID overexpression also was found to increase the aggressiveness of DLBCL in mouse models [81] and has been correlated with a high mutation load (C to G conversion), epigenetic heterogeneity and aggressiveness of human DLBCLs [82-84]. Interestingly however, loss of AID has also been associated with lymphomagenesis: hyper IgM syndrome patients show evidence of lymphoproliferative disease associated with frank lymphoma formation in some cases [54, 85], though the mechanistic basis of this remains unclear.

RNA editing also has a role in cancer pathogenesis although this has been a bit more controversial. Early experiments in mouse models demonstrated an association between ectopic APOBEC1 overexpression (e.g. in liver) with the development of hepatocellular carcinoma [86]. Similarly, early work with human cell lines showed that APOBEC1 could edit coding regions of the NF1 gene and promote tumor formation [87]. More recently, several experiments have revealed a robust association between loss of APOBEC1 in certain tumor models (e.g. *Apc^Min* mice [88] or models of testicular germ cell tumors [89]) and reduced tumor burden [40]. Mechanistically, we understand this to be due to the sequence heterogeneity imparted by editing at the level of mRNA, both in coding regions and untranslated regions [90]. Of course, APOBEC1 can also act on DNA in the context of tumorigenesis [31]: on one hand this requires caution in the interpretation of the cancer data (is oncogenesis mediated through RNA editing or DNA mutation?). But on the other hand, this linkage suggests the potential for a precursor-product relationship whereby RNA editing, perhaps driven by inflammatory stimuli to suppress transformation in non-cancerous cells, diversifies instead the RNA leading to tumorigenesis. Noting that RNA editing is co-transcriptional [91] and that many RNA editing enzymes moonlight as DNA mutators, it is easy to imagine how an RNA editor like APOBEC1 might lose hold of its mRNA substrate *in situ* and mutate the DNA of the cognate locus, thus "fixing" an editing event into the genome (Figure 2).

Figure 2 Co-transcriptional RNA/DNA editing by APOBEC1. RNA editing enzymes edit RNA co-transcriptionally (here APOBEC1 and its cofactors is shown but we envision ADAR working similarly). At the same time, it is well known that editing enzymes can gain access to and mutate DNA (without a co-factor in the case of APOBEC1 or using an R-loop in the case of ADAR). Here we envision that an RNA editing enzyme loses its grip on RNA and then targets DNA *in situ*. Such a DNA molecule in the vicinity of the cognate transcript will be by necessity the gene that encodes the transcript. In this fashion, we hypothesize that the RNA and DNA deamination functions of these enzymes can be temporally linked.

There is also an emergent role of ADAR editing in cancer pathogenesis. Several investigators employing *in silico* approaches have reported that A-to-I editing levels are exceptionally high in several cancer types, with the most prevalent ones to be lung, head and neck, breast and thyroid cancers [92, 93]. Indeed, the editing load in such cancers is estimated to be as high as the mutational load and can occur both in the coding regions and in the transcript UTRs. It is important to stress however that in contrast to mutational load, which can be described as the historical record of all mutations within the cell - some of which are driver mutations but most of which would qualify as "passengers" - the editing load represents an active mark that diversifies the cellular transcriptome only in the presence of the editing deaminase. Thus, whereas some editing events have been considered as "driver" events (e.g. AZIN1 [94] and others [95]) they must be considered *in toto* in the context of oncogenesis. This is of immediate clinical relevance and several attempts are currently made to generate inhibitors of A-to-I deaminases as small molecules with anti-cancer activity [96, 97].

4. Concluding Remarks and Outlook: Sequence Heterogeneity in the Context of Cancer

Sequence heterogeneity can arise from a number of different mechanisms. Classically, it can arise simply by errors during DNA replication [98]. It has been shown that in many types of cancer,

DNA repair mechanisms are dysfunctional, contributing to cancer development in the early stages [99, 100]. A very clear example of this is colorectal cancer (CRC), where mutations in the mismatch repair pathway [101] can lead to the inability of the tumor to correct polymerase slippage errors over tandem repeats [102], causing what has been termed microsatellite instability or MSI [103]. In CRC, MSI occurs in 15-20% of sporadic colorectal tumors and in more than 95% of CRC patients with Lynch syndrome [104], highlighting the role that a single type of repair mutation can play in cancer pathogenesis [105].

