ISSN 1022-7954, Russian Journal of Genetics, 2007, Vol. 43, No. 10, pp. 1093–1107. © Pleiades Publishing, Inc., 2007.
Original Russian Text © A.G. Lada, L.M. Iyer, I.B. Rogozin, L. Aravind, Yu.I. Pavlov, 2007, published in Genetika, 2007, Vol. 43, No. 10, pp. 1311–1327.

THEORETICAL PAPERS AND REVIEWS

Incarnation of Classical Pro- and Eukaryotic Mechanisms of Mutagenesis in Hypermutagenesis and Immunity of Vertebrates

A. G. Lada[a], L. M. Iyer[b], I. B. Rogozin[b, c], L. Aravind[b], and Yu. I. Pavlov[d]

[a] *Department of Genetics and Breeding, Biological Institute, St. Petersburg State University, St. Petersburg, 196240 Russia; fax: +7 (812) 428-77-33; e-mail: biotambovez@mail.ru*

[b] *National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894 United States*

[c] *Institute of Cytology and Genetics, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

[d] *Eppley Institute for Research in Cancer and Allied Diseases, Dapertment of Biochemistry and Molecular Biology, Department of Pathology and Microbiology, University of Nebraska Mediacl Center, Omaha NE 68198 United States*

Received May 11, 2007; in final form, June 5, 2007

Abstract—M.E. Lobashev has brilliantly postulated in 1947 that error-prone repair contribute to mutations in cells. This was shown to be true once the mechanisms of UV mutagenesis in *Escherichia coli* were deciphered. Induced mutations are generated during error-prone SOS DNA repair with the involvement of inaccurate DNA polymerases belonging to the Y family. Currently, several distinct mutator enzymes participating in spontaneous and induced mutagenesis have been identified. Upon induction of these proteins, mutation rates increase by several orders of magnitude. These proteins regulate the mutation rates in evolution and in ontogeny during immune response. In jawed vertebrates, somatic hypermutagenesis occurs in the variable regions of immunoglobulin genes, leading to affinity maturation of antibodies. The process is initiated by cytidine deamination in DNA to uracil by AID (Activation-Induced Deaminase). Further repair of uracil-containing DNA through proteins that include the Y family DNA polymerases causes mutations, induce gene conversion, and class switch recombination. In jawless vertebrates, the variable lymphocyte receptors (VLR) serve as the primary molecules for adaptive immunity. Generation of mature VLRs most likely depends on agnathan AID-like deaminases. AID and its orthologs in lamprey (*PmCDA1* and *PMCDA2)* belong to the AID/APOBEC family of RNA/DNA editing cytidine deaminases. This family includes enzymes with different functions: APOBEC1 edits RNA, APOBEC3 restricts retroviruses. The functions of APOBEC2 and APOBEC4 have not been yet determined. Here, we report a new member of the AID/APOBEC family, APOBEC5, in the bacterium *Xanthomonas oryzae*. The widespread presence of RNA/DNA editing deaminases suggests that they are an ancient means of generating genetic diversity.

DOI: 10.1134/S1022795407100031

INTRODUCTION

In the hereditary material of living organisms (DNA or RNA), continually appear lesions that are spontaneous or induced by external factors. These lesions include base modifications, intra- or inter-strand crosslinks, disturbance of Watson–Crick interactions (mismatch of one or several nucleotide pairs), and single- or double-strand breaks. They can cause mutations or block replication. Most mutations reduce fitness of the organism. Replication block in haploids results in arresting cell division. Multiple DNA lesions in multicellular organisms induce apoptosis and thus decrease proliferation rate or cause death of differentiated cells that are reproductively quiescent. Prokaryotic and eukaryotic cells have mechanisms for protection against DNA-damaging agents. DNA damage repair follows a course of specific enzyme reactions. To reduce deleterious consequences, repair must be error-free, i.e., result in the restoration of the native sequence of the DNA molecule. Indeed, many genes whose mutations have mutator phenotype, encode various repair proteins. High-fidelity repair is not always possible. Some types of DNA damage, e.g., inter-strand crosslinks, distort both DNA strands. At a high level of damage, high-fidelity repair systems may reach saturation. In this case, replication arrest can be prevented by special systems of mutation bypass of the lesions. However, in this case repair fidelity is reduced. Replication continues, but the newly synthesized DNA sequence differs from the original one. Mutants for genes encoding repair systems have antimutator phenotype. Thus, mutations appear as a result of specialized biosynthetic processes in the cell, predicted by Lobashev long before the development of the current views on mutation and repair. As early as in 1947, Lobashev wrote: "establishment of a mutation does not relate to the phase of damage or stimulus to the cell, but to the phase of restoration, i.e., the process of inaccurate repair of substance lesions" [1].

The optimal rate of mutation is regulated by cell systems and is evolutionarily determined. In specialized cells and tissues, mutation rate may by orders of magnitude exceed this basic level. In this paper, we focus on molecular mechanisms of hypermutagenesis regulation in nature.

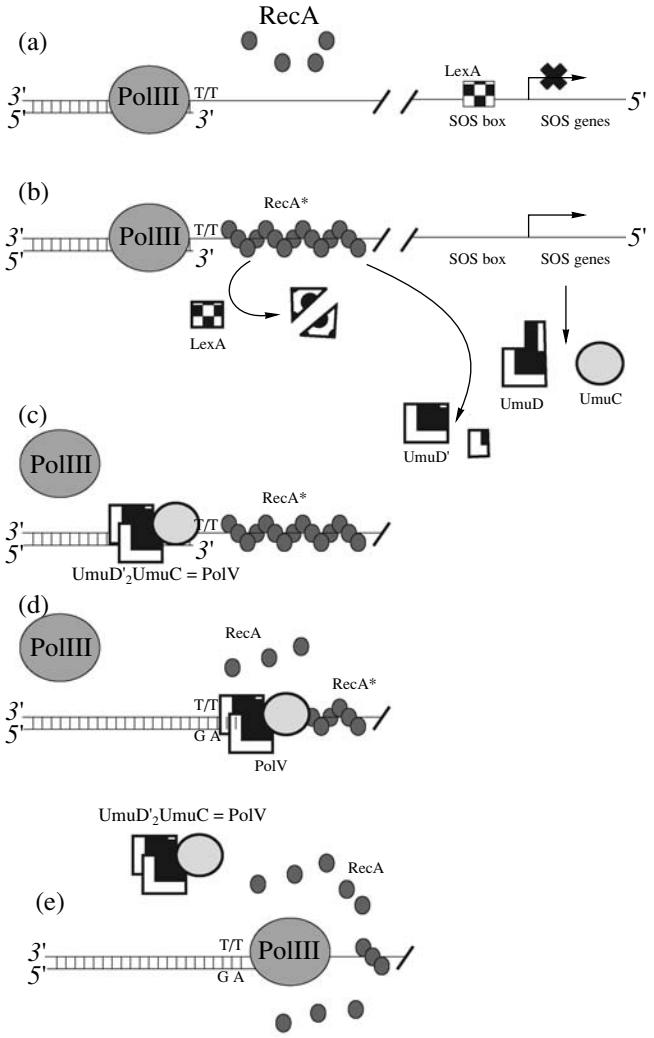

Fig. 1. SOS response in *E. coli* (after [6] with modifications and additions). (a) Normally, when LexA is linked to SOS-boxes in the promoters, the SOS genes are repressed. After encountering a thymine dimer, DNA polymerase III stops; (b) Protein RecA binds to the double-strand DNA stretch in the parental strand, activates, and mediates LexA cleavage, thus derepressing the SOS genes. A severe DNA damage activates the late *Umu* operon. RecA cleaves the newly synthesized UmuD, producing the active protein form UmuD'; (c) DNA polymerase III is substituted in the replication fork by trimer$\mathrm{UmuD'_2UmuC}$, effecting synthesis opposite the lesion (d). Nucleotides G and A are preferentially incorporated opposite a 6–4-photoproduct (5' ⟶ 3'), as shown in the figure, and A and A, opposite a thymine dimer. In eukaryotes, the AA pair is usually incorporated opposite both lesions; (e) after passing the lesion, a reverse polymerase substitution occurs, while reassembly of the RecA filament restores repression of the SOS genes.

SOS-RESPONSE SYSTEM IN *ESHERICHIA COLI* AND THE DISCOVERY OF INACCURATE DNA POLYMERASES

One of the most well-studied physical mutagens, UV light, generates cyclobutane pyrimidine (most often thymine) dimers and 4–6 photoproducts in DNA [2]. *Escherichia coli* has several mechanism of repair of such damage. In photoreactivation, photolyases produce a chemical reaction reverse to the formation of the UV products, thus restoring the initial nucleotide sequence [3]. In excision repair, the damaged DNA strand is cut upstream and downstream of the lesion; after the elimination of the damaged region, the resultant one-strand gap is filled in by DNA polymerase I [3]. However, upon strong UV irradiation, the photoreactivation and the excision repair systems fail to remove all of the photoproducts by the start of replication. Since photoproducts block the replisome movement along DNA, but not the reinitiation of synthesis in the region downstream of the damage [4–6], one-strand gaps of about 1000 bp (secondary UV lesions) are formed in the daughter DNA strands. These gaps can be repaired through post-replicative, *recA*-dependent repair. The main pathway of this repair is homologous recombination, which provides high-fidelity restoration of the genetic information of the damaged region at the expense of the homologous DNA strand. Mutations in genes *recA* or *lexA* arrest not only postreplicative repair, but also UV mutagenesis, at the same time drastically increasing bacterial sensitivity to UV light [7]. Consequently, some of the UV lesions are repaired by a special, error-prone *recA*-dependent system. Defais [8] has suggested that UV irradiation induces in *E. coli* a system of UV mutagenesis, which is suppressed under normal conditions. Radman and Witkin [9–12], further developing Defais's idea, advanced the so-called SOS hypothesis. According to this hypothesis, damaged DNA or blocked replication produce a regulatory signal, which derepresses SOS genes mediating survival in adverse conditions. Some of these damage-inducible *din* genes encodes the components of the error-prone SOS repair system, whose operation underlies SOS mutagenesis. In addition to the RecA protein, generation of SOS mutations requires the products of genes *umuC* and *umuD* (*umu = UV nonmutable)* as well as some components of the DNA polymerase III complex [13]. The molecular mechanism underlying SOS mutagenesis had been unknown until 1999, when two research teams independently have discovered polymerase activity of a SOS system component—the $\mathrm{UmuD'_2C}$ (DNA polymerase V) and showed that the new polymerase could synthesize the daughter DNA strand on a damaged template (translesion synthesis, TLS) [14, 15].

