A-to-I RNA editing — immune protector and transcriptome diversifier

Eli Eisenberg [1]* *and Erez Y. Levanon* [2]*

Abstract | Modifications of RNA affect its function and stability. RNA editing is unique among these modifications because it not only alters the cellular fate of RNA molecules but also alters their sequence relative to the genome. The most common type of RNA editing is A-to-I editing by double-stranded RNA-specific adenosine deaminase (ADAR) enzymes. Recent transcriptomic studies have identified a number of 'recoding' sites at which A-to-I editing results in non-synonymous substitutions in protein-coding sequences. Many of these recoding sites are conserved within (but not usually across) lineages, are under positive selection and have functional and evolutionary importance. However, systematic mapping of the editome across the animal kingdom has revealed that most A-to-I editing sites are located within mobile elements in non-coding parts of the genome. Editing of these non-coding sites is thought to have a critical role in protecting against activation of innate immunity by self-transcripts. Both recoding and non-coding events have implications for genome evolution and, when deregulated, may lead to disease. Finally, ADARs are now being adapted for RNA engineering purposes.

Non-coding RNAs
RNA transcripts that are not translated into proteins but may have a regulatory function.

MicroRNAs
Short non-coding RNAs that regulate gene expression post-transcriptionally, mainly by binding to the 3′ untranslated region of mRNA.

Circular RNAs
RNA molecules that form a covalently closed continuous loop. They were found only recently, and the function of most of them is not known.

[1]*Raymond and Beverly Sackler School of Physics and Astronomy and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.*

[2]*The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel.*

**e-mail: elieis@post.tau.ac.il; erez.levanon@biu.ac.il*

https://doi.org/10.1038/s41576-018-0006-1

The transcriptome is defined as the total set of RNAs expressed in a cell, tissue or organism and comprises protein-coding mRNAs and a variety of non-coding RNAs, including microRNAs (miRNAs) and circular RNAs (circRNAs). The composition and complexity of the transcriptome are believed to account for the increased cellular and functional sophistication of higher eukaryotes[1] and are determined in part by RNA processing events, including alternative splicing[1], alternative polyadenylation[2] and RNA modifications. More than 100 distinct types of RNA modifications have been identified[3], and for some, their effects on the fate of the RNA molecule have been determined; for example, mRNA methylation is involved in transcription, nuclear export, translation and degradation of mRNA[4]. However, our understanding of most modifications is limited by the inability of standard sequencing technologies to distinguish between modified and unmodified RNA molecules[5], which makes them difficult to detect and map. An important exception is RNA editing, a common RNA modification that alters the RNA sequence itself; edited sites can then be detected as mismatches between the RNA sequence and the DNA sequence from which it originated.

Several types of RNA editing have been characterized so far (FIG. 1a). For example, mitochondrial transcripts in *Trypanosoma brucei* undergo extensive insertions and deletions of uracils[6]. Transcripts in plant chloroplasts and mitochondria exhibit multiple cytosine-to-uracil editing events[7]. In mammals, cytosines are deaminated to uracils by members of the apolipoprotein B mRNA editing enzyme catalytic subunit (APOBEC) protein family[8]. This Review focuses on the most abundant form of RNA editing in Metazoans, adenosine-to-inosine (A-to-I) editing, in which enzymes encoded by the adenosine deaminase acting on RNA (*ADAR*) gene family catalyse deamination of adenosine nucleotides to inosines[9–11] (FIG. 1b,c). Three members of this family are encoded in the mammalian genome: double-stranded RNA-specific adenosine deaminase (ADAR1), double-stranded RNA (dsRNA)-specific editase 1 (ADAR2; also known as ADARB1) and ADAR3 (also known as ADARB2) (FIG. 1b). Both ADAR1 and ADAR2 are expressed ubiquitously, but ADAR1 is expressed more strongly and is responsible for the majority of editing activity[12]. The catalytically inactive ADAR3 is expressed primarily in the brain at a relatively low level[12].

ADAR proteins bind a specific dsRNA structure; thus, ADAR-mediated A-to-I editing occurs only on RNAs that adopt this structure (FIG. 2a). There is also an RNA sequence motif associated with editing sites[13–15]. However, our understanding of the features controlling ADAR target recognition is incomplete and does not allow prediction of editing sites. Identification of editing sites is therefore dependent on sequencing data. During RNA sequencing (RNA-seq), reverse transcriptase incorporates guanosines into the cDNA molecule at positions where inosines occur in the RNA; the equivalent positions in the genomic DNA sequence contain adenosines. Thus, A-to-I RNA editing sites can be systematically and directly detected as adenosine-to-guanosine (A → G)

RNA modifications
Changes to the chemical composition of RNA molecules that have the potential to alter their function or stability.

Editome
The entire set of RNA editing events in a genome.

Non-synonymous substitutions
Replacement of one base by another within a coding region of a gene, which results in an amino acid change in the protein sequence.

mismatches in RNA-seq data aligned to a reference genome[16–19] (FIG. 2b). In practice, there are many other sources of DNA–RNA sequence mismatches, including sequence polymorphisms and sequencing errors, which can make it difficult to identify true editing sites. Various algorithmic approaches that have been devised to overcome this difficulty (BOX 1; FIG. 2c–e), in tandem with rapid advances in sequencing technologies, have allowed for systematic transcriptome-wide identification and quantification of editing sites across multiple species[20] (FIG. 1c). They have enabled the detection of millions of human A-to-I editing sites[21,22], making it the best-mapped modification to date and fundamentally changing our understanding of RNA editing.

In this Review, we outline the current view of the A-to-I editome. We summarize the different classes of editing sites and their biological importance, discuss how evolution shapes the editome and how the editome influences evolution and briefly address the implications of RNA editing for pathologies and future genome engineering technologies. The biochemistry of the ADAR enzymes, the role of specific RNA editing sites in disease and the function of recoding of specific targets have been extensively reviewed previously[10,11,13,23–27] and are beyond the scope of this Review.

A new view of the editome

When edited RNAs are processed, the ribosomes and splicing machinery interpret inosines as guanosines[28] instead of the adenosines encoded in the genome. Editing is classified as 'recoding' if these A → G mismatches occur in protein-coding sequences and lead to non-synonymous substitutions that generate novel protein variants. By contrast, 'non-coding' editing occurs in non-coding RNAs or non-coding parts of mRNAs and does not alter the protein sequence. Systematic transcriptome-wide analyses of editomes from multiple species have provided new insights into the prevalence, distribution and functional importance of recoding and non-coding editing.

Recoding sites. The first discoveries of A-to-I RNA editing were serendipitous[29] and identified a handful of mammalian recoding sites in glutamate receptors[30] and the serotonin 5-hydroxytryptamine$_{2c}$ (5-HT$_{2c}$) receptor[31]. Understanding the functional consequences of these recoding events has been the main focus of research for many years, largely owing to interest in their ability to diversify the proteome by introducing novel protein isoforms. The introduction of algorithmic approaches, supplemented by deep sequencing of multiple DNA and RNA samples originating from different tissues and individuals, has greatly increased the variety of known recoding sites. For example, reports have documented more than a thousand recoding sites in humans[21,22] and several *Drosophila* species[32–36], and tens of thousands in various coleoid cephalopod species[37,38]. These studies confirm that recoding is enriched in neural tissues and show that recoding sites are over-represented in transcripts that encode proteins with functions in the nervous system, such as ion channels and neuroreceptors[23]. Notable examples include the mammalian recoding sites in the serotonin 5-HT$_{2c}$ receptor, in which editing reduces the affinity of the receptor for its G protein[31], and the Shaker voltage-dependent potassium channels[29] in mammals, *Drosophila* and cephalopods, in which editing affects the kinetics and voltage dependence of activation[37,39,40].