Sequence heterogeneity can also arise from the accumulation of point mutations. These are generally aggregated into an index called the Tumor Mutational Burden (TMB - which is simply a measurement of the total number of somatic mutations per tumor sample, often expressed in the number of mutations per Mega base of DNA (mut/Mb)) [106]. A high TMB results in heterogeneity and increased tumor evolution and adaptation, and consequently worse disease prognosis. These mutations were thought to be spontaneous, but there is now clear evidence that instead, most DNA mutations are enzymatically acquired through the action of the AID/APOBEC (and perhaps ADAR) deaminases. For example, there is now concrete evidence that APOBEC-introduced mutational burden enhances intratumor heterogeneity [107], which increases the risk of relapse, drug resistance and immune escape [108].

Finally, sequence heterogeneity can also be generated through abundant RNA modification (which, in the case of both ADARs and APOBECs, is robust and reproducible). Here we have discussed the notion that RNA editing (and the associated *editing burden*) is also important for tumor progression and maintenance. Some editing deaminases (e.g. ADAR1, APOBEC3s) are significantly overexpressed in cancer and target hundreds of transcripts, giving rise to editing signatures [92] that could be important predictors of disease progression. At the same time, whereas RNA modification would not immediately imply a potential "bystander" effect on DNA, we note the possibility that these enzymes have the ability to target both nucleic acids. We therefore hypothesize that editing based transcriptional heterogeneity is intimately linked with mutational signatures in cancer, a hypothesis that remains to be tested.

Author Contributions

These authors contributed equally to this work.

Competing Interests

The authors have declared that no competing interests exist.

References

1. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer. 2013; 108: 479-485.
2. F. Tumor heterogeneity in the clinic: is it a real problem? Ther Adv Med Oncol. 2013; 6: 43-51.
3. Jiang L, Nick A, Sood A. Fundamental principles of cancer biology: Does it have relevance to the perioperative period? Curr Anesthesiol Rep. 2015; 5: 250-256.
4. Efferth T, Leber M. Molecular principles of cancer invasion and metastasis (Review). Int J Oncol. 2009; 34.
5. Knudson A. Mutation and cancer: Statistical study of retinoblastoma. Proc Natl Acad Sci. 1971; 68: 820-823.
6. Martincorena I, Campbell P. Somatic mutation in cancer and normal cells. Science. 2015; 349: 1483-1489.
7. Chalmers Z, Connelly C, Fabrizio D, Gay L, Ali S, Ennis R et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med. 2017; 9.
8. Beck B, Blanpain C. Unravelling cancer stem cell potential. Nat Rev Cancer. 2013; 13: 727-738.
9. Chen J, Miller B, Furano A. Repair of naturally occurring mismatches can induce mutations in flanking DNA. eLife. 2014; 3.
10. Branzei D, Foiani M. Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol. 2008; 9: 297-308.
11. Farmer H, McCabe N, Lord C, Tutt A, Johnson D, Richardson T et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005; 434: 917-921.
12. Papaccio F, Paino F, Regad T, Papaccio G, Desiderio V, Tirino V. Concise review: Cancer cells, cancer stem cells, and mesenchymal stem cells: Influence in cancer development. Stem Cells Transl Med. 2017; 6: 2115-2125.
13. Rozhok A, DeGregori J. Toward an evolutionary model of cancer: Considering the mechanisms that govern the fate of somatic mutations. Proc Natl Acad Sci. 2015; 112: 8914-8921.
14. Tirino V, Desiderio V, Paino F, De Rosa A, Papaccio F, Fazioli F et al. Human primary bone sarcomas contain CD133+ cancer stem cells displaying high tumorigenicity in vivo. FASEB J. 2011; 25: 2022-2030.
15. Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S et al. Accumulation of driver and passenger mutations during tumor progression. Proc Natl Acad Sci. 2010; 107: 18545-18550.
16. Greaves M, Maley C. Clonal evolution in cancer. Nature. 2012; 481: 306-313.
17. Abba M, Drake J, Hawkins K, Hu Y, Sun H, Notcovich C et al. Transcriptomic changes in human breast cancer progression as determined by serial analysis of gene expression. Breast Cancer Res. 2004; 6.
18. Sallam R. Proteomics in Cancer Biomarkers Discovery: Challenges and Applications. Dis Markers. 2015; 2015: 1-12.
19. Stephens P, Tarpey P, Davies H, Van Loo P, Greenman C, Wedge D et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012; 486: 400-404.
20. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: Past, present, and the future. Cell. 2017; 168: 613-628.
21. Alexandrov L, Nik-Zainal S, Wedge D, Aparicio S, Behjati S, Biankin A et al. Signatures of mutational processes in human cancer. Nature. 2013; 500: 415-421.

22. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487: 330-337.

23. Wang J, Cazzato E, Ladewig E, Frattini V, Rosenbloom D, Zairis S et al. Clonal evolution of glioblastoma under therapy. Nat Genet. 2016; 48: 768-776.

24. Peng X, Xu X, Wang Y, Hawke D, Yu S, Han L et al. A-to-I RNA editing contributes to proteomic diversity in cancer. Cancer Cell. 2018; 33: 817-828.e7.

25. Eisenberg E, Levanon E. A-to-I RNA editing — immune protector and transcriptome diversifier. Nat Rev Genet. 2018; 19: 473-490.

26. Harjanto D, Papamarkou T, Oates C, Rayon-Estrada V, Papavasiliou F, Papavasiliou A. RNA editing generates cellular subsets with diverse sequence within populations. Nat Commun. 2016; 7.

27. Picardi E, Horner D, Pesole G. Single-cell transcriptomics reveals specific RNA editing signatures in the human brain. RNA. 2017; 23: 860-865.

28. Conticello S. The AID/APOBEC family of nucleic acid mutators. Genome Biol. 2008; 9: 229.

29. Savva Y, Rieder L, Reenan R. The ADAR protein family. Genome Biol. 2012; 13: 252.

30. Dudley D, Chaudhuri J, Bassing C, Alt F. Mechanism and control of V(D)J recombination versus class switch recombination: Similarities and differences. Adv Immunol. 2005; 43-112.

31. Saraconi G, Severi F, Sala C, Mattiuz G, Conticello S. The RNA editing enzyme APOBEC1 induces somatic mutations and a compatible mutational signature is present in esophageal adenocarcinomas. Genome Biol. 2014; 15.

32. Zheng Y, Lorenzo C, Beal P. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. Nucleic Acids Res. 2017; 3369-3377.

33. Sharma S, Patnaik S, Thomas Taggart R, Kannisto E, Enriquez S, Gollnick P et al. APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. Nat Commun. 2015; 6.

34. Keegan L, Gallo A, O'Connell M. The many roles of an RNA editor. Nat Rev Genet. 2001; 2: 869-878.

35. Blanc V, Davidson N. C-to-U RNA editing: Mechanisms leading to genetic diversity. J Biol Chem. 2002; 278: 1395-1398.

36. Chen SH, Yang CY, Chen PF, Setzer D, Tanimura M, Li WH, Gotto A, Chan L. The complete cDNA amino acid sequence of human Apolipoprotein B-100. J Biol Chem. 1986; 261: 12918-12921.

37. Anant S, Davidson N. Molecular mechanisms of apolipoprotein B mRNA editing. Curr Opin Lipidol. 2001; 12: 159-165.

38. Zheng C, Ikewaki K, Walsh B, Sacks F. Metabolism of apoB lipoproteins of intestinal and hepatic origin during constant feeding of small amounts of fat. J Lipid Res. 2006; 47: 1771-1779.

39. Rosenberg B, Hamilton C, Mwangi M, Dewell S, Papavasiliou F. Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. Nat Struct Mol Biol. 2011; 18: 230-236.

40. Blanc V, Henderson J, Newberry R, Xie Y, Cho S, Newberry E et al. Deletion of the au-rich RNA binding protein apobec-1 reduces intestinal tumor burden in apcmin mice. Cancer Res. 2007; 67: 8565-8573.

41. Fossat N, Tourle K, Radziewic T, Barratt K, Liebhold D, Studdert J et al. C to U RNA editing mediated by APOBEC1 requires RNA-binding protein RBM47. EMBO Rep. 2014; 15: 903-910.

42. Snyder E, McCarty C, Mehalow A, Svenson K, Murray S, Korstanje R et al. APOBEC1 complementation factor (A1CF) is dispensable for C-to-U RNA editing in vivo. RNA. 2017; 23: 457-465.