According to the current model of SOS response (Fig. 1), normally transcription of SOS genes is blocked, because repressor protein LexA is bound to the SOS boxes, which are definite sequences in SOS gene promoters. DNA lesions produce single-strand DNA regions, to which bind molecules of the RecA protein. Upon binding to DNA, this protein forms with it a nucleoprotein filament and is transferred to activated state. Activated RecA stimulate LexA autoproteolysis, which results in derepression of SOS gene transcription. SOS boxes of early SOS response genes have less

affinity to LexA than those of the late genes. Consequently, the time and the level of induction of different SOS system components depend on the extent of DNA damage. One of the first genes activated in SOS response encode proteins of error-free (excision and recombinational) repair [16]. In case of multiple damage, the activity of the error-free repair system may be insufficient for restoring the native DNA structure. In this situation, the late SOS genes, including *umuC* and *umuD*, are induced [16]. In the presence of RecA-ATP-ssDNA nucleoprotein filament, the UmuD protein undergoes proteolysis, producing UmuD', whose complex with UmuC ($UmuD'_2UmuC$), as mentioned above, is DNA polymerase V. During replication of the damaged genome, the replication fork stops at the damaged DNA site. Then, in place of DNA polymerase III, which is the main bacterial replicase, SOS-mutagenic DNA polymerase V is RecA-dependently loaded on the sliding clamp and catalyzes DNA synthesis opposite the lesion. The enzyme can add nucleotides opposite cyclobutane dimers, AP sites, 6–4 photoproducts, and some other modified bases. Polymerase V has low fidelity, but exactly this property makes it capable to synthesize DNA from a damaged template. This results in altered sequences (mutations) at the sites of DNA damage. For instance, in synthesis opposite a 6–4 photoproduct (3'-TT-5'), DNA polymerase V predominantly incorporates G opposite 3'-T and A opposite 5'-T (Fig. 1d). After a reverse polymerase change (Fig. 1e), replication proceeds as usual, and gaps in the daughter strand are filled [16].

In the same year 1999, the product of gene *dinB* was identified as DNA polymerase IV [17]. This polymerase, like DNA polymerase V, is induced in SOS response and participates in translesion synthesis [18]. Thus, even such relatively simply organized prokaryotic organism as *E. coli* has at least two (of five) specialized inaccurate DNA polymerases conducting replication on damaged templates.

EUKARYOTIC POLYMERASES OF THE Y FAMILY

In various bacteria, as well as in eukaryotes and archaea, several polymerases orthologous to *E. coli* DNA polymerases IV and V have been found. These enzymes constitute the Y family of DNA polymerases [19–21]. This family also includes polymerases Dbh and Dpo4 of *Sulfolobus solfataricus*; polymerases η (*RAD30A* or XPV, xeroderma pigmentosum variant; homologous gene in *Saccharomyces cerevisiae* is *RAD30*) and ι (*RAD30B*) and κ in human, providing translesion synthesis; and DNA-dependent deoxycytidyltransferase REV1, described in yeast and human.

The ability to bypass various DNA lesions and low fidelity of replication of undamaged templates, characteristic of the Y family polymerases, are explained by their structure (Fig. 2) [19]. Their characteristic feature is the presence of unique domain PAD (polymerase-associated domain, which is also called "little finger"), whose interaction with major DNA groove strengthens polymerase binding to the template. Another characteristic feature is greater openness of the active site and limited restrictions on the form of the synthesized base pair, as compared to polymerases from other families. All without exception polymerases of the Y family lack 3' ⟶ 5' exonuclease proofreading activity, which prevents slowing down DNA synthesis when a mismatched primer should be excised. Special structural elements of the family Y polymerases ensure their proficiency. For instance, owing to an additional clamp (N-clamp), polymerase κ encircles DNA (Fig. 2) [22].

The REV1 protein of yeast and human is also called G-polymerase because of its ability to incorporate a cytosine residue opposite guanine or a lesion, e.g., an AP site or a modified guanine residue [23–25]. The yeast REV1 turns a guanine residue of the template strand by ~90° in the plane perpendicular to the double helix axis. In this process, its atoms N7 and O6 (in other conditions involved in forming Hoogsteen base pairs) make hydrogen bonds with the amid groups of the main strand of polymerase Met685 and Gly686 residues. A number of other amino acid residues of the REV1 protein PAD domain, forming the so-called G-loop, are also involved in the interactions with guanine. The G-loop structure provides optimal pairing only with guanine, thus providing REV1 selectivity with regard to the template. The incoming dCTP does not pair with the templating guanine turned at ~90°. Instead, it makes hydrogene bonds with Arg324, which plays the role of a surrogate template. A number of other amino acid residues are also involved in the dCTP–polymerase interaction. REV1 specificity for dCTP is caused by the fact that none of the three other nucleoside triphosphates can make hydrogen bonds with Arg324. Thus, REV1 provides DNA extension without using the maternal DNA strand sequence as a template. The incorporation of cytosine opposite a templating guanine is determined by the polymerase structure [26]. This fact motivated Nair et al. [26], who have determined the REV1 structure and proposed to above mechanism of functioning of this enzyme, to title their paper "REV1 Employs a Novel Mechanism of DNA Synthesis Using a Protein Template."

Regulation of Y Family Polymerases

The involvement of the Y family polymerases in DNA synthesis beyond damaged regions would induce multiple mutations, thus having adverse consequences for the cell. Notwithstanding, no increase in mutation rate was observed even in case of overproducton of inaccurate polymerase η in yeast or in cell culture [27, 28]. This at the first glance paradoxical result is explained by the fact that prokaryotic and eukaryotic organisms have mechanisms limiting not only the time (for example, switching SOS response and inducing the

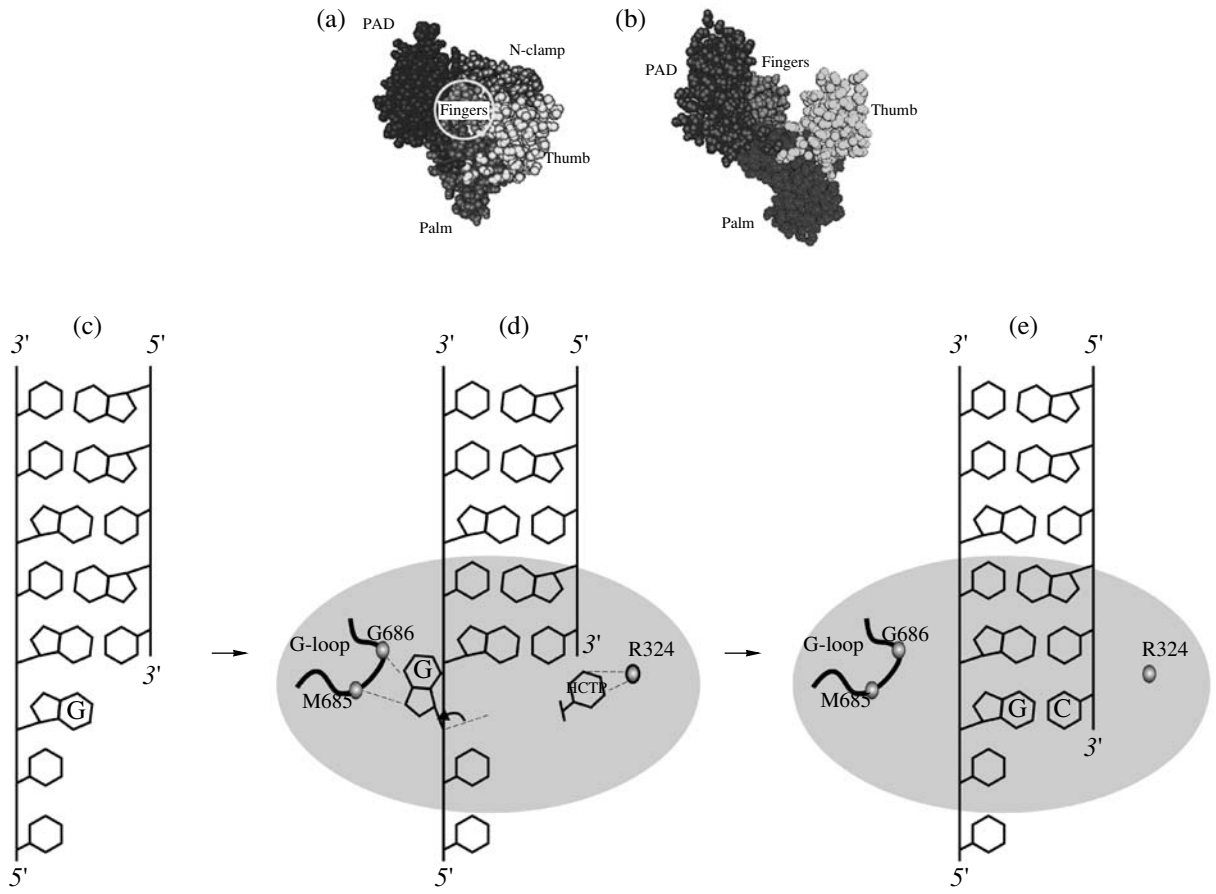

Fig. 2. Structure of some DNA polymerases (a, b) and the mechanism of the polymerase REV1 action (c, d). Human DNA polymerase κ (PDB accession number 2OH2). Apart from typical DNA polymerase domains (palm, thumb, and fingers), this polymerase has an additional domain, PAD, characteristic od Y polymerases, and domain N-clamp, unique for Pol. The position of DNA in transverse section is schematically shown by an open circle; polymerase κ (PDB accession number 1JIH) has the structure similar to that of polymerase η, but lacks the N-clamp domain; (c) polymerase REV1 (shown as a gray oval) can effect replication opposite guanine or a lesion. The guanine resides in the template strand from hydrogen bonds with amino acid residues Met685 and Gly686 of the polymerase G-loop; at that, the guanine residue turns and leaves the template. The incoming dCTP forms hydrogen bonds with Arg324; (d) after the cytosine incorporation in the daughter strand, the guanine residue returns to its normal position (e).