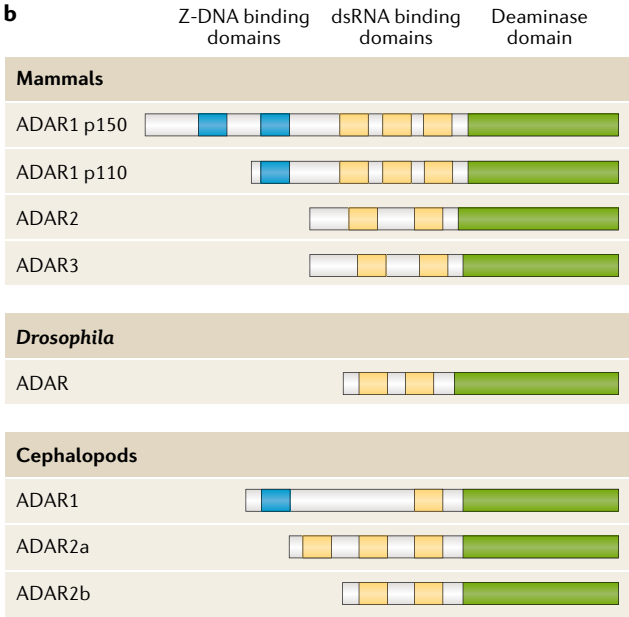
Nevertheless, recoding also occurs in non-neural transcripts, such as the mammalian *FLNA* (REF.[41]), *AZIN1* (REF.[42]) and *NEIL1* (REF.[42]) sites. These recoding targets are ubiquitously expressed, are edited in multiple tissues and have functions seemingly unrelated to the brain. Filamin A (FLNA) is an actin crosslinker that is widely expressed in most tissues. Its transcript is targeted mainly by ADAR2 (REF.[43]), and its editing levels are significantly and consistently altered in various types of cancer[44]. Recoding of antizyme inhibitor 1 (AZIN1) by ADAR1 causes it to translocate from the cytoplasm to the nucleus, increases its affinity to antizyme and promotes cell proliferation through the neutralization of antizyme-mediated degradation of ornithine decarboxylase (ODC) and G1/S-specific cyclin D1 (CCND1)[45]. Endonuclease 8-like 1 (NEIL1) is a DNA repair enzyme, and its recoding by ADAR1 results in a 30-fold reduction in the rate at which it removes thymine glycol from duplex DNA[46].

Fig. 1 | **A-to-I RNA editing is catalysed by ADAR enzymes and is the most common type of RNA editing in Metazoa. a** | Four main types of mRNA editing have been studied in recent decades. A-to-I RNA editing is the most common in terms of the range of organisms affected, the breadth of tissues edited and the number of editing sites. **b** | A-to-I editing is catalysed by double-stranded RNA (dsRNA)-specific adenosine deaminase (ADAR) enzymes. Three members of this highly conserved family are encoded in the mammalian genome: ADAR1, ADAR2 and ADAR3. The enzymes contain a catalytic deaminase domain and two or three dsRNA binding domains and bind to specific dsRNA structures. In addition to the constitutive 110 kDa isoform of ADAR1 (p110), which localizes primarily to the nucleus, the mammalian *ADAR* gene expresses a longer, interferon-inducible isoform, p150 (REF.[169]). ADAR1 p110 and ADAR2 are expressed ubiquitously; however, ADAR1 is more strongly expressed and is responsible for most of the editing activity in mammals[12]. ADAR3 is expressed primarily in the brain at a relatively low level[12]. It is believed to be catalytically inactive[170] and to inhibit editing at specific sites[47,93]. Insects have lost ADAR1, and the ADAR found in *Drosophila* spp. is more similar to mammalian ADAR2 (REF.[171]). Cephalopods have an ADAR1 protein and two variants of ADAR2 (REF.[172]). ADAR2a contains an additional exon that encodes an extra dsRNA binding domain, which increases its in vitro editing activity and confers resistance to the high salt levels found in squid neurons[173], and ADAR2b mimics the conventional ADAR2 family members. Both variants are expressed at comparable levels and are extensively edited, each in a unique pattern. **c** | A-to-I editing has been detected in all Metazoan species screened so far, from coral to mammals[16–19,21,36–38,42,73,74,93,174–179]. APOBEC, apolipoprotein B mRNA editing enzyme catalytic subunit; PPR, pentatricopeptide repeat proteins; RECC, RNA editing core complex. Part **b** is adapted with permission from REF.[104], PLOS, CC-BY-4.0 (https://creativecommons.org/licenses/by/4.0/).

a

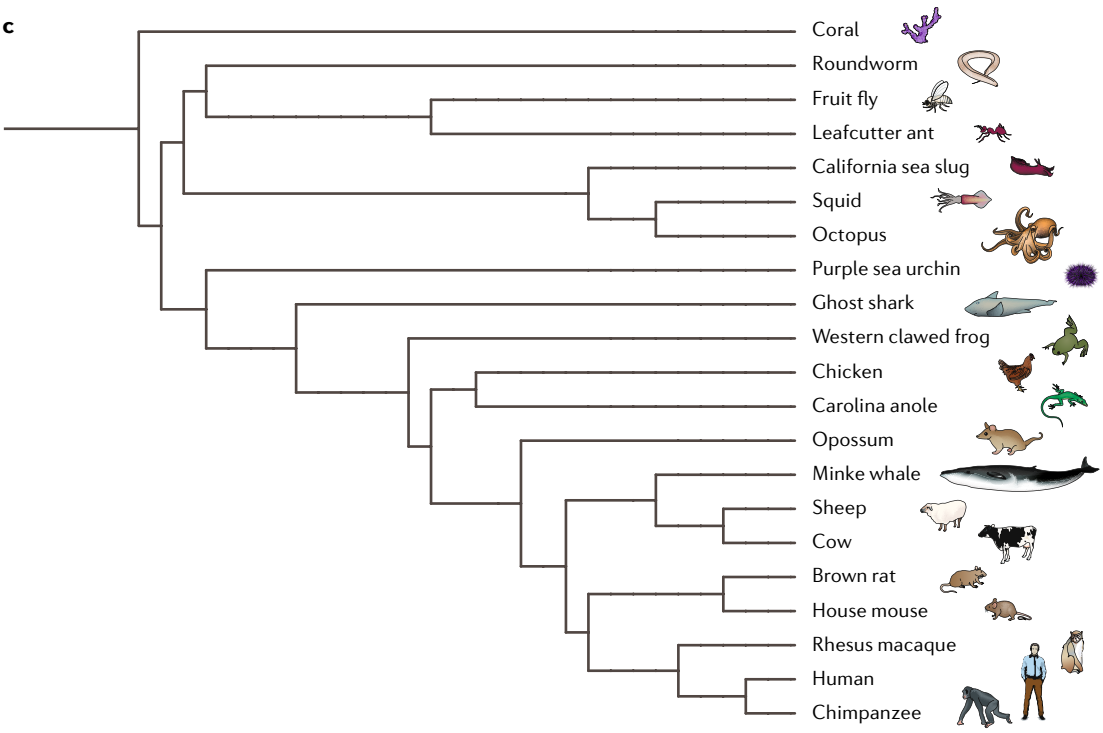
Year discovered	Editing type	Mediated by	Number of editing sites	Targeted transcripts	Organisms
1986	Insertion and deletion of U	RECC	1,000s	Mitochondrial mRNAs; recoding	Kinetoplastids
1987	C-to-U	APOBECs	100s	Mostly non-coding; in a few tissues	Mammals
1989	C-to-U	PPR proteins	100s	Chloroplast and mitochondrial mRNA; mostly recoding	Plants
1991	A-to-I	ADAR proteins	1,000,000s	Mostly non-coding	Metazoa

b



	Expression	Function
Mammals		
ADAR1 p150	Interferon-induced	Non-coding
ADAR1 p110	Ubiquitous	Non-coding
ADAR2	A few select tissues	Recoding
ADAR3	Mostly neural tissues	Possibly editing inhibition
Drosophila		
ADAR	Mostly neural tissues	Mainly recoding
Cephalopods		
ADAR1	Unknown	Recoding and non-coding
ADAR2a	Unknown	Recoding and non-coding
ADAR2b	Unknown	Recoding and non-coding

