Opinion

Cell PRESS

Crosstalk between genetic and epigenetic information through cytosine deamination

Richard Chahwan, Sandeep N. Wontakal and Sergio Roa

Department of Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA

Decades of work have elucidated the existence of two forms of heritable information, namely genetic and epigenetic, which are collectively referred to as the 'dual inheritance'. The underlying mechanisms behind these two modes of inheritance have so far remained distinct. Cytosine deaminases, such as activation-induced cytidine deaminase (AID) and other members of the APOBEC family, have been implicated both in genetic variation of somatic cells and in epigenetic remodeling of germ and pluripotent cells. We hereby synthesize these seemingly dissociated functions into one coherent model, and further suggest that cytosine deaminases, particularly AID, might have a broader influence by modulating epigenetic information in somatic or cancer cells, as well as by triggering genetic variation in germ and pluripotent cells through mutation followed by natural selection. We therefore propose that the AID/APOBEC family of deaminases are likely to have acted as drivers throughout vertebrate evolution.

The dual inheritance

The variation and transmissibility of heritable information is perhaps the primordial principle that drives the evolution of species. It was already clear from the chromosomal theory of inheritance and the subsequent identification of DNA as the heritable material that mutations in DNA can be passed to offspring and propagated within a population under natural selection (i.e. genetic inheritance). However, several phenomena [1–3] have been described over the past decades where phenotypes are transmitted without any underlying changes in the DNA sequence (i.e. epigenetic inheritance), challenging this basic tenet of biology [4]. Because inheritance of both genetic and epigenetic information – referred to as the 'dual inheritance' [5] – determines the phenotype, molecule(s) that could modify both of these forms of heritable information could be important drivers of evolutionary adaptability. However, such a molecule remained elusive until cytosine deaminases, such as AID and other members of the APOBEC family (Glossary), emerged as potential candidates [6–11].

The past decade of research in the fields of immunology and cell biology has significantly advanced our understanding of the genetic role of AID in somatic diversification of B cells and in cancer [12]. Concurrently, the initial idea that AID and APOBEC1 might also deaminate 5'-methyl-cytosines (5meC) [9] has been reinforced by recent advances implicating the AID/APOBEC family in epigenetic remodeling [6–8]. However, the genetic and epigenetic roles remained segregated, with no evidence for AID/APOBEC-mediated demethylation in somatic tissues or AID/APOBEC-generated mutations in germ and pluripotent cells. Here, we hypothesize that because cytosine deaminases, and specifically AID, are unique in that they have the capacity to modify both forms of heritable information, they have the potential to connect genetic and epigenetic inheritance. Therefore, we propose that the AID/APOBEC family of deaminases could be important drivers of evolutionary adaptability in vertebrates.

AID/APOBEC and the modulation of genetic information

The universal function of the proof-reading and repair mechanisms appears to have evolved to prevent deleterious mutations and to maintain genomic stability. Yet, paradoxically, some cellular processes actually require the acquisition of mutations. One such phenomenon is somatic hypermutation (SHM) of B cell immunoglobulin (Ig) genes in the germinal centers of secondary lymphoid

Glossary

Activation-induced cytidine deaminase (AID): a member of the APOBEC family that was first identified in a screen for genes involved in antibody diversification in a murine B-cell line. AID was initially thought to function, like APOBEC1, in editing RNA. However, another model ascribed its function to direct cytosine deamination for DNA, primarily at the Immunoglobulin (Ig) locus in B cells.
APOBEC: a family of zinc-dependent cytosine deaminases named after the first identified member APOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1) which was shown to edit apolipoprotein B mRNA. The APOBEC family comprises 5 groups (AID and APOBEC1–4), with each group containing one or more members. Several members of the APOBEC family have been shown to deaminate cytosines in DNA.
D-demethylation: a term used here to describe the indirect nature of active 5meC demethylation through a deamination (D)-mediated process.
Dual inheritance: a term previously coined to describe the transfer of both genetic and epigenetic information.
Epigenetic inheritance: the transfer of heritable information without an underlying change in the DNA sequence. Epigenetic information can take the form of protein, RNA, or modifications of DNA, such as cytosine methylation.
Genetic inheritance: the transfer of heritable information contained within the DNA sequence (genetic information).
Somatic hypermutation (SHM): an AID-dependent process whereby Ig gene variable regions are subject to a high rate of mutation. SHM, in B cells, leads to diversification of antibody structure and permits the selection of high-affinity antibodies.