TLS polymerase gene expression after DNA damage), but also the activity site of mutagenic polymerases. To date, several mechanisms of polymerase switch upon the encounter of the replication fork with a lesion are known. First, among TLS polymerases, at least the Y family polymerases have low proficiency on undamaged templates and high proficiency on the damaged ones [29, 30]. Second, an important role in providing an access of polymerase to the synthesis is played by the sliding clamp. In bacteria, each of the two β-subunits of the clamp may be bound to polymerase [20]. Because of this, two polymerases (e.g., the main replicase Pol III and Pol V, providing TLS) can be located directly in the synthesis region and replace one another, if necessary. In eukaryotes, the interaction of Y polymerases with PCNA (proliferating cell nuclear antigen, eukaryotic sliding clamp), RFC (replication factor C, the clamp loading complex), and RPA (replication protein A, an eukaryotic analog of the prokaryotic ssDNA-binding SSB protein) are also essential for translesion synthesis [31–33]. In the TLS process, posttranslational PCNA modifications play a major role. DNA damage cause PCNA monoubiquitination at Lys164, effected by the Rad6–Rad18 complex [34–36]. This plays a key role in translesion synthesis. All Y polymerases have ubiquitin-binding domains, required for localizing Y polymerases in the replication machinery [37]. Ubiquitinated PCNA raises the activity of human polymerases η and REV1 in vitro [38]. In the absence of PCNA-deubiquitinating enzyme USP1, the frequency of spontaneous and UV-induced mutations increases. Interestingly, UV light promotes autocatalytic USP1 cleavage, which results in the degradation of the latter in the proteosome. This has remote similarity to the involvement of LexA autoproteolysis in the bacterial SOS response [39, 40].

Thus, low processivity of Y polymerases on undamaged templates, as well as restriction of the time and

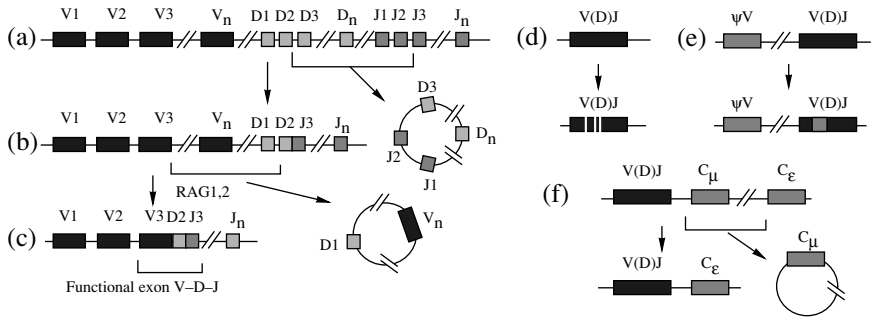

Fig. 3. Immunoglobuline gene formation in mammals. (a–c) RAG1,2-dependent V(D)J recombination. (a) The gametic genome encodes numerous V, D, and J segments. (b) In forming a mature immunoglobulin genes, first D and J segments (here, D2 and J3) are fused, losing the DNA fragment between them. (c) At the next stage, the resultant DJ segment (here, D2–J3) fuses with a V region (V3). This produces a VDJ exon (V3–D2–J3), encoding the variable immunoglobulin domain, while another interstitial DNA fragment is lost. (d) Somatic hypermutagenesis generates point mutations (shown by open vertical bars) into the VDJ exon. (e) Through gene conversion, information from inactive pseudo-V gene (ψV) is introduced into the VDJ exon; the transferred region is shown in light gray color. (f) Upon class switch recombination, the exon coding for the constant M-isotype domain, is replaced by one of the other constant exons (in our case, IgE is produced); the interstitial DNA fragment is also lost.

location of their operation by various cofactors, prevents possible deleterious consequences of spreading of inaccurate replication throughout the genome.

We should like to note an important feature of DNA translesion synthesis. In spite of multiple regulation of Y polymerase activity, the translesion synthesis is often continued several nucleotides apart of the "problem" region, on the native template. This may result in mutations downstream of the lesion, causing so-called nondirectional SOS mutagenesis, which was studied in detail in early works on UV mutagenesis [7, 41]. In eukaryotes, the key role in the synthesis product extension bypassing the lesion is played by inaccurate polymerase ζ from the B family. Error-prone DNA synthesis by polymerase ζ underlies induced mutagenesis in eukaryotes [20]. As shown in further text, spreading of the mutagenic polymerase activity to undamaged template regions play a significant role in immune response of vertebrates.

PROGRAMMED MUTATIONS AND GENOME REARRANGEMENTS IN EUKARYOTES

Higher mutation frequency in adverse conditions increases genetic diversity, thus providing material for selection and increasing adaptation of the population to the changed environment. Error-prone repair involving TLS polymerases, occurring also during SOS response, primarily ensures survival and reproduction of bacteria or cells in a multicellular organism in the case of drastic damage of DNA. Multicellular organisms have mechanism of mutagenesis programmed in ontogeny, resulting in diverse genomic changes, from point mutations to chromosome rearrangements, which enhance general stability of the organism.

Production of Human Immunoglobulins

V(D)J recombination. A classic example of programmed rearrangements of genetic material in higher vertebrates is the formation of mature immunoglobulin genes and T-cell receptors (TcR) in mammals. The number of antibody variants synthesized by the organism significantly exceeds the number of genes in the human genome. Theoretically, according to some estimates, human organism can produce up to 10^{11} different variants of antibodies. This enormous diversity is based on site-specific recombination, occurring in developing lymphocytes of the bone marrow. In cells not producing lymphocytes, the loci encoding constant and variable immunoglobulin domains are located on the chromosome far apart from one another. When a mature immunoglobulin gene is generated, these loci are fused; the variable regions of heavy immunoglobulin chains are produced by uniting three type of immunoglobulin gene segments: V (variability), D (diversity), and J (joining). As the formation of genes for light chains involves only V and J segments, the process of forming variable parts of immunoglobulin genes is referred to as V(D)J recombination (Fig. 3).

Joining of the V, D, and J segments involves the products of genes *RAG1* and *RAG2* (recombination activating genes), which are similar to transposases and retroviral integrases. First the RAG proteins introduce single-strand breaks in the RS (recombination sequences) regions, flanking V, D, and J segments, and then catalyze a reaction of transetherification, which covalently binds the 3' end of the coding strand of each fragment to the 5' end of the noncoding strand (a hairpin appears), resulting in a double-strand break. Next, the hairpin is cleaved at the coding region–RS boundary, the single-strand parts are filled to form double strands, and, finally, the double-strand breaks are repaired with participation of proteins of the NHEJ (nonhomologous end joining) system. More diversity

(N-terminal variability) is generated by nontemplating addition of nucleotides by terminal deoxynucleotidyltransferase. Thus, combination of multiple fragments of immunoglobulin genes generates primary diversity of antibodies [42, 43].

Affinity maturation of antibodies. Mature B lymphocytes that underwent V(D)J recombination migrate to secondary lymphoid organs (spleen and lymphatic nodes, where they encounter antigens provided by T-lymphocytes or follicular dendrite cells. Lymphocytes bounding antigens by means of their immunoglobuline receptors receive a signal for proliferation, which stimulates their further divisions and differentiation. Then, three processes increasing immunoglobuline diversity can occur in the stimulated B-cells. These processes are somatic hypermutagenesis, class switch recombination, and gene conversion [44]. Upon somatic mutagenesis, point mutations occur in variable parts of immunoglobulin genes at a frequency attaining ~10^{-3} per base pair per division, which is million times higher than in other genes, whereas mutation level in other genome regions remains stable [42, 45]. Two statistically preferable "hot motifs" are characteristic of somatic mutagenesis—RGYW/WRCY (the mutating pair G–C is underlined; R, purine; Y, pyrimidine; W, adenine or thymine) for mutations in G–C pairs of the both DNA strands and WA/TW (the mutating pair A—T is underlined) for mutations in pairs A–T of the nontemplate strand [46]. Mutations in G–C and A–T pairs occur at equal frequencies [47, 48].

Consecutive rounds of mutagenesis and proliferative stimulation of cells with "advantageous" mutations that increase immunoglobulin affinity to the antigen, result in the formation of high-affinity antibody molecules. They, so to say, adjust the antigen-binding pocket in the variable immunoglobuline domain to the antigen structure. This process was termed clonal lymphocyte selection [49]. In birds, the main mechanism of antibody affinity maturation is gene conversion, during which the regions of nonfunctional pseudo-V-genes substitute heterologous regions of immunoglobulin genes. This generates point substitutions, as in somatic hypermutagenesis [50]. Switching antibody isotypes, i.e., formation of various classes of antibodies with the same specificity, occurs by substitution of constant gene regions in immunoglobuline genes in the course of isotype switching (CSR, class switch recombination) (Fig. 3) [44].

AID and its role in antibody affinity. A significant step towards understanding the molecular mechanisms underlying the three processes—somatic hyper mutagenesis, class switch recombination, and gene conversion—was made by the team headed by Honjo in 1999. Using subtractive mRNA hybridization, they identified gene *AID* (activation-induced deaminase), induced by CSR activation [51]. *Aid* deletion in mice resulted in the CSR absence and somatic hypermutagenesis [52]. Moreover, it has been shown that mutation in the human *AID* gene are associated with a special form of immunodeficite—hyper-IgM syndrome (HIGM2). In patients having this syndrome, no class switch recombination occur and all immunoglobulines produced by the organism belong to the prototype M class. These patients also lack somatic hypermutagenesis [53]. It was shown that the *AID* gene is required for gene conversion in birds [54, 55]. Thus, the *AID* expression proved to be needed for all of the three known mechanisms of additional immunoglobulin diversification.

The *AID* gene encodes a 28-kDa protein, whose closest homolog is RNA-editing enzyme APOBEC1 (apolipoprotein B mRNA-editing catalytic subunit 1). In the so-called editing complex, APOBEC1 deaminates cytosine in position 6666 of the open reading frame of apolipoprotein B100 mRNA, which results in the transformation of glutamine codon CAA to stop codon UAA and the production of truncated tissue-specific protein apoB48 [56, 57]. Based on the homology between proteins AID and APOBEC1, Muramatsu et al. proposed a hypothesis of RNA edition, according to which AID deaminates an unknown mRNA, which results in producing a hypothetic recombinase, encoded in this mRNA and involved in immunoglobulin gene diversification [52]. An alternative hypothesis of DNA edition, advanced by Neuberger [58], suggests that AID directly deaminates cytosine in DNA, and the resultant uracil switches the reaction underlying hypermutagenesis, class switch recombination, and gene conversion [58]. Most of the available data supports the second hypothesis. For instance, the *AID* expression in wild-type *E. coli* cells increases the mutation rate 3–6 times an, more importantly, shifts the mutation spectrum towards C ⟶ T transitions. The mutagenic effects of AID increased in strains carrying a deletion of uracil-DNA-glycosylase gene (Δ*ung1*) and incapable for eliminating uracil from DNA, which is a significant argument in favor of DNA deaminase activity of AID [58]. In subsequent biochemical studies, the ability of AID to deaminate cytosine in DNA in vitro was directly shown [43].