c



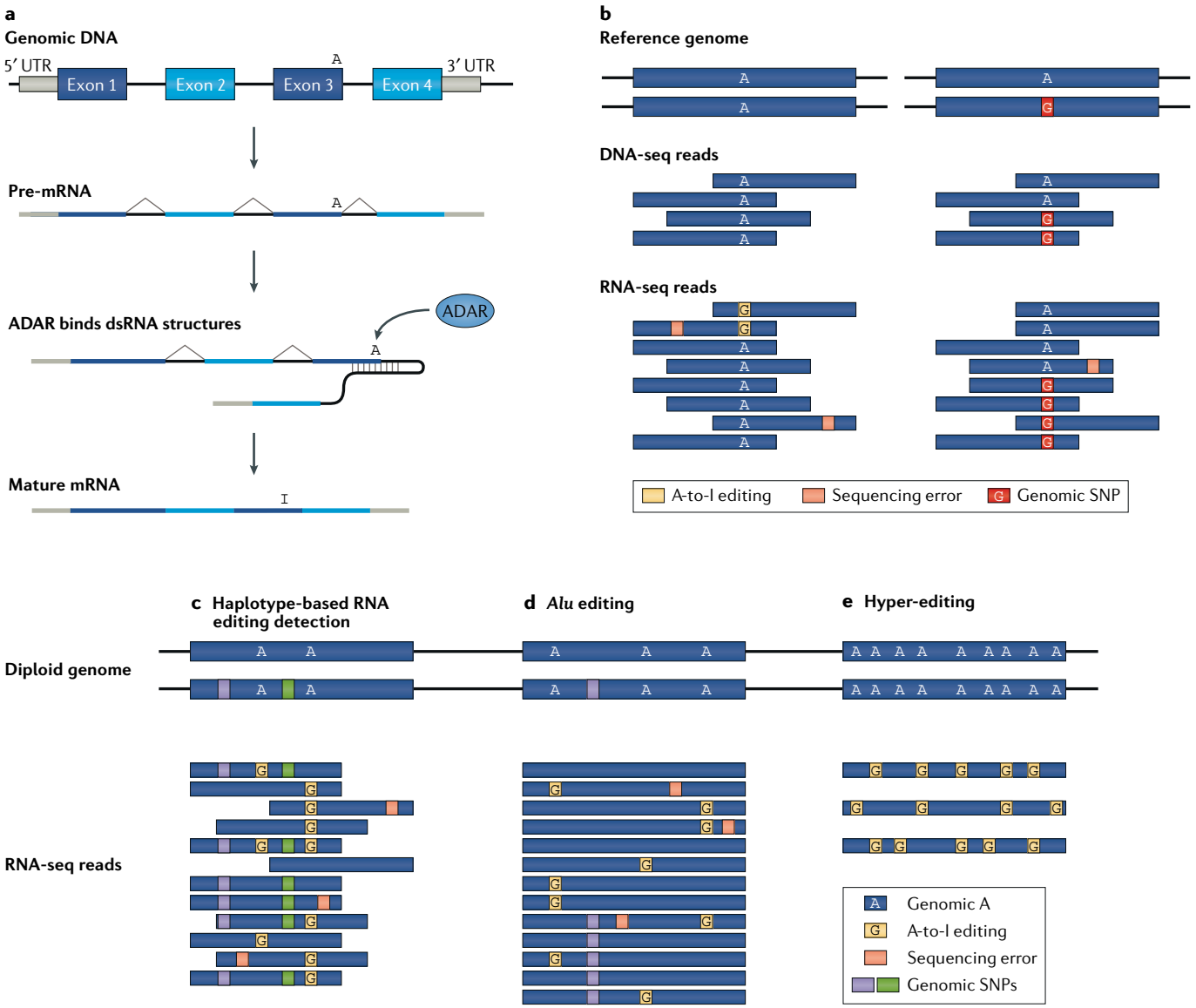


Fig. 2 | **A-to-I RNA editing and how it is detected. a** | Pre-mRNA molecules transcribed from the genome may fold to form double-stranded RNA (dsRNA) secondary structures. dsRNA-specific adenosine deaminase (ADAR) enzymes bind these structures and deaminate some adenosines to inosines. If these inosines are located in an exon, they will be present in the mature mRNA. **b** | Reverse transcription replaces inosines in mRNA with guanosines in the cDNA. Thus, the hallmark of RNA editing is a consistent A → G mismatch between RNA sequencing (RNA-seq) data and the reference genomic sequence to which it is aligned. However, most of these mismatches arise from sequencing errors and genomic polymorphisms, including somatic mutations and incorrect alignment[145–150]. Matched DNA sequencing (DNA-seq) data may be utilized to distinguish between editing events and genomic polymorphisms. At an editing site, the DNA reads agree with the genome reference (left), while a genomically polymorphic site exhibits mismatches in both DNA-seq and RNA-seq data (right). **c** | In the absence of matching DNA data, genomic polymorphisms may be filtered based on their allelic linkage pattern (haplotype), which is distinctive from the correlation observed between neighbouring editing sites. **d** | The majority of editing sites reside within repetitive elements, such as the primate-specific *Alu* element. Searching for mismatches confined to repetitive genomic loci helps identify editing sites, which may cause mismatches in these regions more frequently than biological noise. **e** | RNA-seq reads harbouring a large number of mismatches of the same type (for instance, hyper-edited reads) almost always result from a cluster of editing events. Identifying these highly edited reads allows for high-accuracy detection of clusters of editing sites in the absence of matching DNA data. SNP, single nucleotide polymorphism; UTR, untranslated region.

In mammals, recoding of editing sites is the main function of the ADAR2 protein[47]. This enzyme edits the majority of neural recoding sites. However, gene knockout studies in mice[48] suggest that the only essential target of ADAR2 is the Q/R recoding site within *GRIA2* transcripts[30]. *GRIA2* encodes the main subunit of glutamate receptor 2 (GRIA2), an α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor that mediates most fast excitatory synaptic transmission. Mice lacking the *Adar2* gene suffer from progressive

Box 1 | **Detection and quantification of RNA editing sites**

Accurate identification of editing sites is a challenging task. Technical noise (such as sequencing errors and incorrect alignment) and biological noise (such as genomic polymorphisms, including somatic mutations) vastly outweigh the editing signal[145–150]. Furthermore, most sites are edited infrequently, making them difficult to detect from low-coverage RNA sequencing (RNA-seq) data. Over recent years, various statistical models, filtering approaches and clustering schemes have been devised to determine which A → G mismatches in a data set represent bona fide editing sites[151–153].

Sequencing errors are usually curated by focusing on high-quality reads and bases and retaining only events supported by multiple reads[151]. Given a known a priori error rate, a statistical model (for example, a binomial model for statistically independent errors) can filter out those mismatches that are likely to be caused by sequencing errors. This method becomes more effective with increasing coverage and sequencing accuracy. A DNA sequencing (DNA-seq) sample from the same individual greatly facilitates identification and removal of genomic polymorphisms[151,154,155] (FIG. 2b). However, advances in sequencing technologies have increased the availability of high coverage multi-sample data sets, which allows most polymorphisms and somatic mutations to be identified and removed even in the absence of matched DNA-seq data. Mismatches that recur in RNA data originating from multiple individuals are unlikely to result from rare genomic polymorphisms or somatic mutations[156]. In addition, neighbouring genomic polymorphisms exhibit a distinct correlation pattern (haplotype) that differs from the partial correlation typical of editing sites (FIG. 2c) and can be used to filter them out[157]. Systematic alignment errors are particularly challenging, as they are consistent across samples and may be supported by a large number of reads[148–150]. They often occur near exon–intron junctions when short overhangs of a few nucleotides are misattributed to an incorrect exon–exon junction[146] or in duplicated genomic regions (mainly pseudogenes) when alignment programmes map the read to the wrong copy. In order to tackle this problem, regions that are prone to systematic alignment errors can be masked, discarding all mismatches in these regions.

Different computational approaches have been devised and optimized for specific sets of editing sites. For example, many editing events, such as those in *Alu* sequences, occur in clusters (FIG. 2d), but mismatches resulting from noise usually do not. This feature may be used to substantially improve the signal-to-noise ratio[16–19,72,76,157,158]. Looking at the accumulated distribution of mismatches in a given locus, instead of testing each position separately[72], allows for the additive signal of multiple editing events in a cluster to be clearly detected. A notable example is the case of hyper-edited reads (FIG. 2e), where individual RNA reads contain so many editing-induced mismatches that they cannot be mapped by standard alignment tools. These reads are usually discarded, leaving this class of editing events completely overlooked. However, a designated alignment scheme has been developed that allows these reads to be mapped correctly by ignoring A → G mismatches[159,160], yielding a very low false-positive rate for clusters of editing sites and providing a powerful tool for high-accuracy detection of clusters of editing sites in the absence of matching DNA data. Another example of an approach that is specific to a subset of the editome is the detection of conserved recoding sites. Recoding sites are often few in number and dispersed, making them challenging to identify with confidence. In fact, newly developed all-purpose detection schemes often show an impressive performance, which mostly represents the large number of *Alu* sites that can be readily detected but do not perform well at detecting the challenging recoding sites. However, conserved recoding sites can be located using a cross-genome approach[29,32,41,55]. These examples emphasize the need for dedicated approaches for identifying different classes of editing sites with different characteristics.