Corresponding authors: Chahwan, R. (richard.chahwan@einstein.yu.edu); Wontakal, S.N. (sandeep.wontakal@med.einstein.yu.edu); Roa, S. (sergio.roa@einstein.yu.edu)

0168-9525/$ – see front matter doi:10.1016/j.tig.2010.07.005 Trends in Genetics 26 (2010) 443–448

organs [13–15]. This genetic process allows the maturation and diversification of antibody affinity through the accumulation of mutations in Ig gene variable regions, and these occur at frequencies of 10^{-5} to 10^{-3} per nucleotide and are subject to clonal selection [16,17]. Initial investigations in the murine B-cell line CH12F3 [18], in AID-deficient mice [19], and in humans with hyper-IgM immunodeficiency [20] led to the characterization of AID as the mutagenic enzyme responsible for initiating the genetic modifications required for an efficient immune response [12,13,15].

Although it was originally hypothesized that AID edits RNA [18,21], a number of groups have provided evidence to support a DNA-editing model [13,22–29]. This model illustrates how B cells have efficiently adopted cytosine deamination into a mechanism to generate DNA mismatches in order to trigger genetic variation (Figure 1a). Any AID-targeted cytosine (C) can be readily deaminated to form uracil (U) [30], turning C:G pairs into U:G mispairs that can be detected and repaired in an error-free manner by the base excision-repair (BER) and/or mismatch-repair (MMR) machinery [13,15,31]. However, U:G mispairs can also escape detection and be replicated to produce transitions (T:A in Figure 1), or can be processed through a mutagenic pathway that recruits error-prone polymerases to produce transitions or transversions at the deaminated site (N:N in Figure 1) [24,32,33]. It is therefore possible that any cell that expresses AID could also be subject to deamination-mediated genetic mutations. Indeed, ectopic expression of AID in mammalian cells [34,35], yeast [36], or even bacteria [23] can lead to DNA mutations. Similarly, other members of the APOBEC family have been implicated in deaminating cytosine, leading to mutations in DNA [10,11,37]. Moreover, several studies have shown that AID can act as a broad genomic mutator [38], and its inappropriate expression in various tissues has been linked to tumorigenesis [39–43]. Interestingly, the tight regulation of AID expression and its mutagenic activity can also be compromised in B cells and has been associated with lymphomagenesis [44–51].