The production of the AID protein in yeast *S. cerevisiae* also has a mutagenic effect, which is higher in *ung1*$^-$ strains. Analysis of the nucleotide context of the mutation sites demonstrated that yeast mutations caused by the *AID* gene expression occur preferentially in sequences DGYW/WRCH) (the mutating pair G–C is underlined; R, purine; Y, pyrimidine; D, guanine, adenine, or thymine) [59]. Thus, the *AID* expression in yeast generates mutations in the nucleotide context similar to the context of mutation occurring in G–C pairs of immunoglobulin genes in somatic mutagenesis. The mutation spectra, obtained by in vitro AID interaction with single-strand DNA, show hot spot for mutations in G–C pairs in the same motifs as in vivo [60]. These results indicate that AID substrate-specificity determines the position of hot spots in somatic hypermutagenesis. To understand possible mechanisms

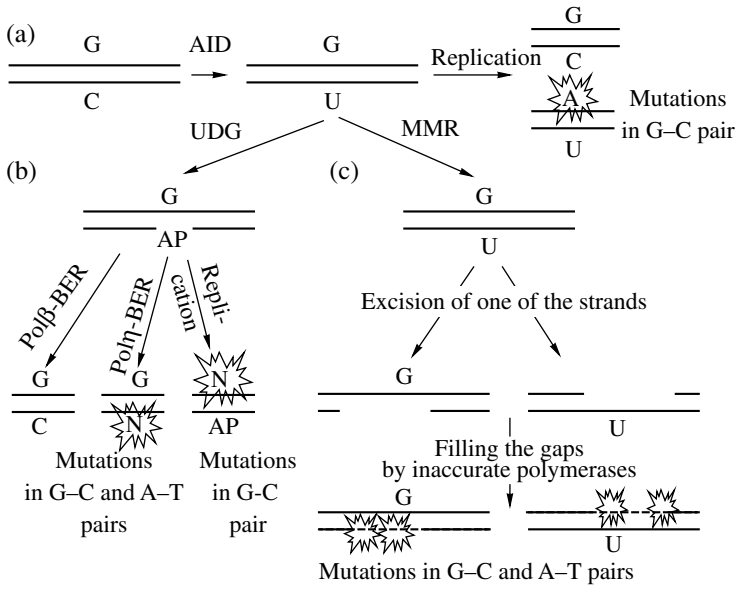

Fig. 4. Processing of uracil in DNA upon somatic mutagenesis. (a) AID deaminates cytosine in DNA, producing a G–U pair. Replication will result in GA transition, because thymine will be incorporated opposite uracil. (b) after switching the BER system, uracil-DNA-glycosilase (UDG) excises uracil, forming an apyrimidine site. The restoration of the native DNA structure involves participation of accurate polymerase β at subsequent stages of repair. The involvement of TLS polymerase η at the filling in stage results in mutations in pairs G–C and A–T. Upon DNA replication with the AR site mutations occur in pairs G–C. (c) If an unmatched G–U pair is recognized by the MMR system, one strand of the region containing the unmatched pair is excised. Filling in of the gaps by inaccurate polymerases results in mutations in G–C and A–T pairs.

of the mutation formation in A–T pairs, the consequences of cytosine deamination in DNA should be considered.

The role of mutagenic DNA polymerases and mismatch repair in somatic hypermutagenesis. Uracil formed in DNA can be recognized and excised by uracil-DNA-glycosilase (UNG), which is followed by base excision repair reactions that restore the original base pair G–C, so that no mutation occur (Fig. 4). If uracil-containing DNA is involved in replication, adenine is incorporated opposite uracil in the daughter strain, and a C ⟶ T transition is fixed in the next replication round. In addition, in translesion replication, a random nucleotide can incorporate opposite the apyrimidine site produced by the UNG protein activity, which leads to mutation C ⟶ N (where N is any nucleotide) in the deamination site. Spreading of inaccurate replication in the 3' direction from the AP site may induce mutations in A–T pairs. The stage of gap filling in base excision repair usually involve relatively accurate polymerase β. However, the activity of Pol β is suppressed in B-lymphocytes, which may result in the involvement of Y polymerases in uracil-containing DNA repair [61].

The involvement of TLS polymerases in somatic mutagenesis is also associated with operation of mismatch repair (MMR) system. Mice with deletions of genes coding for various components of the MMR system exhibit low somatic mutagenesis and higher frequency of mutations in G–C then in A–T pairs [43]. Binding of mismatched U–G base pairs by the Msh2–Msh6 complex triggers MMR, in which a part of one of the DNA strands in the mismathc region is excised (Fig. 4c). Filling the resultant gaps by TLS polymerases leads to mutations and premutation states (mismatches), which are repaired or fixed as mutations in the next replication round. At that, mutations both in G–C and in A–T pairs can occur. The mechanisms of recruiting TLS polymerases in variable immunoglobulin gene regions instead of accurate replicative polymerases is still unknown.

A number of mutagenic DNA polymerases were suggested as candidates for involvement in somatic mutagenesis [62]. These are primarily polymerases η and ι from the Y family, θ from the A family, and ζ from the B family, as well as κ from the Y family, and μ and λ from the X family. So far, reliable experimental support of the involvement in SHM was obtained only for polymerase η. A deletion Pol η results in a drastic reduction in mutation frequency in immunoglobulin genes in A–T pairs [63, 64]. The same effect is observed in patients with xeroderma pigmentosum carrying a mutation in the polymerase η gene [65]. The mutation spectra generated by polymerase η are characterized by the presence of W<u>A</u>/<u>T</u>W hot spots, which supports the assumption that this polymerase is involved in translesion synthesis in SHM [66–68]. However, since Pol η deletion does not lead to the total absence of mutations in A–T pairs, as well as because of insufficient information available on the error spectra of other DNA polymerases, it may well be that other mutagenic poly-

merases are involved in somatic hypermutagenesis, being able to "substitute" for Pol η in its absence.

The level of somatic mutagenesis in B-lymphocyte cultures depends on the production of the AID protein [69]. Normally, *AID* is specifically expressed in B lymphocytes. Artificial overproduction of AID triggers somatic mutagenesis in fibroblast cultures [70] and hybridomas [71]. Therefore, fibroblasts and hybridomas have all components and conditions for hypermutagenesis except the *AID* gene expression. Thus, the AID protein is an only factor of hypermutagenesis that is absent in most cells. Abnormal *AID* regulation, leading to the ectopic deaminase expression, causes cancer [72], whereas constitutive AID production in mice raises the tumor frequency [73]. Recently, Matsumoto et al. [74] have shown that infection by *Helicobacter pylori* induce *AID* expression in human gastric epithelium cells. This is accompanied by mutations appearing in the tumor suppressor gene *TP53*, whose inactivation is often observed in case of cancer. Thus, malignization upon *H. pylori* infection is likely to be mediated by the AID protein [74].

Role of AID in class switch recombination. Class switch recombination requires a substitution of the coding constant domain IgM by one of the four other domains: C_α, C_γ, C_δ, or C_ε (Fig. 3). Apparently, deamination of the neighboring cytosine residues in different DNA strands in the switch regions located downstream of the exons encoding constant domains and subsequent operation of systems BER and MMR generate double-strand breaks (DSB) [75, 76]. Joining nonhomologous ends by the NHEJ repair system produces two products, a mature immunoglobulin gene of a new isotype and a circular fragment containing DNA that was located between two breaks; the latter is subsequently lost [77, 78]. The CSR level decreases with impairment of various MMR and BER system components [79]. If MMR and BER are both absent, the CSR level is very low [80].

The data obtained in a yeast system support the above model of AID involvement in initiation of recombination rearrangements. In yeast, homologous recombination is initiated by a double-strand DNA break. The *AID* expression in yeast induces intragenic recombination. The inactivation of uracil-DNA-glycosilase (*ung1*) blocks recombinogenic AID activity [81]. The necessity of the AID protein for gene conversion can be explained in a similar manner: AID initiates repair that generates breaks. In this case, they are cured by the system of homologous recombination [79].

Formation of Lymphocyte Receptors in Vertebrates

The adaptive immune response in all jawed vertebrates is associated with immunoglobulin lymphocyte receptors, whose diversity can be achieved in various ways. For instance, in sharks and scates, the main unit V_H-D_{H1}-D_{H2}-J_H-C_H is multiply repeated in the genome, but recombination occurs only within each of such units, which dramatically reduces the possible antibody number [82]. In chicken, the number of immunoglobulin-coding genes is low and various antibodies are produced largely by gene conversion [49]. *AID* paralogs and orthologs were identified in the genomes of members of various vertebrate groups, including cartilaginous and bony fishes, amphibians, and birds [82, 83].

Adaptive Immune Response in Jawless Vertebrates

Jawless vertebrates (lampreys and hagfishes) also have lymphocytes and are capable of adaptive immune response. However, no immunoglobulin receptors have been found in these organisms; they also lack *RAG* genes. Very recently, it has been reported that in jawless vertebrates, the role of immunoglobulins is played by variable lymphocyte receptors (*VLR*) [84, 85]. Immature (embryonic) *VLR* genes in lamprey and hagfish encode N- and C-terminal regions of future receptors, divided by stretches of noncoding DNA. However, in mature lymphocytes, the *VLR* genes contain the central variable region consisting of various leucine-rich repeats (LRR) (Fig. 5). It was suggested that mature *VLR* genes result from combination of various LRR modules flanking embryonic *VLR* genes [83, 85, 86]. What is the mechanism of LRR combination that produces a mature *VLR* gene? In 2006, Nagawa et al. [87] have suggested a copy choice model to explain this process. However, more promising seems the model of consecutive LLR assembly producing VLR by use of a mechanism similar to class switch recombination in jawed vertebrates.