Accurate quantification of editing levels and detection of differential editing across the transcriptome are even more challenging. Typical coverage of 100 reads per site results in an absolute error of up to 10% owing to statistical noise. Ultra-deep sequencing of a targeted pre-determined subset of editing sites[123] or enrichment approaches[161–164] can be used for high-precision studies of differential editing levels. However, quantification of sites that are edited at a very low frequency is inaccurate even with reasonably high-coverage sequencing data. In these cases, an editing index can be used[76], which is an averaged measure that aggregates the accumulated data from multiple sites. Such indices, if used consistently and with proper normalization, could allow systematic global comparison of editing activity across the increasingly growing volume of publicly available RNA-seq data.

Some useful resources for investigating A-to-I editing are listed in the table below.

Tool	URL	Description
Databases		
RADAR[21]	http://rnaedit.com/	Database of RNA editing sites in humans, mice and flies
REDIportal[22]	http://srv00.recas.ba.infn.it/atlas/	Database of RNA editing sites in humans; Includes information about editing in various tissues
DARNED[166]	http://darned.ucc.ie/about/	First RNA editing database on the web
RNA editing detection packages		
JACUSA[167]	https://github.com/dieterich-lab/JACUSA	Tool for site-specific identification of RNA editing events from replicate sequencing data
RNAEditor[159]	http://rnaeditor.uni-frankfurt.de/	A tool to detect editing sites from deep sequencing data
RES-Scanner[168]	https://github.com/ZhangLabSZ/RES-Scanner	A software package for genome-wide identification of RNA editing sites
GIREMI[158]	https://www.ibp.ucla.edu/research/xiao/GIREMI.html	Genome-independent identification of RNA editing by mutual information
REDItools[153]	http://reditools.sourceforge.net/	Suites of python scripts developed to study RNA editing at a genomic scale by next-generation sequencing data
SPRINT[169]	https://github.com/jumphone/SPRINT	An integrated tool to detect RNA editing and hyper-edited reads in the absence of SNP database

SNP, single nucleotide polymorphism.

Low-coverage
Arises when the number of reads that include a given nucleotide is insufficient to provide reliable variant calling at that position in the reconstructed sequence.

Ultra-deep sequencing
The application of massively parallel sequencing methods to a small set of targets, yielding much higher read coverage than that obtained from standard whole-transcriptome RNA sequencing data.

Purifying selection
Selective removal of deleterious alleles from the general population.

Retro-elements
Mobile elements that move around the genome through transcription into RNA followed by reverse transcription.

Mobile elements
DNA fragments that can move around within the genome. Most of the mammalian genome is composed of sequences derived from mobile genetic elements.

seizures and die within 3 weeks of birth, but this severe phenotype is completely rescued by altering their genome to encode an arginine instead of a glutamine at the GRIA2 Q/R recoding site[48,49]. It thus follows that editing of other ADAR2 targets is not essential in mice.

Overall, only a small proportion of recoding sites have been characterized[30,31,45,46,50–54], and the functional implications of recoding remain largely unknown. Furthermore, the functional importance of many recoding sites is doubtful. Of the more than 1,000 recoding sites reported in humans, only a few dozen are conserved across mammals[55] — that is, editing occurs at the same location in multiple species. The vast majority of human recoding sites seems to be restricted to human or the primate lineage. These non-conserved recoding sites do not show signs of selection — that is, they are less abundant and more-weakly edited compared with editing at synonymous sites, and they are under-represented in essential genes, highly expressed genes and genes that are under purifying selection. These observations imply that most non-conserved recoding sites are unlikely to be of functional importance[56]. By contrast, comparative studies of editomes in other lineages show a different picture: hundreds of sites are conserved across the *Drosophila* lineage[32–34], and more than 10,000 conserved sites are documented in cephalopods[37], which implies that functional recoding is abundant, both in terms of the actual number of sites and their proportion of total recoding sites.

Non-coding sites. Another important outcome of the systematic searches for editing sites was the understanding that the vast majority of RNA editing activity occurs in non-coding parts of the transcriptome. Although non-coding editing does not directly alter the protein-coding capacity of an mRNA like editing at recoding sites (FIG. 3a), it has a number of functional implications. Editing can generate new protein isoforms by altering the splicing pattern of the pre-mRNA[57] (FIG. 3b). Furthermore, the cellular fate of an mRNA and/or its likelihood of being translated can be affected by editing of miRNA binding sites in its 3′ untranslated region (UTR)[58] or by editing of the cognate miRNAs themselves[59] (FIG. 3c,d). Indeed, large-scale screens have revealed that there are dozens of editing sites within miRNAs and, if these sites occur within the seed region, editing can alter the set of target mRNAs recognized by them[58,60,61]. Furthermore, it was recently suggested that RNA editing affects the biogenesis of circRNA, a newly discovered type of long-lived RNA molecule; the biological function of most circRNAs is still unclear (FIG. 3e). Editing disrupts annealing between the introns that flank the circRNA, thus antagonizing its formation and expression[62]. However, these roles are unlikely to be the main purpose of the abundant non-coding editing activity, which takes place mostly in non-conserved, probably non-functional, parts of the transcriptome.

It has recently been shown that ADAR1-mediated editing of endogenous dsRNA is required to prevent activation of the cytosolic innate immune system and that this is, most probably, the essential function of ADAR1 editing[63–65]. Endogenous (or 'self') dsRNAs resemble structures commonly formed by viruses[66] and trigger the innate immune system by activating melanoma differentiation-associated protein 5 (MDA5; also known as IFIH1), a dsRNA sensor. MDA5 binds long (hundreds of bps)[67], near-perfect dsRNAs (FIG. 3f). Upon activation, MDA5 interacts with the mitochondrial antiviral signalling protein (MAVS), leading to an interferon response that severely damages the host cells[65]. In human cells, a lack of ADAR1 results in translational shutdown and cell death owing to hyperactivation of protein kinase RNA-activated (PKR; also known as EIF2AK2), a dsRNA sensor[68]. ADAR1-mediated editing of the offending self dsRNAs disrupts their base pairing to the extent that they are no longer recognized by MDA5, and thereby prevents inappropriate activation of the antiviral cellular immune system[63–65]. Indeed, deletion of the *Mda5* gene in mice that have editing-deficient ADAR1 enzymes rescues their lethal phenotype and results in a normal lifespan and normal behaviour, suggesting that protection against autoimmunity is the only essential role of ADAR1 editing[69,70,71].

A typical ADAR target consists of a long dsRNA duplex, usually formed by paired inverted copies of genomic retro-elements in introns or UTRs of the transcript (FIG. 4). For instance, >99% of the millions of editing sites detected in the human genome[72] are located in the primate-specific *Alu* repeats[16–19]. *Alu* repeats are typically clustered in gene-rich genomic regions, especially in introns. Pairs of inverted copies of the *Alu* repeat residing within the same pre-mRNA transcript may bind together to form dsRNA structures within the pre-mRNA. These dsRNAs are the ideal target for ADAR enzymes. All multicellular Metazoans screened so far[73], including corals[74] (belonging to Cnidaria, an early-branching Metazoan phylum), exhibit extensive editing of non-coding sequences (FIG. 1c). The extent of non-coding editing varies considerably across organisms and strongly depends on the properties of the mobile elements present in the genome[73,75] (FIG. 4A): high numbers of repetitive elements from lowly diverged families will generate large numbers of dsRNA structures that may be potential ADAR targets. However, in all cases studied to date, non-coding editing substantially outweighs recoding activity, even in cephalopods, in which recoding is extraordinarily abundant[37].