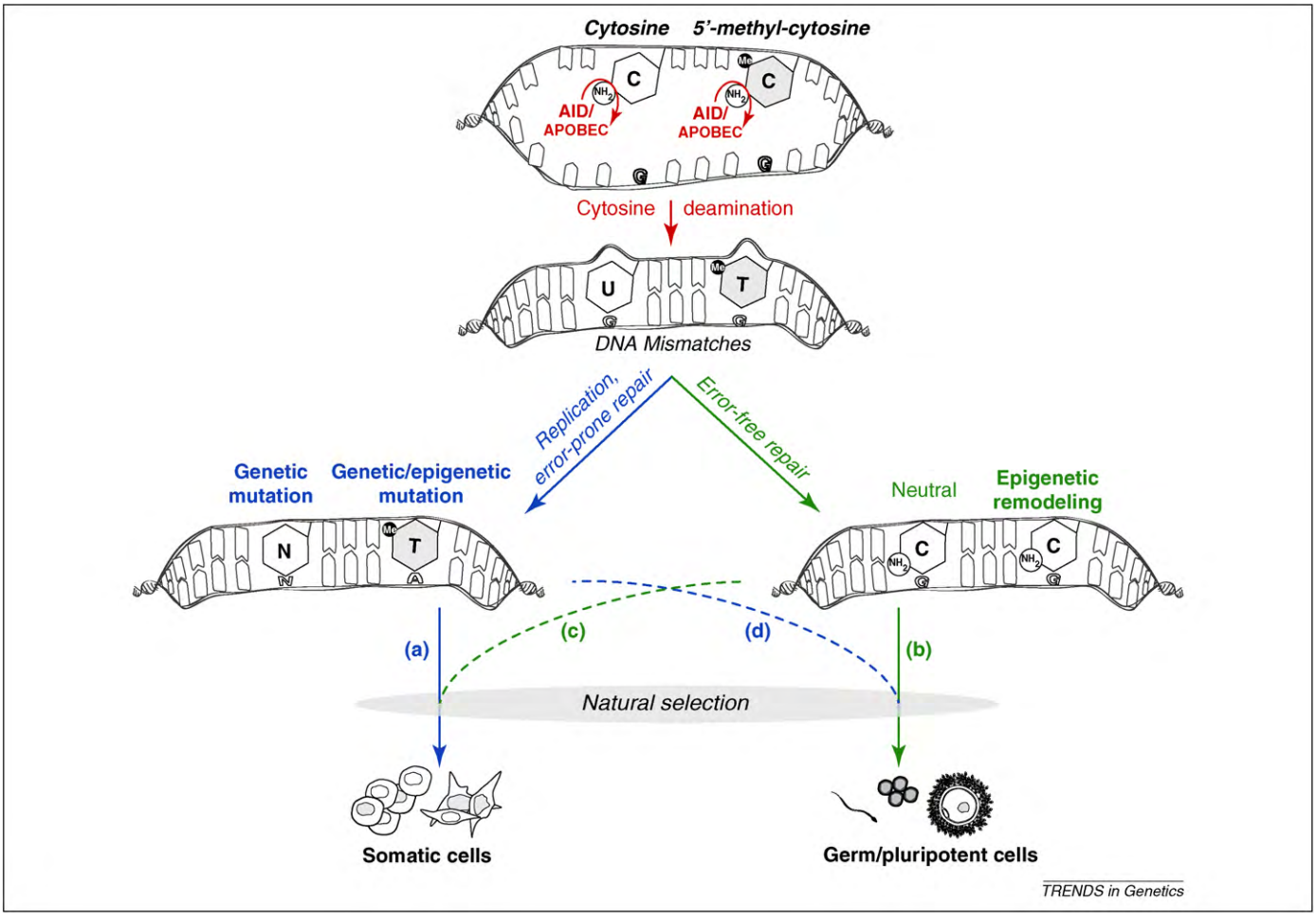

Figure 1. A unified paradigm for cytosine deamination in the modulation of genetic and epigenetic information. This model depicts how AID/APOBEC can function on exposed ssDNA to deaminate cytosines (white C) or 5′-methyl-cytosines (grey C), thereby creating U:G or T:G mismatches in the DNA. **(a)** Replication of uncorrected mismatches or error-prone repair has been shown to produce mutations (transitions or transversions) at deaminated sites in somatic cells. These can be classified as either genetic mutations (C-to-N) or genetic/epigenetic mutations (5meC-to-T). **(b)** Error-free repair of mismatches results in the incorporation of unmethylated Cs, and these can have either neutral effects (C-to-C) or lead to epigenetic remodeling (5meC-to-C). The role of AID/APOBEC as DNA D-demethylases has been suggested to result in epigenetic remodeling of germ/pluripotent cells. **(c)** It is plausible to assume a role for cytosine deamination in epigenetic remodeling of somatic cells that express AID/APOBEC. **(d)** We further postulate that AID/APOBEC expression in germ and pluripotent cells could also be a source of deamination-induced mutations that can impact upon the germinal transmission of genetic information. The model assumes that any genetic or epigenetic variation triggered by AID/APOBEC is subject to selection pressures that dictate the transmissibility of such information.

AID/APOBEC and the modulation of epigenetic information

As mentioned above, several phenomena have been described where the transmission of heritable information – between somatic or germ cells – takes place without underlying changes to the DNA sequence, referred to as epigenetic inheritance. Unlike the genetic material, epigenetic information can be stored and transmitted in the form of any of the three macromolecules of the cell (i.e. DNA, RNA and protein) [4]. One mechanism for epigenetic inheritance involves DNA methylation [4,52]. In this respect, cytosines in DNA have the remarkable capacity to hold genetic as well as epigenetic information in the form of methylation [52]. In vertebrates, methylation of cytosines is normally associated with CpG dinucleotides [53]. The presence or absence of 5'-methyl-cytosine (5meC) can lead to changes in gene expression, and thus alterations in the methylation pattern can result in phenotypic changes without affecting the actual DNA sequence [52,54]. Studies during embryogenesis have shown that gametes contain relatively high levels of DNA methylation, but global demethylation takes place after fertilization, and eventually methylation patterns are re-established in somatic tissues [54]. This cycle strongly suggests that there are mechanisms to both 'write' and 'erase' methylated DNA signatures. Despite the successful identification and characterization of enzymes involved in methylating DNA, the molecular mechanisms involved in DNA demethylation have remained elusive in vertebrates.