43. Harris R, Petersen-Mahrt S, Neuberger M. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. Mol Cell. 2002; 10: 1247-1253.

44. Chester A, Somasekaram A, Tzimina M, Jarmuz A, Gisbourne J, O'Keefe, Scott J, Navaratnam N. The apolipoprotein B mRNA editing complex performs a multifunctional cycle and suppresses nonsense-mediated decay. EMBO J. 2003; 22: 3971-3982.

45. Vanharanta S, Marney C, Shu W, Valiente M, Zou Y, Mele A et al. Loss of the multifunctional RNA-binding protein RBM47 as a source of selectable metastatic traits in breast cancer. eLife. 2014; 3.

46. Maul R, Gearhart P. AID and somatic hypermutation. Adv Immunol. 2010; 159-191.

47. Liu M, Duke J, Richter D, Vinuesa C, Goodnow C, Kleinstein S et al. Two levels of protection for the B cell genome during somatic hypermutation. Nature. 2008;451: 841-845.

48. Bransteitter R, Pham P, Calabrese P, Goodman M. Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. J Biol Chem. 2004; 279: 51612-51621.

49. Yu K, Chedin F, Hsieh C, Wilson T, Lieber M. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol. 2003; 4: 442-451.

50. Ronai D, Iglesias-Ussel M, Fan M, Li Z, Martin A, Scharff M. Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. J Exp Med. 2007; 204: 181-190.

51. Longerich S, Basu U, Alt F, Storb U. AID in somatic hypermutation and class switch recombination. Curr Opin Immunol. 2006; 18: 164-174.

52. Zheng S, Vuong B, Vaidyanathan B, Lin J, Huang F, Chaudhuri J. Non-coding RNA generated following lariat debranching mediates targeting of AID to DNA. Cell. 2015; 161: 762-773.

53. Cammas A, Millevoi S. RNA G-quadruplexes: emerging mechanisms in disease. Nucleic Acids Res. 2016; 45: 1584-1595.

54. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. 2000; 102: 553-563.

55. Melchers F. Checkpoints that control B cell development. J Clin Invest. 2015; 125: 2203-2210.

56. Chapman CJ, Zhou JX, Gregory C, Rickinson AB, Stevenson FK. VH and VL gene analysis in sporadic Burkitt's Lymphoma shows somatic hypermutation, intraclonal heterogeneity, and a role for antigen selection. Blood. 1996; 88: 3562-3568.

57. Sale J, Neuberger M. TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. Immunity. 1998; 9: 859-869.

58. Duggal N, Fu W, Akey J, Emerman M. Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. Virology. 2013; 443: 329-337.

59. Borzooee F, Asgharpour M, Quinlan E, Grant M, Larijani M. Viral subversion of APOBEC3s: Lessons for anti-tumor immunity and tumor immunotherapy. Int Rev Immunol. 2017; 37: 151-164.

60. Siriwardena S, Chen K, Bhagwat A. Functions and malfunctions of mammalian DNA-cytosine deaminases. Chem Rev. 2016; 116: 12688-12710.

61. Nik-Zainal S, Alexandrov L, Wedge D, Van Loo P, Greenman C, Raine K et al. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012; 149: 979-993.

62. Sato Y, Probst H, Tatsumi R, Ikeuchi Y, Neuberger M, Rada C. Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. J Biol Chem. 2009; 285: 7111-7118.

63. Vonica A, Rosa A, Arduini B, Brivanlou A. APOBEC2, a selective inhibitor of TGFβ signaling, regulates left–right axis specification during early embryogenesis. Dev Biol. 2011; 350: 13-23.

64. Conticello S, Langlois M, Yang Z, Neuberger M. DNA deamination in immunity: AID in the context of its APOBEC relatives. Adv Immunol. 2007; 94: 37-73.

65. Rogozin I, Basu M, Jordan I, Pavlov Y, Koonin E. APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases predicted by computational analysis. Cell Cycle. 2005; 4: 1281-1285.

66. Salter J, Bennett R, Smith H. The APOBEC protein family: United by structure, divergent in function. Trends Biochem Sci. 2016; 41: 578-594.