It remains unclear what proportion of non-coding editing is involved in this immune-protective function or which transcripts are likely to be recognized by MDA5 (REF.[63]). Pairing of neighbouring inverted repetitive elements results in long dsRNA duplexes, typically hundreds of nucleotides long, which are usually stable to a single nucleotide change. Assuming that ADAR1-mediated immune protection is achieved by destabilizing the dsRNA structure, multiple editing events would be required in virtually every copy of a given target's pre-mRNA to unwind it sufficiently so that it is not recognized by MDA5. However, the majority of ADAR1 targets in repetitive elements are only weakly edited; that is, the number of inosines per RNA molecule is often very low. For instance, human *Alu* editing sites are typically edited at levels <1%, and even though dozens of sites exist in any given *Alu*, the average number of

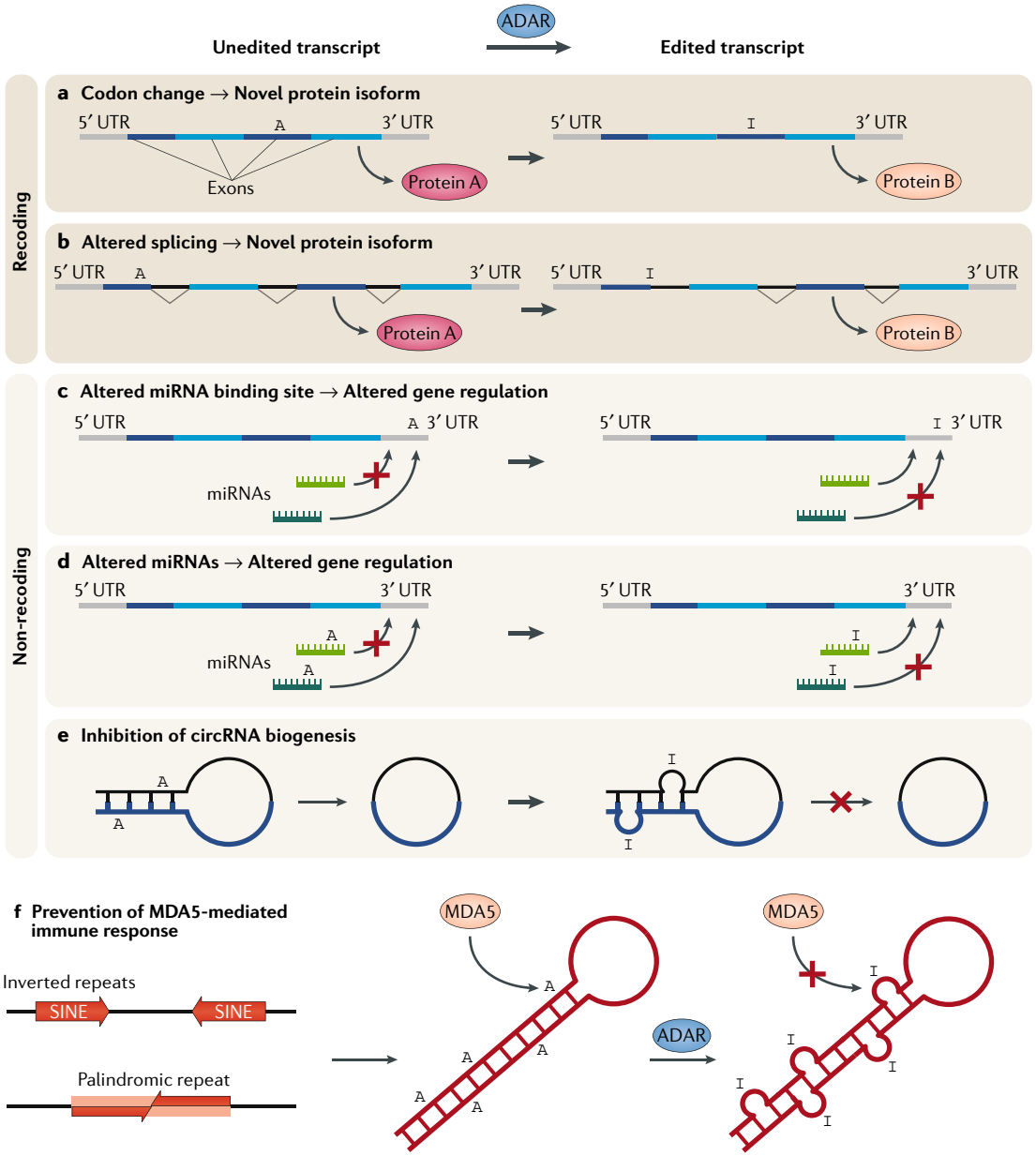


Fig. 3 | **Editing can modify protein function, generate new protein products, alter gene regulation and provide immune protection against endogenous dsRNAs. a** | If editing occurs within the coding part of a transcript, it can result in an amino acid substitution (recoding) and a novel protein isoform. **b** | Editing can also generate novel protein isoforms by disrupting splicing signals or creating new splice sites. **c** | Editing in non-coding RNA may have regulatory implications. It can create or destroy microRNA (miRNA) recognition sites in the untranslated region (UTR) of the transcript that may be involved in either translational repression or mRNA degradation. **d** | Editing of the miRNA sequence itself could reprogramme the set of targets that it recognizes. **e** | Circular RNA (circRNA) processing could also be affected and disrupted by editing of the parent sequence. **f** | However, the main and probably ancestral role of double-stranded RNA (dsRNA)-specific adenosine deaminase (ADAR) editing is to prevent dsRNA triggering an innate immune response mediated by melanoma differentiation-associated protein 5 (MDA5). Editing tends to unwind the secondary structure of dsRNAs, which prevents their misrecognition as non-self by MDA5 and false activation of the innate immune system. SINE, short interspersed nuclear element.

inosines per transcribed *Alu* is less than one in some tissues[72]. This level of editing is unlikely to have a substantial effect on secondary structure. Moreover, the majority of individual RNA molecules harbouring editable *Alu* elements are not edited at all[72,76]. It follows that most of the dsRNA structures targeted and edited by ADAR1 do not pose a real risk of activating the innate immune system, presumably reflecting the requirement of MDA5 for longer, more strongly paired dsRNA structures. Pinpointing the critical ADAR1 targets that do pose such a risk is still an unmet challenge, but the prime candidates are those targets that are edited multiple times per copy and that are expressed only in an extensively edited form.

RNA editing and evolution

Editing of invading mobile elements: genome evolution and novel recoding sites. RNA editing does not change the genome sequence itself. Yet the editing process has an important effect on genome evolution. In addition to preventing an immune response to transcripts containing repeat elements that have been long established in the genome, RNA editing also facilitates the introduction of new repeat elements to the genome. When a new active mobile element integrates into a genome, it can quickly replicate over an evolutionarily short period of time and generate nearly identical copies at multiple genomic locations that are often quite close together. Unlike their more established counterparts that have accumulated multiple mutations and diverged, these new nearly identical repeat elements can form extremely stable dsRNA structures when expressed as inverted repeats; these structures are more likely to elicit a strong MDA5-mediated immune response, resulting in severe damage to the host cells. Thus, integration of new mobile elements would be expected to be subject to a strong negative selection pressure. However, these dsRNA structures are recognized and extensively edited by ADAR1, thereby avoiding an immune response. Thus, RNA editing allows host organisms to tolerate the integration of new mobile elements[73]. The rapid turnover of mobile elements in the genome contributes to genomic plasticity and accelerates evolution by introducing novel sequences, triggering recombination events and altering the function of insertion sites[77–79]. RNA editing, therefore, facilitates global genome evolution.

The extensive ADAR1 response to invasion of a genome by a new active mobile element not only promotes genome evolution but also presents an opportunity to introduce novel recoding sites (FIG. 4B). These new sites are created when the mobile element is incorporated into protein-coding sequences by exonization or when paired mobile elements create a duplex that acts as an editing enhancer or inducer. It has been shown recently that introducing inverted *Alu* repeats into a transcript can induce site-selective editing at novel recoding sites several hundred nucleotides away on the same transcript; conversely, deletion of inverted repeats can virtually abolish editing at nearby known recoding sites[80–82]. For example, editing of a recoding site in NEIL1 depends on a primate-specific *Alu*-derived inducer in an adjacent intron. Addition of this primate-specific *Alu* sequence induces editing in a mouse construct[83]. Furthermore, many of the most efficiently edited (>50% editing) recoding sites conserved across mammals are located in proximity to a nearby editing-inducing element[81]. Presumably, these long duplex RNAs promote editing by acting as ADAR recruitment elements. Thus, a pair of inverted mobile elements newly introduced near a coding exon could form a dsRNA structure that would enhance editing of a neighbouring pre-existing recoding site or even initiate recoding at a site that was not edited before insertion of the repetitive element[83] (FIG. 4Bd).