One way to demethylate 5meC is through a passive mechanism involving DNA replication. Although the evidence suggests that passive DNA demethylation through replication does take place [55,56], it is unlikely to be the only mechanism because there are several examples of replication-independent demethylation [9,57–59]. A number of mechanisms have therefore been postulated to account for an active demethylation process [52,54]. One such mechanism postulates an indirect enzymatic model whereby 5meC is deaminated [9,52,54]. However, deaminated 5meC is, in fact, thymine (T). Such a conversion results in a T:G mispair, and this can subsequently be recognized by the DNA repair machinery and converted back to a C:G pair by incorporating an unmethylated C residue (Figure 1). To distinguish this from other mechanisms of active demethylation [52,54] we hereafter use the term D-demethylation to specifically refer to deamination-mediated demethylation.

Previous work showed that AID and APOBEC1 can deaminate 5meC *in vitro*, suggesting that the AID/APOBEC family of cytosine deaminases can act as DNA D-demethylases [9]. In support of this observation, a recent study reported that both AID and APOBEC2a/b can demethylate DNA in zebrafish embryos [7]. Although no direct biochemical activity has yet been ascribed to APOBEC2a/b, the fact that mutants in the conserved putative catalytic domain are unable to induce demethylation suggests that these proteins could have the ability to deaminate 5meC [7]. This function requires two additional components: MBD4, a DNA glycosylase that can act on T:G mismatches, and GADD45 family members which are induced by DNA damage. Overexpression and knockdown experiments showed that these proteins not only deaminate exogenous methylated DNA but also function on endogenous loci. Furthermore, these proteins seem to interact with one another in heterologous cells and can localize to exogenous as well as endogenous loci that undergo active demethylation. This study not only provided the first putative mechanism of active DNA demethylation in higher eukaryotes, but it also highlighted the importance of the DNA repair machinery in this process [7,9].

Further evidence suggesting that AID can modulate epigenetic information through D-demethylation comes from studies of cellular reprogramming of interspecies heterokaryons formed by the fusion of mouse embryonic stem cells with human fibroblasts [6]. A prerequisite for cellular reprogramming is the activation of the pluripotency genes *OCT4* and *NANOG* which are normally silenced in somatic cells through DNA methylation. AID was shown to localize to the promoters of *OCT4* and *NANOG* in human fibroblasts [6]. Subsequent D-demethylation during heterokaryon formation was demonstrated to be replication-independent and to require AID. In parallel, another report showed an increase in global DNA methylation in primordial germ cells (PGCs) of AID knockout mice, including modest changes in imprinted loci and transposable elements [8]. However, AID deficiency seemed to only affect the level of methylation in PGCs. We believe it would be of great interest to explore if the effect on other cell types could be accentuated by either the loss of AID through multiple generations or by the combinatorial loss of AID along with potential compensating APOBEC family member(s). Intriguingly, there seems to be an effect on the litter sizes of the F1 generation from AID-deficient homozygous mice. It is tantalizing to speculate, as suggested by the authors [8], that the changes in litter size might be due to epigenetic modifications in the PGCs of AID-deficient mice, although there is as yet no direct evidence. Taken together, these studies establish a role for AID and other APOBEC family members in DNA D-demethylation, and thereby highlight the potential of the AID/APOBEC family in modulating epigenetic information (Figure 1b,c).