Two genes orthologous to *AID*, *PmCDA1*, and *PmCDA2*, were identified in the lamprey genome. Both genes are specifically expressed in lymphocytes and gematopoietic tissues, which suggests the involvement of the corresponding proteins in the immune response [88]. By analogy with CSR, one could conjecture that cytosine deamination in DNA by these deduced hypothetical enzymes causes double-strand breaks, whose repair leads to LRR insertion in the produced *VLR* gene. To test this hypothesis of the formation of variable lymphocyte receptors, experiments with unicellular organisms were carried out. As in the case of AID, the PmCDA1 production in *E. coli* caused C $\longrightarrow$ T transitions, whose frequency was higher in strains carrying mutations at uracil-DNA-glycosilase. The PmCDA1 production in yeast *S. cerevisiae* had a similar effect. Moreover, PmCDA1 produced in yeast induced intragenic recombination, for which, as with AID, uracil-DNA-glycosilase was required. These results support the hypothesis on recombinogenic VLR formation [88].

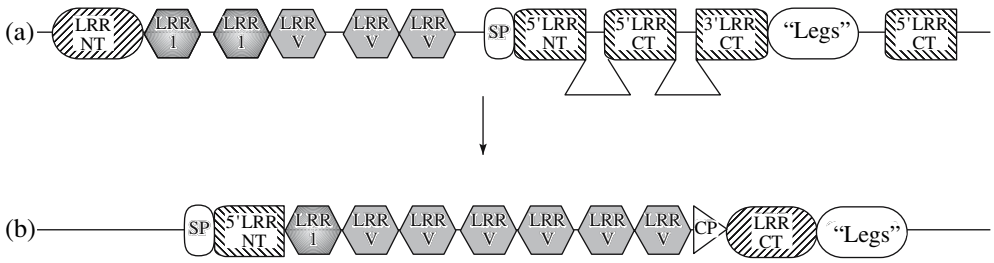

Fig. 5. Formation of variable lymphocyte receptors in jawless vertebrates. (a) In immature *VLR* (Variable Lymphocyte Receptors) genes, N- and C-terminal parts of the future variable receptor regions (LRR NT and LRR CT, respectively) are separated by noncoding DNA stretches. In generating mature *VLR* gene, various leucine-rich repeats flanking the primordial gene (LRR V) form the variable region of the *VLR* gene. (b) Mature *VLR* gene consists of a signal peptide (SP), a variable region with a connecting peptide (CP), and "legs," used for attaching the protein to the membrane.

THE AID/APOBEC FAMILY OF DNA/RNA DEAMINASES

As mentioned above, proteins AID, APOBEC1, PmCDA1, and PmCDA2 are members of a large deaminase family, AID/APOBEC. Their closest relatives are adenosine deaminases (see below the section on the molecular evolutionary tree of deaminases). The family is divided into subfamilies (Fig. 6). The subfamily APOBEC1 includes RNA editing enzymes APOBEC1. The functions of proteins from the subfamily APOBEC2 are unknown. Proteins of the AID subfamily are involved in vertebrate immunity. Mechanisms of protection against retroviruses by means of proteins of the subfamily APOBEC3 are discussed in the next section. In 2005, the fifth subfamily, APOBEC4, was identified using computer-aided methods. APOBEC4 proteins occur in all quadruped vertebrates. In human, APOBEC4 is expressed in testis, which suggests its possible role in spermatogenesis [89].

Subfamily APOBEC3 and Protection against Retroviruses

Retroviruses consist of two copies of single-strand RNA genome, packed into the virion together with a reverse transcriptase. Upon cell infection, after the fusion of the lipoproteid viral capsule with the cell membrane, the viral genome gets into the cytoplasm, and reverse transcriptase synthesizes the minus strand of DNA using the RNA plus strand as a template. Then the DNA plus strand is synthesized on the DNA minus strands, thus forming a double-strand DNA molecule, which integrates into the host genome with help of integrase. Transcription of the integrated sequences produces new copies of the viral genome, which, together with the RNA plus strand-encoded proteins are packed into new virus particles. Well-known examples of retroviruses are human immunodeficite virus (HIV) and T-cell leukemia virus [90]. The APOBEC3G protein was identified using subtractive cDNA hybridization as a factor inhibiting propagation of HIV, which is the most well-known retrovirus [91]. Studying the molecular mechanism of the inhibition has shown that APOBEC3G exhibits deaminase activity on the single-strand DNA and causes C ⟶ U transitions in the minus strand of the viral DNA. Processing of the resultant uracil by the BER system results in production of an AP site with subsequent degradation of the viral DNA by AP endonucleases. If the uracil-containing minus strand escapes degradation and replicates, then an insertion of adenine opposite uracil leads to multiple G ⟶ A transitions in the plus strand (hypermutagenesis), resulting in a loss of function of the proteins encoded in the viral DNA. Interestingly, APOBEC3G is packaged in virus particles by viral protein Gag and deaminates the synthesized DNA minus strand upon viral infection from another cell [92].

Viruses can counteract destructive APOBEC3G activity. Most natural HIV strains have a special protein, Vif (virus infectivity factor). Vif viruses are resistant to the effect of APOBEC3G. Apparently, Vif blocks the deaminase activity by inducing polyubiquitinylation of APOBEC3G and its degradation in the proteosome [91, 93].

The human APOBEC3 gene is located on chromosome 22 in a cluster of eight homologous genes (*APOBEC3A–APOBEC3H*), belonging to the subfamily APOBEC3 of the AID/APOBEC family. The products of four of these genes (*APOBEC3B, APOBEC3G, APOBEC3F*, and *APOBEC3C*) suppress retroviruses probably by means of the above mechanisms, differing only in hot motifs of hypermutagenesis. APOBEC3A suppresses retrotransposons; the functions of the other three genes are unknown [94, 95]. From the evolutionary viewpoint, it is interesting that, in contrast to human and chimpanzee that have eight genes of the APOBEC3 subfamily, rodents have only one such gene. Apparently, the additional *APOBEC3* gene copies formed via tandem duplications and unequal crossing over. It is believed that such strong divergence was caused by the necessity to suppress endogenous retrotransposons to ensure genomic stability [96]. Indeed, the divergence of the APOBEC3 subfamily and a decline in retrotransposon activity coincide in time [97].

Several authors have shown that mutant APOBEC3G and APOBEC3F, incapable of deamina-

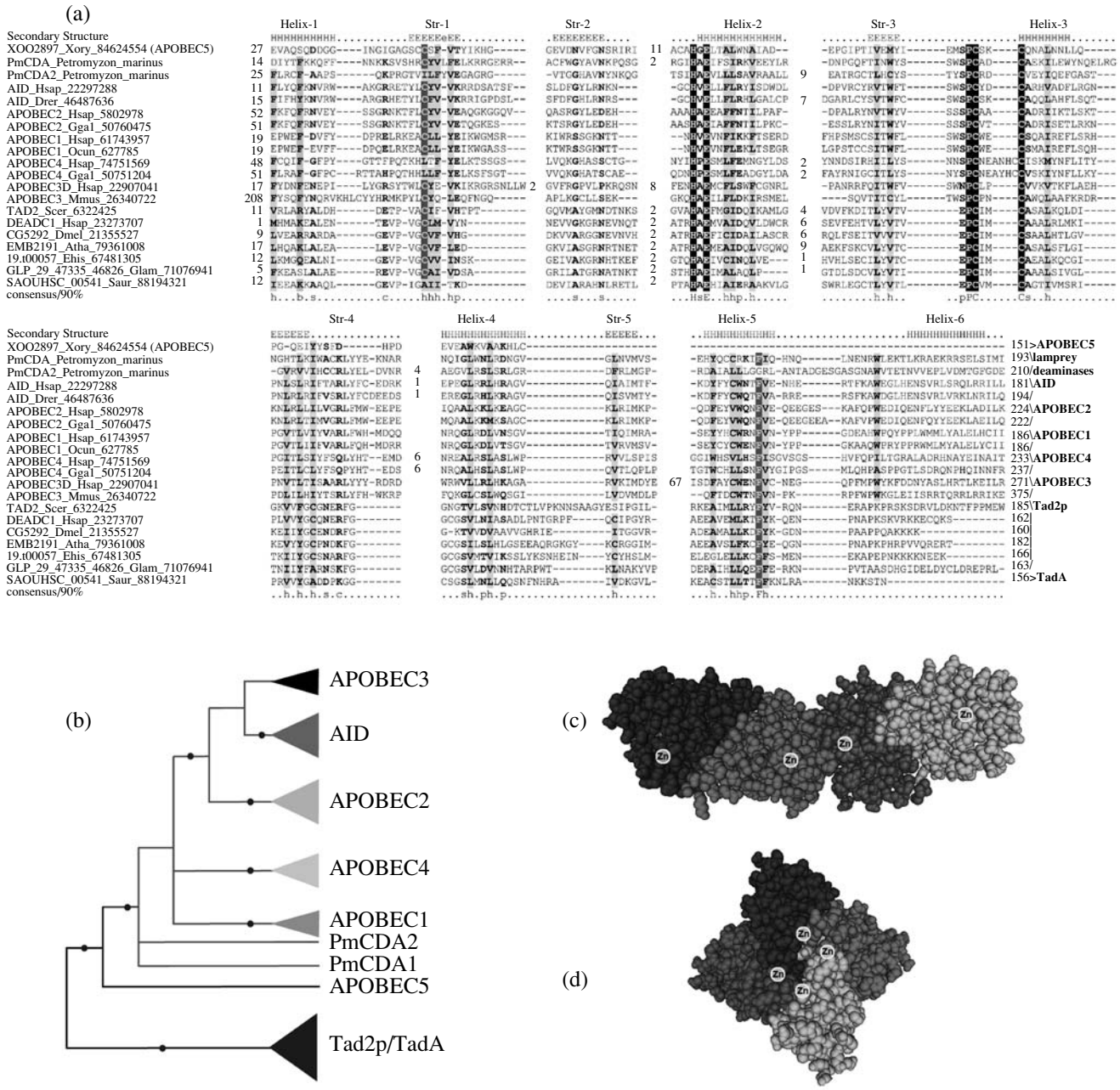

Fig. 6. AID/APOBEC family of RNA/DNA deaminases. Multiple alignment (a) and a phylogenetic tree (b) of the AID/APOBEC family. Conserved amino acids are shown in gray or black, depending on the degree of conservatism. Consensus with a 90% conservatism threshold is presented in the last line of alignment [h: hydrophobic amino acids (ACFILMVWY), b: large amino acids (EFHIKLMQRWY), s: small amino acids (AGSVCDN), c: charged amino acids (EDKR), p: polar amino acids (STEDKRNQHC)]. Abbreviations: Atha, *Arabidopsis thaliana*; Dmel, *Drosophila melanogaster*; Drer, *Danio rerio*; Ehia, *Entamoeba histolytica*; Ggal, *Gallus gallus*; Glam, *Giardia lamblia*; Hsap, *Homo sapiens*; Mmus, *Mus musculus*; Ocun, *Orystolagus cuniculus*; Saur, *staphylococcus aureus*; Scer, *Saccharomyces cerevisiae*; Xory, *Xanthomonas oryzae.* (b) Maximum likelihood phylogenetic tree schematically shows the evolutionary relationships between the AID/APOBEC subfamilies. The family of RNA-editing enzymes Tad2p/TadA was used for finding the root of the AID/APOBEC phylogenetic tree. (c) Three-dimensional structure of the human APOBEC2 (PDB accession number 2NYT). The position of Zn^{2+} ions playing the key role in catalysis is shown. (d) Three-dimensional structure of murine cytidine deaminase (PDB accession number 2FR6).

tion, still retain a certain level of antiviral activity, while mutant APOBEC3A and APOBEC3B efficiently blocks transposon movement. Apparently, the mechanisms of antiretroviral activity of APOBEC3 proteins are not restricted to DNA deamination (see [98]).