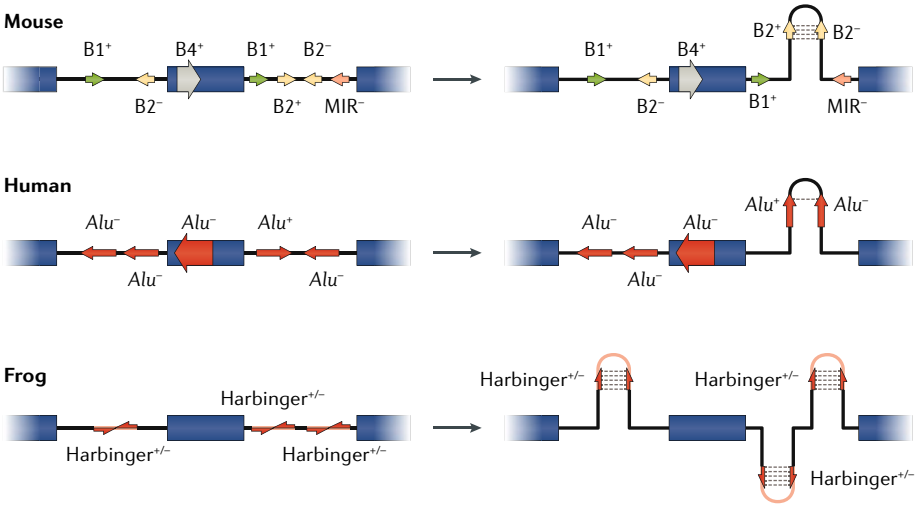
Alternatively, repetitive elements — and the editing sites within them — may be incorporated into coding sequences as novel exons[57,84]. Indeed, hundreds of *Alu* elements have been exonized into coding regions of the human transcriptome[85], and they are likely to have created novel recoding sites (FIG. 4Be). A notable example is the *NARF* gene. It contains a pair of inverted *Alu* repeats in one of its exons, which are extensively edited in *NARF* transcripts. In primates, editing of *NARF* pre-mRNA creates a novel splicing site and recodes a stop codon, resulting in a novel primate-specific alternatively spliced exon, which itself contains additional recoding sites[57] (FIG. 4Bf). If novel recoding sites introduced by exonization or insertion of an editing-inducing element result in a beneficial phenotype, they may be selected for and maintained in the genome.

Functional recoding in different lineages. The most probable ancestral function of ADAR1 is to protect against an innate immune response to self dsRNAs. It is therefore possible that all existing recoding sites originated as satellites of this ancestral ADAR1 activity through the mechanisms described above and that fixation of a few of these beneficial satellites allowed recoding, which in mammals is mediated mainly by a distinct protein, ADAR2, to evolve as a secondary ADAR

Exonization
Recruitment of a new exon from non-protein-coding intronic DNA, mostly from mobile elements.

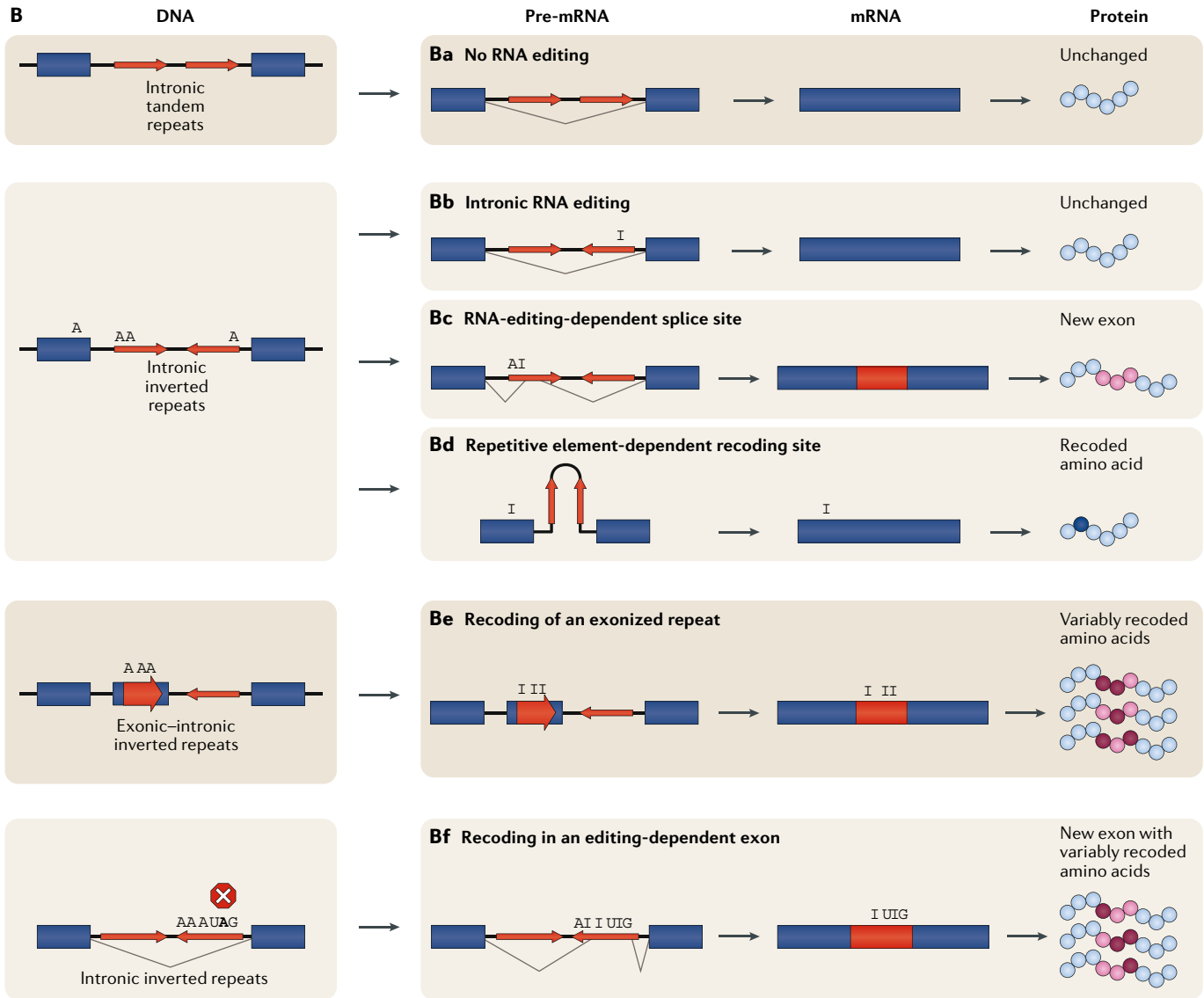
Fig. 4 | **Extent and consequences of editing in repetitive elements. A** | The extent of editing in repetitive elements strongly depends on the repeat repertoire. In most organisms, pairing of inverted repeats is a major source of endogenous double-stranded RNAs (dsRNAs). However, the number of these structures and the scope of editing within them vary considerably across animals, with frogs showing an order of magnitude more editing than mice[73]. This variability could be traced back to the properties of the repeats. In mice, there are many families of abundant but diverse repeats, including short interspersed nuclear elements (SINEs), and the typical length of a repeat is rather short, roughly 100 bp. Thus, pairing of similar, reversely oriented repeats of the same family is less likely, and the resulting structures are less stable[75]. By contrast, there is only one dominant repeat family in the human genome, the *Alu* SINE. Low divergence between *Alu* copies and the 300-bp length result in more abundant and more stable editing targets. Finally, the frog Harbinger repeat has a palindromic structure, enabling it to fold to a dsRNA structure even in the absence of a neighbouring repeat, leading to an even stronger signal[73,175]. Data in the table are adapted from REF.[73]. **B** | Double-stranded RNA-specific adenosine deaminase (ADAR) activity in genomic repeats may introduce novel proteins: **Ba-Bb** | Editing of a dsRNA structure formed by repetitive elements residing within an intron may have no effect on the resulting protein. **Bc** | However, in some cases, it may create an editing-dependent splicing site. For example, in the appropriate sequence context, the genomic dinucleotide AA edited to AI may be interpreted by the splicing machinery as an acceptor splice site, resulting in a novel exon being introduced into the protein. **Bd** | Alternatively, long dsRNA structure within the intron may serve as an editing enhancer, promoting recoding in the neighbouring exons. **Be** | In addition, retroposed mobile elements are often exonized, and, if an inverted repeat is present in the adjacent intron, the exonized repeat is likely to include multiple editing sites that may become recoding sites. **Bf** | Finally, an exonized repeat may contain a premature stop codon, editing of which may be necessary to prevent nonsense-mediated decay of the transcript and enable its translation, as is the case with the human *NARF* gene[57].