Crosstalk between the genetic and epigenetic functions of AID/APOBEC

The recognition of AID/APOBEC as components of an active D-demethylase complex through their ability to deaminate 5meCpG in germ and pluripotent cells [7,8] raises the possibility that this mechanism could also function in somatic cells (Figure 1c). Indeed, this hypothesis is supported by data showing AID activity on 5meC in bacteria and *in vitro* [9], and is consistent with recent work in heterokaryons demonstrating a role for AID in cell reprogramming through D-demethylation of the *OCT4* and *NANOG* promoters [6]. It is therefore reasonable to postulate a role for cytosine deamination in the epigenetic remodeling of any somatic cells that appropriately (e.g. B cells) or inappropriately (e.g. tumor cells) express a member(s) of the AID/APOBEC family. Such somatic remodeling can have either beneficial consequences if AID/APOBEC-dependent D-demethylation is physiologically

induced and tightly regulated, or deleterious consequences if inappropriate expression or regulation of AID/APOBEC pushes the epigenetic remodeling towards malignancy. Indeed, a number of reports have shown that AID plays a role in tumorigenesis [39–51]. Therefore, it would be interesting to study if changes in the DNA methylation profiles of somatic tissues and cancer cells are AID/APOBEC-dependent (Figure 1c). Interestingly, it has been suggested that AID genetic and epigenetic mutations at CpG sites might lead to premature stop codons of the *APC* tumor-suppressor gene in some colorectal cancers [9] (Figure 1a).

Although AID can deaminate both C and 5meC, it exhibits 3-to-10-fold lower specific activity for methylated cytosines *in vitro* [9,27,60]. Despite this, AID has been implicated in the deamination of 5meC in different cellular contexts [6–9]. D-demethylation will occur when this deamination event is coupled to error-free repair, perhaps involving the DNA repair proteins GADD45, MBD4, and/or XPG [7,61], and converts the T back into C, thereby demethylating a 5meC:G pair into a C:G pair (Figure 1). However, consistent with the mutator role of AID discussed above, a T:G mispair can be erroneously converted into a T:A mutated pair (Figure 1). If this occurs, the deaminated DNA sequence would lose the original 5meCpG and gain a TpG dinucleotide. Indeed, CpG depletion with TpG gain is a hallmark of many vertebrate genomes that undergo DNA methylation, and this seems to be present in species as far back as zebrafish and fugu (Refs [62,63] and J. Glass, personal communication). This finding suggests that D-demethylation is a significant mode of demethylating 5meC in vertebrates. Moreover, TpG gain is also of particular importance because 5meC loss alters not only epigenetic information but also the underlying genetic information. Interestingly, the AID/APOBEC family of cytosine deaminases seems to have originated in the vertebrate lineage [37,64]. Because the AID/APOBEC family is thus far the only known group of cytosine deaminases in vertebrates [37], and given the coincidence of the evolutionary emergence of the family along with CpG depletion, it is possible to postulate that members of the AID/APOBEC family of deaminases are, at least in part, responsible for the phenomenon of CpG depletion with TpG gain (Figure 1d). If this is true, the AID/APOBEC family of deaminases expressed in the germline could have acted as evolutionary drivers throughout vertebrate evolution.

Undoubtedly, excessive DNA mutation can be detrimental to an organism, but mutations can also confer a marked advantage under some conditions. A good example of this is APOBEC3G, which normally confers resistance against viral infections (such as HIV) through its deaminase activity [11]. The HIV protein termed VIF degrades APOBEC3G and thereby counteracts its effects. Even so, a naturally occurring attenuated VIF can be permissive for APOBEC3G-mediated mutations, paradoxically allowing HIV to acquire drug resistance [11]. As with AID in B cells, this example supports the idea that balanced enzymatic mutations by deaminases could confer a marked evolutionary advantage through natural selection. The intriguing fact that AID/APOBEC enzymes are expressed not only in somatic cells but also in germ and pluripotent cells [6,7,9,37,65] raises the possibility of a common enzymatic behavior for the AID/APOBEC family in all these scenarios. One should not disregard the possibility that, analogous to what happens at the immunoglobulin locus in B cells, AID expression in germ and pluripotent cells could also induce deamination-mediated mutations. These mutations would subsequently be subject to selective pressures and result in genetic variation during germline inheritance (Figure 1d). In fact, a recent computational study suggested that the evolution of protein–protein interfaces shares some similarities with AID-mediated diversification of antibodies by SHM [66]. Moreover, a recent analysis of the mouse and human genomes identified mutations in retrotransposable elements that are reminiscent of AID/APOBEC-induced mutations [67], consistent with the idea that this family of deaminases might act as defenders of the genome by mutating parasitic genetic elements [67,68]. It would be of great interest to analyze the genomes of other vertebrate species to see if a similar phenomenon is present. In addition, a link between AID-induced deamination and meiotic recombination, which is thought to be the main source of genetic diversification during germline inheritance, has been recently described [69]. Consistent with this notion, some AID/APOBEC family member(s) are expressed in oocytes and testis [9,37,65], suggesting a potential role for the mutagenic activity of AID/APOBEC during the germinal transmission of genetic information. Therefore, if this hypothesis holds true, we would predict an enrichment of mutations resulting from deamination of C or 5meC in the genome of species that express AID/APOBEC, and these would then be subject to natural selection, potentially leading to a marked advantage for the organism.