The Structure of APOBEC2 and APOBEC3G

Three-dimensional structure has been established for a number of members of the deaminase suprafamily, CDAs (cytodone deaminases), which have free nucle-

otides and nucleotides as substrates. These proteins are characterized by dimeric or tetrameric organization, in which the active center is located on the inner surface of the oligomer (Fig. 6d). This makes possible to binding nucleotides and fulfilling the corresponding biosynthetic functions. The structure of human APOBEC2, a member of the AID/APOBEC family, has been established. The functions and biochemical activities of this protein produced in muscle are unknown. However, APOBEC2 is the only protein of the AID/APOBEC family with the known three-dimensional structure. The mutuial positions of subunits in tetrameric APOBEC2 differs from that in classic deaminases operating on free nucleotides. The APOBEC2 tetramer has elongated shape with the catalytic center exposed on the outer complex surface (Fig. 6c). This structure probably provides the possibility to bind and transform polynucleotide substrates by AID/APOBEC deaminases. Impairment of the AID activity in mutants carrying mutations in certain amino acid residues of the protein conforms to this structural model [99].

A similar structural model was proposed for APOBEC3G on the basis of the known APOBEC2 structure. To design this model, alignment of amino acid sequences and parts of the secondary structure, as well as information derived from mutation analysis of APOBEC3G, were used. This model was employed to deduce the regions of the N-terminal cytidine deaminase domain required for the protein packing in the viral capside [100].

EXTENSION OF THE MOLECULAR EVOLUTIONARY TREE OF DEAMINASES FROM THE AID/APOBEC FAMILY

We have conducted a bioinformative screening for novel members of the AID/APOBEC family, using the PSI-BLAST software (www.ncbi.nlm.nih.gov/BLAST/). Novel AID/APOBEC family members were sought in the GenBank database on the basis of statistically nonrandom similarity of proteins with unknown functions with AID/APOBEC family members. In this way, a new member of this deaminase family, hereafter referred to as APOBEC5, was found. Interestingly, the gene for APOBEC5 (XOO2897, GI: 64624554) was revealed in bacterium *Xanthomonas oryzae* (Fig. 6). The member of the AID/APOBEC family most similar to APOBEC5 proved to be protein APOBEC2 of fish *Tetraodon nigroviridis.* Similarity between APOBEC5 (XOO2897) and APOBEC2 was statistically significant even at the first step of the PSI-BLAST procedure (the probability of this similarity due to random causes $E < 5 \times 10^{-3}$), whereas the other known AID/APOBEC family members were detected at subsequent steps of the program. Alignment of APOBEC5 with other deaminases (Fig. 6a) showed that, along with motifs common for all deaminases, APOBEC5 have two features that are characteristic only for the AID/APOBEC family: helix-4 (Fig. 6a) and an insertion of three amino acids upstream the PCxxC motif (Fig. 6a). This confirms that APOBEC5 indeed is a new member of the AID/APOBEC family. The presence of highly conserved motifs HxE and PCxxC (where x is any amino acid), which are characteristic of various deaminases (Fig. 6a) indicates that APOBEC5 may be a functional deaminase. Interestingly, the C-end of this protein is shorter than those of the known AID/APOBEC family members (Fig. 6a). This may suggest a change of specific function of this protein.

Phylogenetic analysis with the use of various methods (for their description, see [88]) consistently classified APOBEC5 as the most distant AID/APOBEC family member (Fig. 6b). The absence of other members of APOBEC5 subfamily may indicate a relatively recent transfer of this gene to *X. oryzae* from the vertebrate genome. The AID/APOBEC family members of lamprey were not assigned to any of the known deaminase subfamilies of jawed vertebrates (Fig. 6b), forming a separate subfamily. The divergence between subfamilies APOBEC1, APOBEC4, APOBEC2, AID, and APOBEC3, probably, have occurred after splitting of jawed and jawless vertebrates. Previous analysis [87] have shown that the family of RNA-editing enzymes Tad2p/TadA are closest to the AID/APOBEC family (Fig. 6b). These proteins have a number of traits characteristic of AID/APOBEC, including helix-4 and helix-5 (see Fig. 6a). Such structural elements are absent in other deaminase families, including cytidine deaminases (CDA, EC 3.5.4.5), which were previously thought closest to the AID/APOBEC family. Moreover, Tad proteins that edit tRNA in trypanosome, apart from deaminating adenine to inosine in tRNA, can deaminate cytosine to uracil in single-strand DNA [101]. Establishing the relationships between Tad, AID/APOBEC, and CDA, one should also take into account demonstrated and deduced differences in the quaternary structure of AID/APOBEC proteins (see the previous section). The wide distribution of DNA- and RNA-editing deaminases suggest that they are involved in the ancient mechanism of regulating genetic variation, which has appeared at the dawn of the vertebrate radiation.

MUTAGENESIS AS AN ELEMENT OF CELL PHYSIOLOGY

Hypermutagenesis that enhances adaptation of an organism to the environmental conditions can occur in organisms of different levels of organization, from intestinal bacteria to human. Hypermutagenesis is characterized by physiological lability, manifested in the dependence of the time and action of mutator systems on the requirements of the environment, such as strength of the effect of DNA-damaging agents or the complexity of the antigen structure. In addition to mutator proteins, for instance, DNA deaminases and inaccurate DNA polymerases, hypermutagenesis involves the repair and replication systems. Fixation of

mutations occurs upon inaccurate restoration of the DNA structure with coordinated operation of various molecular systems. The main aspects of hypermutagenesis are in good agreement with the theory of mutation developed by Lobashev, in which mutagenesis was regarded as a complex physiological process [1, 102, 103].