A



Main edited repeats family	Editing sites found in one brain sample	Sites that overlap with repeats (% of total)
B1 and B2	13,748	9,497 (69)
Alu	75,984	70,160 (92)
Harbinger	147,172	69,943 (48)

B



Synonymous substitutions
Replacement of one base by another within a coding region of a gene, which does not result in an amino acid change in the protein sequence.

Exapted
A trait, a gene or a cellular process that has changed function during evolution.

Adaptation
The evolutionary process by which the genetic information carried by a population of organisms is adjusted to improve their fitness to the environment.

function. If this is the case, it would be expected that at a sufficiently short evolutionary timescale after the creation of the new recoding sites, most would be evolutionarily neutral or possibly mildly deleterious: only a tiny fraction would be expected to give rise to beneficial phenotypes and be subject to positive selection. Indeed, a bioinformatic analysis of human recoding sites supports this idea[56]. Only a few dozen coding sites in humans are conserved across mammalian species[55], and these highly conserved sites are believed to be functional. By contrast, the thousands of human-specific editing sites in protein-coding sequences are generally non-adaptive and slightly deleterious[56]. Nevertheless, a small fraction of these species-specific sites may be beneficial. Indeed, because of the ever-changing landscape of *Alu* elements in primate genomes, there are likely to be numerous functional recoding sites that are specific to primates and possibly specific to humans[86]. Given the extremely high similarity between the human genome and the chimpanzee genome, identification and characterization of even a few human-specific functional sites could be an important contribution to our understanding of primate evolution. However, reliable identification of such sites is currently an unmet challenge.

In contrast to humans, hundreds of editing sites in coding sequence have been identified that are conserved across *Drosophila* spp. spanning ~60 million years of evolution[32–34], and more than 10,000 are conserved across cephalopod species that diverged from each other ~120–220 Myr ago[37]. These sites show signs of positive selection, including enrichment for non-synonymous substitutions over synonymous substitutions. Interestingly, virtually all recoding sites in mammals, *Drosophila* and cephalopods are lineage-specific, with only a single target (the Shaker potassium channel) known to be shared by all[29]. This observation is consistent with the view that recoding sites were not part of the ancestral set of ADAR targets but rather were exapted into the genomes of the different lineages subsequent to their divergence, possibly following a lineage-specific large-scale genome invasion of mobile elements. If this is the case, screening of more lineages is expected to reveal completely independent sets of recoding sites of widely varying size.

The adaptive potential of recoding. Importantly, recoding has the capacity to create a range of proteins from a single DNA sequence, which provides the organism with a new means for adaptation and the rapid development of new traits. Unlike genomic mutations, RNA edits are not directly transmitted to the next generation of cells. However, editing relies on the target RNA adopting a specific sequence-dependent secondary structure and possibly adjacent editing-enhancing dsRNA structures, which are determined by the genomic sequence surrounding the editing sites[52,83,87,88]. Genomic mutations within these regions can create novel editing sites or fine-tune the editing levels of existing ones, and these genomically encoded changes are heritable.

RNA editing creates a much greater level of transcript (and therefore proteome) complexity than do genomic mutations. A genomic mutation will be present in all pre-mRNAs transcribed from that allele, in all cells of the organism and under all conditions. By contrast, not every recoding site in a given pre-mRNA molecule will be edited, and the balance between edited and unedited transcript versions, as well as the repertoire of different combinations of editing in neighbouring sites, is often dependent on the tissue[47], the developmental stage[89] or external conditions (FIG. 5). Thus, recoding facilitates a much wider range of possibilities for adjusting the transcriptome than do genomic mutations. RNA editing is therefore a mechanism for heritable proteome diversification and has the potential to lead to adaptation in response to external pressures[90].

One interesting example of how RNA editing might contribute to adaptation occurs in corals. ADAR and editing levels in corals peak during spawning at the time of gamete release, resulting in over a thousand recoding events[74]. Without affecting the genome, RNA editing generates an increased tissue-specific and condition-specific post-transcriptional diversity in the numerous gametes being released, which can increase the probability that at least some gametes will survive the dynamic and varied environmental conditions they encounter while in the early developmental stages (FIG. 6a). However, despite the advantages of this seemingly large potential for adaptation, recoding is believed to be rather limited in the animal kingdom, except for in cephalopods, as mentioned above. This observation suggests that there are underlying mechanisms that prevent recoding from becoming as widespread as other means of generating transcriptome and proteome diversity, such as alternative splicing.

Constraints on recoding as a means of adaptation. It has recently been suggested[37] that maintaining a fixed recoding site within the genome imposes an evolutionary constraint on the genomic region that encodes the dsRNA structure recognized by ADAR proteins. Mutations that affect the stability of this secondary structure might modify the level of editing or abolish editing altogether[87,91]. Such mutations will undergo purifying selection so that the delicate balance between the edited and unedited versions of the protein is maintained. If numerous functional recoding sites exist throughout the genome, this constraint would result in a global slow-down of proteome evolution. In cephalopods, it was estimated that 3–15% of the interspecies mutations and 10–26% of the intraspecies polymorphisms were purified owing to constraints associated with maintenance of editing[37]. Conversely, creation of a new editing site requires a structure to evolve, imposing evolutionary constraints on the surrounding sequence. This trade-off between the transcriptome plasticity provided by RNA editing and the genomic variation required to drive adaptation and evolution might explain why extensive recoding is disfavoured in most Metazoan lineages.

Another possible explanation for recoding not being used more widely is related to its regulation. As far as is known, editing levels are mainly regulated by two factors: first, the local RNA sequence and structural motifs surrounding each specific editing site; and second, the expression level of ADAR proteins and their regulators. The surrounding sequence is hard wired