Open questions and future perspectives

Insights from the mutagenic role of AID in B cells strongly support the involvement of transcription in its function [15]. In fact, AID has a binding preference for single-stranded DNA (ssDNA) and binds very poorly to double-stranded DNA (dsDNA) [25–29]. For this reason the generation of ssDNA through transcription of Ig variable or switch regions is thought to be an important prerequisite for AID function during antibody diversification in B cells [12,15]. The highly repetitive Ig switch regions are characterized by the transcription-dependent formation of stable RNA:DNA hybrids known as R-loops [70] which expose the non-transcribed DNA strand to AID-mediated deamination. Paradoxically, the observation in heterokaryons that the methylated and initially untranscribed human *NANOG* and *OCT4* loci could be prone to AID-mediated deamination indicates that transcription might not be associated with AID-mediated epigenetic reprogramming. Another obvious alternative to generate ssDNA would be replication. However, both active demethylation [57–59] and D-demethylation of the *NANOG* and *OCT4* promoters has been shown to take place in the absence of replication [6]. This hints that it might be possible to generate ssDNA through a yet unknown mechanism, perhaps involving basal unproductive transcription [71,72], G-quadruplex DNA [73], or other DNA-unwinding activity at methylated CpG sites.

The ability of AID/APOBEC to modulate both genetic and epigenetic information sets forth a slew of fundamental questions. For example, how does AID/APOBEC differentiate between DNA D-demethylation and DNA mutation? In other words, could AID/APOBEC make the choice between genetic and epigenetic modifications? And, if so, on what grounds is the selection made? Is it cell-type specificity, expression levels of AID/APOBEC, or targeting and/or chaperoning factors that can modulate AID/APOBEC activity accordingly? In addition to the aforementioned DNA-mediated mechanism, it was originally postulated that AID could also function through an RNA-mediated mechanism [18,21]. Although this remains controversial, if it holds true then this might offer another layer of complexity to AID function. In either case, future studies on the potential role of the DNA repair machinery in mediating AID/APOBEC activity might help to shed more light on these questions.

Concluding remarks

At first glance, genetic variation and epigenetic reprogramming seem to be two distinct processes. However, looking at them through the AID/APOBEC lens the two seem to be intertwined, akin to opposite faces of the same coin. Because AID/APOBEC has the capacity to modulate both the genetic and epigenetic programs it is possible that they could alter both forms of information storage in a given cell; whether it be by mutating and perhaps demethylating DNA in somatic cells (Figure 1a,c) or by demethylating and perhaps mutating DNA in germ or pluripotent cells (Figure 1b,d). Whereas the former might be involved in normal cell physiology and possibly tumorigenesis, the latter scenario is of particular significance because germ cells will eventually give rise to future progeny and thus, by modifying both the genetic and epigenetic programs in such cells, AID/APOBEC has the unique potential to influence the dual forms of inheritance. It is therefore plausible that the AID/APOBEC family might embody the molecular convergence of evolutionary adaptability by both genetic variation and epigenetic inheritance. The latter, seemingly a short-term fix, could occur through heritable epigenetic modifications, whereas the former might offer a longer-term advantage through genetic mutability of germ cells followed by rounds of natural selection.