67. Zinshteyn B, Nishikura K. Adenosine-to-inosine RNA editing. WIRES Syst Biol Med. 2009; 1: 202-209.

68. Samuel C. ADARs: Viruses and Innate Immunity. Curr Top Microbiol Immunol. 2011; 163-195.

69. Orecchini E, Doria M, Antonioni A, Galardi S, Ciafrè S, Frassinelli L et al. ADAR1 restricts LINE-1 retrotransposition. Nucleic Acids Res. 2016; 45: 155-168.

70. Higuchi M, Maas S, Single F, Hartner J, Rozov A, Burnashev N et al. Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature. 2000;406(6791):78-81.

71. Liscovitch-Brauer N, Alon S, Porath H, Elstein B, Unger R, Ziv T et al. Trade-off between transcriptome plasticity and genome evolution in cephalopods. Cell. 2017; 169: 191-202.e11.

72. Athanasiadis A, Rich A, Maas S. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol. 2004; 2: e391

73. Montiel-Gonzalez M, Vallecillo-Viejo I, Yudowski G, Rosenthal J. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. Proc Natl Acad Sci. 2013; 110: 18285-18290.

74. Cox D, Gootenberg J, Abudayyeh O, Franklin B, Kellner M, Joung J et al. RNA editing with CRISPR-Cas13. Science. 2017; 358: 1019-1027.

75. Rebhandl S, Huemer M, Greil R, Geisberger R. AID/APOBEC deaminases and cancer. Oncoscience. 2015; 2: 320.

76. Okazaki I, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K et al. Constitutive expression of AID leads to tumorigenesis. J Exp Med. 2003; 197: 1173-1181.

77. Starrett G, Luengas E, McCann J, Ebrahimi D, Temiz N, Love R et al. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. Nat Commun. 2016; 7.

78. D'Antonio M, Tamayo P, Mesirov J, Frazer K. Kataegis expression signature in breast cancer is associated with late onset, better prognosis, and higher HER2 levels. Cell Rep. 2016; 16: 672-683.

79. Henderson S, Fenton T. APOBEC3 genes: retroviral restriction factors to cancer drivers. Trends Mol Med. 2015; 21: 274-284.

80. Pettersen H, Galashevskaya A, Doseth B, Sousa M, Sarno A, Visnes T et al. AID expression in B-cell lymphomas causes accumulation of genomic uracil and a distinct AID mutational signature. DNA Repair. 2015; 25: 60-71.

81. Teater M, Dominguez P, Redmond D, Chen Z, Ennishi D, Scott D et al. AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis. Nat Commun. 2018; 9.

82. Chambwe N, Kormaksson M, Geng H, De S, Michor F, Johnson N et al. Variability in DNA methylation defines novel epigenetic subgroups of DLBCL associated with different clinical outcomes. Blood. 2014; 123: 1699-1708.

83. De S, Shaknovich R, Riester M, Elemento O, Geng H, Kormaksson M et al. Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. PLoS Genet. 2013; 9: e1003137.

84. Dominguez P, Shaknovich R. Epigenetic function of activation-induced cytidine deaminase and its link to lymphomagenesis. Front Immunol. 2014; 5.

85. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). Cell. 2000; 102: 565-575.

86. Yamanaka S, Balestra M, Ferrell L, Fan J, Arnold K, Taylor S et al. Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. Proc Natl Acad Sci. 1995; 92: 8483-8487.

87. Mukhopadhyay D, Anant S, Lee R, Kennedy S, Viskochil D, Davidson N. C→U editing of neurofibromatosis 1 mRNA occurs in tumors that express both the type II transcript and apobec-1, the catalytic subunit of the apolipoprotein B mRNA–Editing enzyme. Am J Hum Genet. 2002; 70: 38-50.

88. Moser AR, Luongo C, Gould KA, McNeley MK, Shoemaker AR, Dove WF. ApcMin: a mouse model for intestinal and mammary tumorigenesis. Eur J Cancer. 1995; 31A: 1061-1064.

89. Nelson V, Heaney J, Tesar P, Davidson N, Nadeau J. Transgenerational epigenetic effects of the Apobec1 cytidine deaminase deficiency on testicular germ cell tumor susceptibility and embryonic viability. Proc Natl Acad Sci. 2012; 109: E2766-E2773.