REFERENCES

1. Lobashev, M.E., Physiological (Paranecrotic) Hypothesis of the Mutation Process, *Vestn. Leningr. Univ.*, 1947, no. 8, pp. 10–29.
2. Setlow, R. and Setlow, J.K., Effects of Radiation on Polynucleotides, *Annu. Rev. Biophys. Bioeng.*, 1972, vol. 1, pp. 293–346.
3. Altshuler, M., Recovery of DNA Replication in UV-Damaged *Escherichia coli, Mutat. Res.,* 1993, vol. 294, no. 2, pp. 91–100.
4. Rupp, W.D. and Howard-Flanders, P., Discontinuities in the DNA Synthesized in an Excision-Defective Strain of *Escherichia coli* Following Ultraviolet Irradiation, *J. Mol. Biol.*, 1968, vol. 31, pp. 291–304.
5. Heller, R.C. and Marians, K.J., Replication ork Reactivation Downstream of a Blocked Nascent Leading Strand, *Nature*, 2006, vol. 439, pp. 557–562.
6. Lopes, M., Foiani, M., and Sogo, J.M., Multiple Mechanisms Control Chromosome Integrity after Replication Fork Uncoupling and Restart at Irreparable UV Lesions, *Mol. Cell*, 2006, vol. 21, pp. 15–27.
7. Witkin, E.M., Ultraviolet Mutagenesis and Inducible DNA Repair in *Escherichia coli, Bact. Rev.*, 1976, vol. 40, no. 4, pp. 869–907.
8. Defais, M., Fauquet, P., Radman, M., and Errera, M., Ultraviolet Reactivation and Ultraviolet Mutagenesis of λ in Different Genetic Systems, *Virology*, 1971, vol. 43, pp. 495–503.
9. Radman, M., Phenomenology of an Inducible Mutagenic DNA Repair Pathways in *Escherichia coli* SOS Hypothesis, *Molecular and Environmental Aspects of Mutagenesis*, Tabor, H.W., Ed., Springfild: Charles C. Thomas Publ., 1974, pp. 128–142.
10. Radman, M., SOS Hypothesis: Phenomenology of an Inducible Mutagenic DNA Repair which is Accompanied by Mutagenesis, *Molecular Mechanisms for Repair of DNA*, Hanawalt, P. and Setlow, R.B., Eds., New York: Plenum, 1975, part A, pp. 355–367.
11. Witkin, E.M. and George, D.L., Ultraviolet Mutagenesis in *polA* and *uvrA polA* Derivatives of *Escherichia coli* B/r: Evidence for an Inducible Error-Prone Repair System, *Genetics*, 1973, vol. 73, suppl., pp. 91–108.
12. Witkin, E.M., Elevated Mutability of *polA* and *uvrA polA* Derivatives of *Escherichia coli* B/r at Sublethal Doses of Ultraviolet Light: Evidence for an Inducible Error-Prone Repair System ("SOS Repair") and Its Anomalous Expression in these Strains, *Genetics*, 1975, vol. 79 suppl., pp. 199–213.
13. Friedberg, E.C., Walker, G.C., and Siede, W., *DNA Repair and Mutagenesis*, Washington: ASM, 1995, pp. 407–522.
14. Tang, M., Shen, X., Frank, E.G., et al., UmuD'$_2$C Is an Error-Prone DNA Polymerase, *Escherichia coli* pol V, *Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, pp. 8919–8924.
15. Reuven, N.B., Arad, G., Maor-Shoshani, A., and Livneh, Z., The Mutagenic Protein UmuC Is a DNA Polymerase Activated by UmuD', RecA, and SSB and Is Specialized for Translesion Replication, *J. Biol. Chem.*, 1999, vol. 274, pp. 31 763–31 766.
16. Goodman, M.F. and Tippin, B.T., Sloppier Copier DNA Polymerases Involved in Genome Repair, *Curr. Opin. Gen. Dev.*, 2000, vol. 10, pp. 162–168.
17. Wagner, J., Gruz, P., Kim, S.R., et al., The *dinB* Gene Encodes a Novel *E. coli* DNA Polymerase, DNA pol IV, Involved in Mutagenesis, *Mol. Cell*, 1999, vol. 4, pp. 281–286.
18. Kim, S.-R., Matsui, K., Yamada, M., et al., Roles of Chromosomal and Episomal *dinB* Genes Encoding DNA pol IV in Targeted and Untargeted Mutagenesis in *Escherichia coli, Mol. Genet. Genom*, 2001, vol. 266, pp. 207–215.
19. Yang, W., Portraits of a Y-Family DNA Polymerase, *FEBS Lett.*, 2005, vol. 579, pp. 868–872.
20. Shcherbakova, P.V. and Fijalkowska, I.J., Translesion Synthesis DNA Polymerases and Control of Genome Stability, *Front. Biosc.*, 2006, vol. 11, pp. 2496–2517.
21. Pavlov, Y.I., Shcherbakova, P.V., and Rogozin, I.B., Roles of DNA Polymerases in Replication, Repair and Recombination in Eukaryotes, *Int. Rev. Cytol.*, 2006, vol. 255, pp. 41–132.
22. Lone, S., Townson, S.A., Uljon, S.N., et al., Human DNA Polymerase Encircles DNA: Implications for Mismatch Extension and Lesion Bypass, *Mol. Cell*, 2007, vol. 25, pp. 601–614.
23. Nelson, J.R., Lawrence, C.W., and Hinkle, D.C., Deoxycytidyl Transferase Activity of Yeast REV1 Protein, *Nature*, 1996, vol. 382, pp. 729–731.
24. Haracska, L., Prakash, S., and Prakash, L., Yeast Rev1 Protein Is a G Template-Specific DNA Polymerase, *J. Biol. Chem.*, 2002, vol. 277, pp. 15 546–15 551.
25. Lin, W., Xin, H., Zhang, Y., et al., The Human *REV1* Gene Codes for a DNA Template-Dependent dCP Transferase, *Nucleic Acids Res.*, 1999, vol. 27, pp. 4468–4475.
26. Nair, D.T., Johnson, R.E., Prakash, L., et al., Rev1 Employs a Novel Mechanism of DNA Synthesis Using a Protein Template, *Science*, 2005, vol. 309, pp. 2219–2222.
27. Pavlov, Y.I., Nguyen, D., and Kunkel, T.A., Mutator Effects of Overproducing DNA Polymerase η(Rad30) and Its Catalytically Inactive Variant in Yeast, *Mutat. Res.*, 2001, vol. 478, pp. 129–139.
28. King, N.M., Nikolaishvili-Feinberg, N., Bryant, M.F., et al., Overproduction of DNA Polymerase Does not Raise the Spontaneous Mutation Rate in Diploid Human Fibroblasts, *DNA Repair* (Amsterdam), 2005, vol. 4, pp. 714–724.
29. Kusumoto, R., Masutani, C., Shimmyo, S., et al., DNA Binding Properties of Human DNA Polymerase eta: Implications for Fidelity and Polymerase Switching of Translesion Synthesis, *Genes Cells*, 2004, vol. 9, pp. 1139–1150.
30. McCulloch, S.D., Kokoska, R.J., Chilkova, O., et al., Enzymatic Switching for Efficient and Accurate Trans-

lesion DNA Replication, *Nucleic Acids Res.,* 2004, vol. 32, pp. 4665–4675.

31. Haracska, L., Acharya, N., Unk, I., et al., A Single Domain in Human DNA Polymerase iota Mediates Interaction with PCNA: Implications for Translesion DNA Synthesis, *Mol. Cell. Biol.*, 2005, vol. 25, pp. 1183–1190.
32. Haracska, L., Johnson, R.E., Unk, I., et al., Physical and Functional Interactions of Human DNA Polymerase eta with PCNA, *Mol. Cell Biol.*, 2001, vol. 21, pp. 7199–7206.
33. Haracska, L., Johnson, R.E., Unk, I., et al., Targeting of Human DNA Polymerase iota to the Replication Machinery via Interaction with PCNA, *Proc. Natl. Acad. Sci. USA*, 2001, vol. 98, pp. 14256–14261.
34. Hoege, C., Pfander, B., Moldovan, G.L., et al., RAD6-Dependent DNA Repair Is Linked to Modification of PCNA by Ubiquitin and SUMO, *Nature*, 2002, vol. 419, pp. 135–141.
35. Kannouche, P.L., Wing, J., and Lehmann, A.R., Interaction of Human DNA Polymerase eta with Monoubiquitinated PCNA: A Possible Mechanism for the Polymerase Switch in Response to DNA Damage, *Mol. Cell*, 2004, vol. 14, pp. 491–500.
36. Stelter, P. and Ulrich, H.D., Control of Spontaneous and Damage-Induced Mutagenesis by SUMO and Ubiquitin Conjugation, *Nature*, 2003, vol. 425, pp. 188–191.
37. Bienko, M., Green, C.M., Crosetto, N., et al., Ubiquitin-Binding Domains in Y-Family Polymerases Regulate Translesion Synthesis, *Science*, 2005, vol. 310, pp. 1821–1824.
38. Garg, P. and Burgers, P.M., Ubiquitinated Proliferating Cell Nuclear Antigen Activates Translesion DNA Polymerases eta and REV1, *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, pp. 18361–18366.
39. Huang, T.T., Nijman, S.M., Mirchandani, K.D., et al., Regulation of Monoubiquitinated PCNA by DUB Autocleavage, *Nat. Cell. Biol.,* 2006, vol. 8, pp. 341–347.
40. Ulrich, H.D., Deubiquitinating PCNA: A Downside to DNA Damage Tolerance, *Nat. Cell. Biol.,* 2006, vol. 8, pp. 303–305.
41. Miura, A., Ultraviolet-Induced Mutation of Bacteriophage λ, *Mol. Gen. Genet.*, 1974, vol. 134, pp. 21–27.
42. Milstein, C. and Rada, C., *The Maturation of Immune Response*, London: APress, 1995.
43. Samaranayake, M., Bujnicki, J.M., Carpenter, M., and Bhagwat, A.S., Evaluation of Molecular Models for the Affinity Maturation of Antibodies: Roles of Cytosine Deamination by AID and DNA Repair, *Chem. Rev.*, 2006, vol. 106, pp. 700–719.
44. Papavasiliou, F.N. and Schatz, D.G., Somatic Hypermutation of Immunoglobulin Genes: Merging Mechanisms for Genetic Diversity, *Cell*, 2002, vol. 109, pp. S35–S44.
45. Berek, C. and Milstein, C., The Dynamic Nature of the Antibody Repertoire, *Immunol. Rev.*, 1988, vol. 105, pp. 5–26.
46. Rogozin, I.B. and Kolchanov, N.A., Somatic Hypermutagenesis in Immunoglobulin Genes: II. Influence of Neighbouring Base Sequences on Mutagenesis, *Biochim. Biophys. Acta*, 1992, vol. 1171, pp. 11–18.
47. Neuberger, M.S. and Milstein, C., Somatic Hypermutation, *Curr. Opin. Immunol.*, 1995, vol. 7, pp. 248–254.
48. Storb, U., The Molecular Basis of Somatic Hypermutation of Immunoglobulin Genes, *Curr. Opin. Immunol.*, 1996, vol. 8, pp. 206–214.
49. Talmage, D.W., Immunological Specificity, Unique Combinations of Selected Natural Globulins Provide an Alternative to the Classical Concept, *Science*, 1959, vol. 129, pp. 1643–1648.
50. Sharma, J.M., The Structure and Function of the Avian Immune System, *Acta Vet. Hung.,* 1997, vol. 45, pp. 229–238.
51. Muramatsu, M., Sankaranand, V.S., Anant, S., et al., 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.,* 1999, vol. 274, pp. 18470–18476.
52. Muramatsu, M., Kinoshita, K., Fagarasan, S., et al., Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme, *Cell*, 2000, vol. 102, pp. 553–563.
53. Revy, P., Muto, T., Levy, Y., et al., Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2), *Cell*, 2000, vol. 102, pp. 565–575.
54. Arakawa, H., Hauschild, J., and Buerstedde, J.M., Requirement of the Activation-Induced Deaminase (AID) Gene for Immunoglobulin Gene Conversion, *Science*, 2002, vol. 295, pp. 1301–1306.
55. Harris, R.S., Sale, J.E., Petersen-Mahrt, S.K., and Neuberger, M.S., AID Is Essential for Immunoglobulin V Gene Conversion in a Cultured B Cell Line, *Curr. Biol*, 2002, vol. 12, pp. 435–438.
56. Teng, B., Burant, C.F., and Davidson, N.O., Molecular Cloning of an Apolipo-Protein B Messenger RNA Editing Protein, *Science*, 1993, vol. 260, pp. 1816–1819.
57. Scott, J., A Place in the World for RNA Editing, *Cell*, 1995, vol. 81, pp. 833–836.
58. Petersen-Mahrt, S.K., Harris, R.S., and Neuberger, M.S., AID Mutates *E. coli* Suggesting a DNA Deamination Mechanism for Antibody Diversification, *Nature*, 2002, vol. 418, pp. 99–104.
59. Mayorov, V.I., Rogozin, I.B., Adkison, L.R., et al., Expression of AID in Yeast Induces Mutations in the Context of Somatic Hypermutation in IG Genes, *BMC Immunol.*, 2005, vol. 6, p. 10.
60. Pham, P., Bransteitter, R., Petruska, J., and Goodman, M.F., Processive AID-Catalysed Cytosine Deamination on ssDNA Simulates SHM, *Nature*, 2003, vol. 424, pp. 103–107.
61. Poltoratsky, V., Prasad, R., Horton, J.K., and Wilson, S.H., Down-Regulation of DNA Polymerase Accompanies Somatic Hypermutation in Human BL2 Cell Lines, *DNA Repair*, 2007, vol. 6, pp. 244–253.
62. Diaz, M. and Lawrence, C., An Update on the Role of Translesion Synthesis DNA Polymerases in Ig Hypermutation, *Trends Immunol.*, 2005, vol. 26, pp. 215–220.
63. Delbos, F., De Smet, A., Faili, A., et al., Contribution of DNA Polymerase eta to Immunoglobulin Gene Hypermutation in the Mouse, *J. Exp. Med.,* 2005, vol. 201, pp. 1191–1196.