Acknowledgements

We are grateful to the Albert Einstein College of Medicine community and especially to our mentors Matthew Scharff and Arthur Skoultchi for their support, enthusiasm, and encouragement. We also thank Alberto Moldón, Jacob Glass, Thomas MacCarthy, Aviv Bergman, and The Relativity Club for stimulating and insightful discussions.

References

1 Prusiner, S.B. (1982) Novel proteinaceous infectious particles cause scrapie. *Science* 216, 136–144
2 Buiting, K. *et al.* (2003) Epimutations in Prader-Willi and Angelman syndromes: a molecular study of 136 patients with an imprinting defect. *Am. J. Hum. Genet.* 72, 571–577
3 Rassoulzadegan, M. *et al.* (2006) RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 441, 469–474
4 Jablonka, E. and Raz, G. (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* 84, 131–176
5 Maynard Smith, J. (1990) Models of a dual inheritance system. *J. Theor. Biol.* 143, 41–53
6 Bhutani, N. *et al.* (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 463, 1042–1047
7 Rai, K. *et al.* (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* 135, 1201–1212
8 Popp, C. *et al.* (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463, 1101–1105
9 Morgan, H.D. *et al.* (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* 279, 52353–52360
10 Harris, R.S. *et al.* (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* 10, 1247–1253
11 Mulder, L.C. *et al.* (2008) Cytidine deamination induced HIV-1 drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5501–5506
12 Delker, R.K. *et al.* (2009) A coming-of-age story: activation-induced cytidine deaminase turns 10. *Nat. Immunol.* 10, 1147–1153
13 Di Noia, J.M. and Neuberger, M.S. (2007) Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* 76, 1–22
14 MacLennan, I.C. (2005) Germinal centers still hold secrets. *Immunity* 22, 656–657
15 Peled, J.U. *et al.* (2008) The biochemistry of somatic hypermutation. *Annu. Rev. Immunol.* 26, 481–511
16 Rajewsky, K. *et al.* (1987) Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science* 238, 1088–1094
17 Neuberger, M.S. (2008) Antibody diversification by somatic mutation: from Burnet onwards. *Immunol. Cell Biol.* 86, 124–132
18 Muramatsu, M. *et al.* (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274, 18470–18476
19 Muramatsu, M. *et al.* (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553–563
20 Revy, P. *et al.* (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* 102, 565–575
21 Shivarov, V. *et al.* (2009) Molecular mechanism for generation of antibody memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 569–575
22 Poltoratsky, V. *et al.* (2000) Error prone candidates vie for somatic mutation. *J. Exp. Med.* 192, F27–F30
23 Petersen-Mahrt, S.K. *et al.* (2002) AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418, 99–103
24 Rada, C. *et al.* (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* 12, 1748–1755
25 Pham, P. *et al.* (2003) Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 424, 103–107
26 Dickerson, S.K. *et al.* (2003) AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med.* 197, 1291–1296
27 Bransteitter, R. *et al.* (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4102–4107
28 Ramiro, A.R. *et al.* (2003) Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat. Immunol.* 4, 452–456
29 Chaudhuri, J. *et al.* (2003) Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* 422, 726–730
30 Kavli, B. *et al.* (2007) Uracil in DNA – general mutagen, but normal intermediate in acquired immunity. *DNA Repair* 6, 505–516
31 Visnes, T. *et al.* (2009) Uracil in DNA and its processing by different DNA glycosylases. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 563–568
32 Saribasak, H. *et al.* (2009) Hijacked DNA repair proteins and unchained DNA polymerases. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 605–611
33 Jansen, J.G. *et al.* (2006) Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. *J. Exp. Med.* 203, 319–323