90. Baysal B, Sharma S, Hashemikhabir S, Janga S. RNA editing in pathogenesis of cancer. Cancer Res. 2017; 77: 3733-3739.

91. Laurencikiene J, Källman A, Fong N, Bentley D, Öhman M. RNA editing and alternative splicing: the importance of co-transcriptional coordination. EMBO Rep. 2006.

92. Paz-Yaacov N, Bazak L, Buchumenski I, Porath H, Danan-Gotthold M, Knisbacher B et al. Elevated RNA editing activity is a major contributor to transcriptomic diversity in tumors. Cell Rep. 2015; 13: 267-276.

93. Fumagalli D, Gacquer D, Rothé F, Lefort A, Libert F, Brown D et al. Principles governing A-to-I RNA editing in the breast cancer transcriptome. Cell Rep. 2015; 13: 277-289.

94. Chen L, Li Y, Lin C, Chan T, Chow R, Song Y et al. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. Nat Med. 2013; 19: 209-216.

95. Lazzari E, Mondala P, Santos N, Miller A, Pineda G, Jiang Q et al. Alu-dependent RNA editing of GLI1 promotes malignant regeneration in multiple myeloma. Nat Commun. 2017; 8.

96. Chen W, He W, Cai H, Hu B, Zheng C, Ke X et al. A-to-I RNA editing of BLCAP lost the inhibition to STAT3 activation in cervical cancer. Oncotarget. 2017; 8.

97. Nakajima Y, Kanno T, Nagaya T, Kuribayashi K, Nakano T, Gotoh A et al. Adenosine deaminase inhibitor EHNA exhibits a potent anticancer effect against malignant pleural mesothelioma. Cell Physiol Biochem. 2015; 35: 51-60.

98. Robertson A, Klungland A, Rognes T, Leiros I. DNA Repair in Mammalian Cells. Cell Mol Life Sci. 2009; 66: 981-993.

99. Goode E, Ulrich C, Potter J. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev. 2002; 11: 1513-1530.

100. Helleday T, Petermann E, Lundin C, Hodgson B, Sharma R. DNA repair pathways as targets for cancer therapy. Nat Rev Cancer. 2008; 8: 193-204.

101. Schloetterer C, Harr B. Microsatellite Instability. Encyclopedia of Life Sciences. 2004

102. Kim T, Laird P, Park P. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. Cell. 2013; 155: 858-868.

103. Hause R, Pritchard C, Shendure J, Salipante S. Classification and characterization of microsatellite instability across 18 cancer types. Nat Med. 2016; 22: 1342-1350.

104. Grady W, Carethers J. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology. 2008; 135: 1079-1099.

105. Chang L, Chang M, Chang H, Chang F. Expending role of microsatellite instability in diagnosis and treatment of colorectal cancers. J Gastrointest Cancer. 2017; 48: 305-313.

106. Goodman A, Kato S, Bazhenova L, Patel S, Frampton G, Miller V et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. Mol Cancer Ther. 2017; 16: 2598-2608.

107. de Bruin E, McGranahan N, Mitter R, Salm M, Wedge D, Yates L et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science. 2014; 346: 251-256.

108. Venkatesan S, Rosenthal R, Kanu N, McGranahan N, Bartek J, Quezada S et al. Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. Ann Oncol. 2018; 29: 563-572.
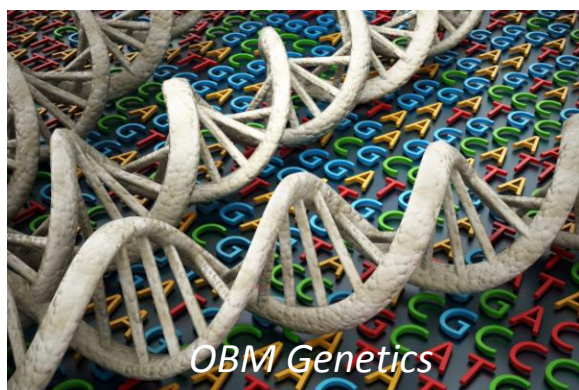
Enjoy *OBM Genetics* by:

1. Submitting a manuscript
2. Joining in volunteer reviewer bank
3. Joining Editorial Board
4. Guest editing a special issue

For more details, please visit:
http://www.lidsen.com/journals/genetics