64. Martomo, S.A., Yang, W.W., Wersto, R.P., et al., Different Mutation Signatures in DNA Polymerase eta and MSH6-Deficient Mice Suggest Separate Roles in Antibody Diversification, *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, pp. 8656–8661.

65. Zeng, X., Winter, D.B., Kasmer, C., et al., DNA Polymerase eta Is an A-T Mutator in Somatic Hypermutation of Immunoglobulin Variable Genes, *Nat. Immunol.*, 2001, vol. 2, pp. 537–541.

66. Matsuda, T., Bebenek, K., Masutani, C., et al., Error Rate and Specificity of Human and Murine DNA Polymerase, *J. Mol. Biol.*, 2001, vol. 312, pp. 335–346.

67. Rogozin, I.B., Pavlov, Y.I., Bebenek, K., et al., Somatic Mutation Hotspots Correlate with DNA pol η Error Spectrum, *Nat. Immunol*, 2001, vol. 2, pp. 530–536.

68. Pavlov, Y.I., Rogozin, I.B., Galkin, A.P., et al., Correlation of Somatic Hypermutation Specificity and A-T Base Pair Substitution Errors by DNA Polymerase η during Copying of a Mouse Immunoglobulin Kappa Light Chain Transgene, *Proc. Natl. Acad. Sci. USA*, 2002, vol. 99, pp. 9954–9959.

69. Zhang, W., Bardwell, P.D., Woo, C.J., et al., Clonal Instability of V Region Hypermutation in the Ramos Burkitt's Lymphoma Cell Line, *Int. Immunol.*, 2001, vol. 13, pp. 1175–1184.

70. Yoshikawa, K., Okazaki, I.M., Eto, T., et al., AID Enzymeinduced Hypermutation in an Actively Transcribed Gene in Fibroblasts, *Science*, 2002, vol. 296, pp. 2033–2036.

71. Martin, A., Bardwell, P.D., Woo, C.J., et al., Activation-Induced Cytidine Deaminase Turns on Somatic Hypermutation in Hybridomas, *Nature*, 2002, vol. 415, pp. 802–806.

72. Neuberger, M.S., Harris, R.S., Di Noia, J., and Petersen-Mahrt, S.K., Immunity through DNA Deamination, *Trends Biochem. Sci.,* 2003, vol. 28, pp. 305–312.

73. Okazaki, I.M., Hiai, H., Kakazu, N., et al., Constitutive Expression of AID Leads to Tumorigenesis, *J. Exp. Med.*, 2003, vol. 197, pp. 1173–1181.

74. Matsumoto, Y., Marusawa, H., Kinoshita, K., et al., *Helicobacter Pylori* Infection Triggers Aberrant Expression of Activation-Induced Cytidine Deaminase in Gastric Epithelium, *Nature Medicine*, 2007, vol. 13, pp. 470–476.

75. Rada, C., Williams, G.T., Nilsen, H., et al., Immunoglobulin Isotype Switching Is Inhibited and Somatic Hypermutation Perturbed in UNG-Deficient Mice, *Curr. Biol.*, 2002, vol. 12, pp. 1748–1755.

76. Stavnezer, J. and Schrader, C., Mismatch Repair Converts AID Instigated Nicks to Double-Strand Breaks for Antibody Classswitch Recombination, *Trends Genet.*, 2006, vol. 22, pp. 23–28.

77. Reina-San-Martin, B., Chen, H.T., Nussenzweig, A., and Nussenzweig, M.C., ATM is Required for Efficient Recombination between Immunoglobulin Switch Regions, *J. Exp. Med.,* 2004, vol. 200, pp. 1103–1110.

78. Franco, S., Gostissa, M., Zha, S., et al., Histone H2AX Prevents DNA Breaks from Progressing to Chromosome Breaks and Translocations, *Mol. Cell*, 2006, vol. 21, pp. 201–214.

79. Vallur, A.C., Yabuki, M., Larson, E.D., and Maizels, N., AID in Antibody Perfection, *Cell. Mol. Life Sci.*, 2007, vol. 64, pp. 555–565.

80. Rada, C., Di Noia, J., and Neuberger, M., Mismatch Recognition and Uracil-Excision Provide Complementary Paths to Both Immunoglobulin Switching and the Second (dA:dT-Focussed) Phase of Somatic Mutation, *Mol. Cell*, 2004, vol. 16, pp. 163–171.

81. Poltoratsky, V.P., Wilson, S.H., Kunkel, T.A., and Pavlov, Y.I., Recombinogenic Phenotype of Human Activation-Induced Cytosine Deaminase, *J. Immunol.*, 2004, vol. 172, pp. 4308–4313.

82. Litman, G.W., Evolution of Antigen Binding Receptors, *Annu. Rev. Immunol.*, 1999, vol. 17, pp. 109–147.

83. Conticello, S.G., Thomas, C.J.F., Petersen-Mahrt, S.K., and Neuberger, M.S., Evolution of AID/APOBEC Family of Polynucleotide (Deoxy)Cytidine Deaminases, *Mol. Biol. Evol.*, 2006, vol. 22, pp. 367–377.

84. Pancer, Z., Amemiya, C.T., Ehrhardt, G.R., et al., Somatic Diversification of Variable Lymphocyte Receptors in the Agnathan Sea Lamprey, *Nature*, 2004, vol. 430, pp. 174–180.

85. Alder, M.N., Rogozin, I.B., Iyer, L.M., et al., Diversity and Function of Adaptive Immune Receptors in a Jawless Vertebrate, *Science*, 2005, vol. 310, pp. 1970–1973.

86. Pancer, Z., Saha, N.R., Kasamatsu, J., et al., Variable Lymphocyte Receptors in Hagfish, *Proc. Natl. Acad. Sci. USA,* 2005, vol. 102, pp. 9224–9229.

87. Nagawa, F., Kishishita, N., Shimizu, K., et al., Antigen-Receptor Genes of the Agnathan Lamprey Are Assembled by a Process Involving Copy Choice, *Nat. Immunol.*, 2007, vol. 8, no. 2, pp. 206–213.

88. Rogozin, I.B., Lakshminarayan, M.I., Liang, L., et al., Evolution and Diversification of Lamprey Antigen Receptors: Evidence for Involvement of an AID-APOBEC Family Cytosine Deaminase, *Nat. Immunol.*, 2007, vol. 8, no. 6, pp. 647–656.

89. Rogozin, I.B., Basu, M.K., Jordan, I.K., et al., APOBEC4, a New Member of the AID/APOBEC Family of Polynucleotide (Deoxy)Cytidine Deaminases Predicted by Computational Analysis, *Cell Cycle*, 2005, vol. 4, pp. 1281–1285.

90. Lewine, B., *Genes VII*, Oxford: Oxford Univ. Press, 2000.

91. Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H., Isolation of a Human Gene that Inhibits HIV-1 and Is Suppressed by the Viral Vif, *Nature*, 2002, vol. 418, pp. 646–650.

92. Douaisi, M., Dussart, S., Courcoul, M., et al., HIV-1 and MLV GAG Proteins Are Sufficient to Recruit APOBEC3G into Virus-Like Particles, *Biochem. Biophys. Res. Comm.,* 2004, vol. 321, pp. 566–573.

93. Sheehy, A.M., Gaddis, N.C., and Malim, M.H., The Antiretroviral Enzyme APO-BEC3G Is Degraded by the Proteasome in Response to HIV-1 VIF, *Nature Med.*, 2003, vol. 9, pp. 1404–1407.

94. Chiu, Y.-L. and Greene, W.C., APOBEC3 Cytidine Deaminases: Distinct Antiviral Actions along the Retroviral Life Cycle, *JBC*, 2006, vol. 281, no. 13, pp. 8309–8312.

95. Franca, R., Spadari, S., and Maga, G., APOBEC Deaminases as Cellular Antiviral Factors: A Novel Natural

Host Defence Mechanism, *Med. Sci. Monit.*, 2006, vol. 12, no. 5, pp. RA92–98.

96. Turelli, P. and Trono, D., Editing at the Crossroad of Innate and Adaptive Immunity, *Science*, 2005, vol. 307, pp. 1061–1065.
97. Waterston, R.H., Lindblad-Toh, K., Birney, E., et al., Initial Sequencing and Comparative Analysis of the Mouse Genome, *Nature*, 2002, vol. 420, pp. 520–562.
98. Bogerd, H.P., Wiegand, H.L., Doehle, B.P., and Cullen, B.R., The Intrinsic Antiretroviral Factor APOBEC3B Contains Two Enzymatically Active Cytidine Deaminase Domains, *Virology*, 2007 (in press).
99. Prochnow, C., Bransteitter, R., Klein, M.G., et al., The APOBEC-2 Crystal Structure and Functional Implications for the Deaminase AID, *Nature*, 2007, vol. 445, pp. 447–451.
100. Zhang, K.-L., Mangeat, B., Ortiz, M., et al., Model Structure of Human APOBEC3G, *PLoS ONE*, 2007, vol. 4, p. e378.
101. Rubio, M.A.T., Pastar, I., Gaston, K.W., et al., An Adenosine-to-Inosine tRNA-Editing Enzyme that Can Perform C-to-U Deamination of DNA, *Proc. Natl. Acad. Sci. USA,* 2007, vol. 104, pp. 7821–7826.
102. Lobashev, M.E., Physiological Hypothesis of the Mutation Process, in *Issledovaniya po genetike* (Research in Genetics), Leningrad: Leningr. Gos. Univ., 1976, issue 6, pp. 3–15.
103. Khromov-Borisov, N.N., Physiological Theory of Mutation Process a Quarter of a Century Later, in *Issledovaniya po genetike* (Research in Genetics), Leningrad: Leningr. Gos. Univ., 1976, issue 6, pp. 16–31.