34 Yoshikawa, K. *et al.* (2002) AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* 296, 2033–2036
35 Martin, A. and Scharff, M.D. (2002) Somatic hypermutation of the AID transgene in B and non-B cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 12304–12308
36 Rogozin, I.B. and Pavlov, Y.I. (2006) The cytidine deaminase AID exhibits similar functional properties in yeast and mammals. *Mol. Immunol.* 43, 1481–1484
37 Conticello, S.G. (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* 9, 229
38 Liu, M. *et al.* (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451, 841–845
39 Okazaki, I.M. *et al.* (2003) Constitutive expression of AID leads to tumorigenesis. *J. Exp. Med.* 197, 1173–1181
40 Morisawa, T. *et al.* (2008) Organ-specific profiles of genetic changes in cancers caused by activation-induced cytidine deaminase expression. *Int. J. Cancer* 123, 2735–2740
41 Lin, C. *et al.* (2009) Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* 139, 1069–1083
42 Matsumoto, Y. *et al.* (2007) *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat. Med.* 13, 470–476
43 Endo, Y. *et al.* (2007) Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene* 26, 5587–5595
44 Pasqualucci, L. *et al.* (2001) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412, 341–346
45 Greeve, J. *et al.* (2003) Expression of activation-induced cytidine deaminase in human B-cell non-Hodgkin lymphomas. *Blood* 101, 3574–3580
46 Smit, L.A. *et al.* (2003) Expression of activation-induced cytidine deaminase is confined to B-cell non-Hodgkin's lymphomas of germinal-center phenotype. *Cancer Res.* 63, 3894–3898
47 Pasqualucci, L. *et al.* (2008) AID is required for germinal center-derived lymphomagenesis. *Nat. Genet.* 40, 108–112
48 Takizawa, M. *et al.* (2008) AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J. Exp. Med.* 205, 1949–1957
49 Klemm, L. *et al.* (2009) The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. *Cancer Cell* 16, 232–245
50 Komeno, Y. *et al.* (2010) AID-induced T-lymphoma or B-leukemia/lymphoma in a mouse BMT model. *Leukemia* 24, 1018–1024
51 Ramiro, A.R. *et al.* (2006) Role of genomic instability and p53 in AID-induced *c-myc-Igh* translocations. *Nature* 440, 105–109
52 Zhu, J.K. (2009) Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* 43, 143–166
53 Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21
54 Morgan, H.D. *et al.* (2005) Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* 14 (Suppl 1), R47–58
55 Howlett, S.K. and Reik, W. (1991) Methylation levels of maternal and paternal genomes during preimplantation development. *Development* 113, 119–127
56 Carlson, L.L. *et al.* (1992) Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* 6, 2536–2541
57 Oswald, J. *et al.* (2000) Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* 10, 475–478
58 Mayer, W. *et al.* (2000) Demethylation of the zygotic paternal genome. *Nature* 403, 501–502
59 Bruniquel, D. and Schwartz, R.H. (2003) Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat. Immunol.* 4, 235–240
60 Larijani, M. *et al.* (2005) Methylation protects cytidines from AID-mediated deamination. *Mol. Immunol.* 42, 599–604
61 Barreto, G. *et al.* (2007) Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671–675
62 Bird, A.P. (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* 8, 1499–1504
63 Glass, J.L. *et al.* (2007) CG dinucleotide clustering is a species-specific property of the genome. *Nucleic Acids Res.* 35, 6798–6807
64 Conticello, S.G. *et al.* (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol. Biol. Evol.* 22, 367–377
65 Schreck, S. *et al.* (2006) Activation-induced cytidine deaminase (AID) is expressed in normal spermatogenesis but only infrequently in testicular germ cell tumours. *J. Pathol.* 210, 26–31
66 Li, B. *et al.* (2010) The protein–protein interface evolution acts in a similar way to antibody affinity maturation. *J. Biol. Chem.* 285, 3865–3871
67 Zaranek, A.W. *et al.* (2010) A survey of genomic traces reveals a common sequencing error, RNA editing, and DNA editing. *PLoS Genet.* 6, e1000954
68 Conticello, S.G. *et al.* (2007) DNA deamination in immunity: AID in the context of its APOBEC relatives. *Adv. Immunol.* 94, 37–73
69 Pauklin, S. *et al.* (2009) Alternative induction of meiotic recombination from single-base lesions of DNA deaminases. *Genetics* 182, 41–54
70 Yu, K. *et al.* (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* 4, 442–451
71 Roa, S. *et al.* (2008) Does antisense make sense of AID targeting? *Proc. Natl. Acad. Sci. U. S. A.* 105, 3661–3662
72 Perlot, T. *et al.* (2008) Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. *Proc. Natl. Acad. Sci. U. S. A.* 105, 3843–3848
73 Lipps, H.J. and Rhodes, D. (2009) G-quadruplex structures: *in vivo* evidence and function. *Trends Cell Biol.* 19, 414–422