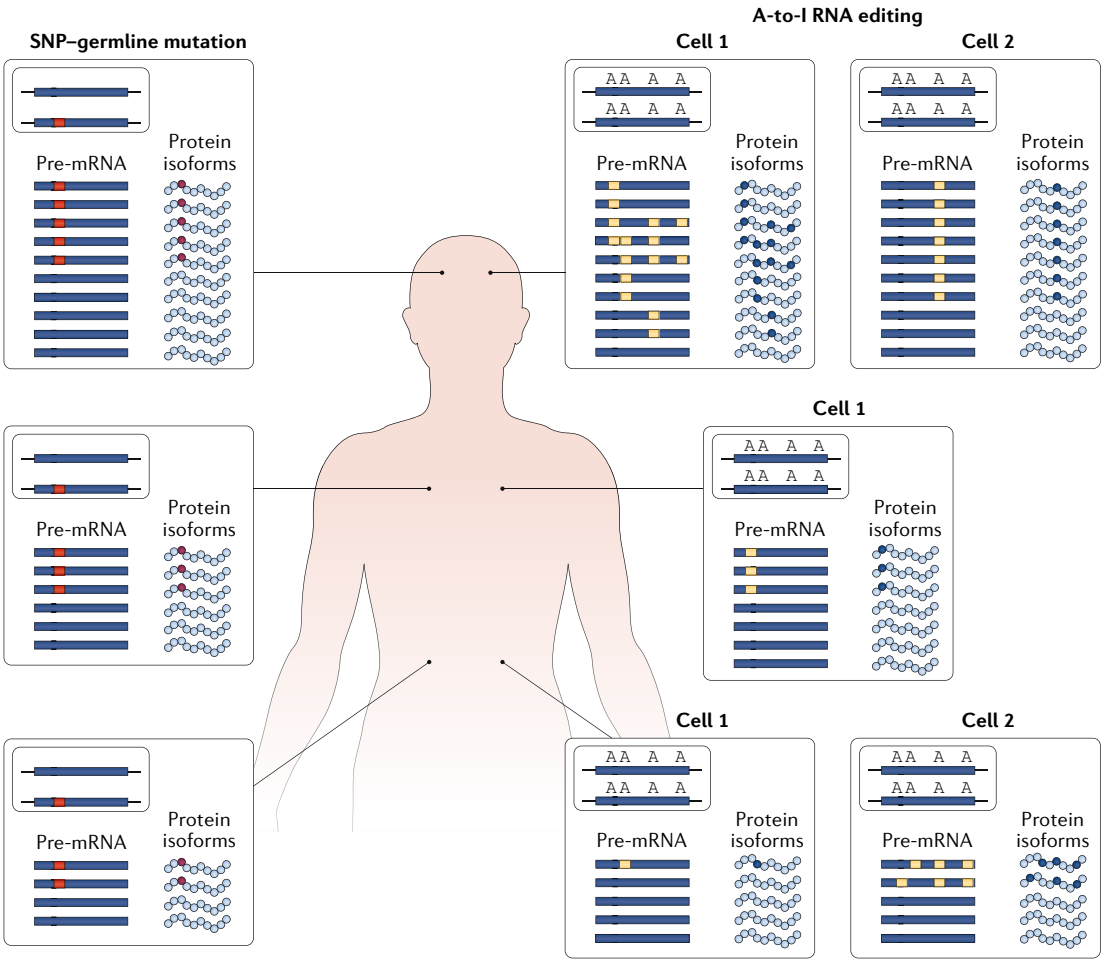


Fig. 5 | **RNA editing generates transcriptomic diversity.** Unlike genomic mutations, which usually affect either half or all the transcripts expressed from a gene, editing levels vary between 0 and 100%. Moreover, RNA edits, which can be viewed as RNA mutations, may lead to a different balance of edited and non-edited protein isoforms in different cells or tissues or under different conditions. Newly introduced edits can therefore have a more subtle effect on phenotype than do genomic mutations. Furthermore, while genomic variability introduced by a mutation is typically limited to two isoforms (haplotypes), multiple editing events in a given transcript can lead to combinatorial diversity. SNP, single nucleotide polymorphism.

because it is encoded in the genome, and its effects are the same regardless of the cell type or tissue type, developmental stage or external conditions. Indeed, editing levels at specific mammalian sites are largely consistent across tissue-matched samples from different individuals[92]. Such hard-wired regulation has little or no added adaptive value over genomic mutations. By contrast, changes in ADAR levels might allow intricate tissue-dependent or condition-dependent regulation[93], but all editing sites would be affected (FIG. 6b). It is possible that there are *trans*-regulators of RNA editing that would allow for more complex regulation. For example, ADAR3 serves as an editing inhibitor, probably by competing for dsRNA substrates[47,94]. NEDD4-like E3 ubiquitin-protein ligase WWP2 (REF.[95]) and aminoacyl tRNA synthase complex-interacting multifunctional protein 2 (AIMP2) (REF.[47]) also act as negative regulators of ADAR activity, while peptidyl-prolyl cis-trans isomerase NIMA-interacting 1 (PIN1) (REF.[95]), importin-α4 (REF.[96]), 26S proteasome complex subunit SEM1 (REF.[97]) and heterogeneous nuclear ribonucleoproteins A2/B1 (HNRNP A2/B1) (REF.[97]) are positive regulators. However, these ADAR regulators mainly affect editing globally and probably do not allow for site-specific control of editing levels. If indeed editing regulation, by and large, does not provide site-specific resolution, this sets a major limitation on the use of recoding for adaptation and acclimation. The higher the number of functional recoding sites, the more constraints there will be on adjusting global regulators of recoding.

Nevertheless, global regulation of recoding may still be useful for adaptation in cases where all sites, or many of them, are affected equally by a change in external conditions, such as temperature or acidity. Indeed, editing has been shown to be involved in temperature responses in both *Drosophila* and cephalopods[98–100]. Presumably, a decrease in the external temperature perturbs the energy–entropy balance controlling protein-folding and might be mitigated by a global increase in editing, which tends to recode to smaller, less stabilizing amino acids[101]. Similarly, changes in RNA editing have been associated with

Acclimation
The process by which an individual organism adjusts to a short-term change in its environment (as opposed to genomic changes on evolutionary timescales, called adaptation).

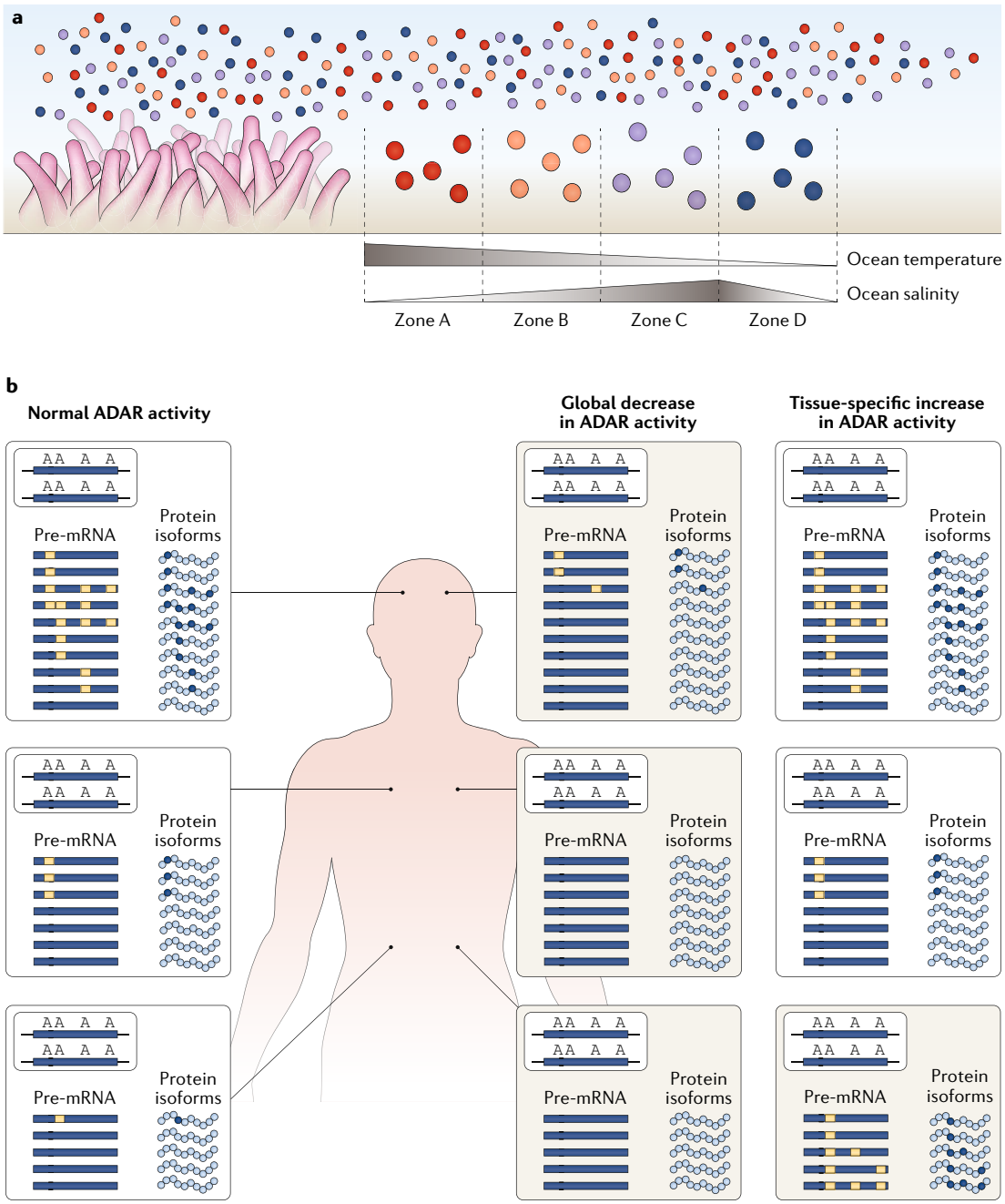


Fig. 6 | **Capacity and limitations of RNA editing as a means for adaptation. a** | Extensive recoding, such as that observed in coral gametes during spawning, can increase diversity and adaptability without manipulating the underlying genome. Numerous gametes (represented by small circles) are produced, with many different patterns of recoding (denoted by different colours). The proteomic diversity in this population may contribute to the probability that at least some gametes will overcome the challenges they encounter in their early developmental stages (represented by large coloured circles) when they face dynamic and varied environmental conditions (represented by zones A–D), such as ocean temperature and salinity. These proteomic changes are all temporal and are not present in surviving adults. **b** | Dynamic regulation of editing is achieved mainly by changes in double-stranded RNA-specific adenosine deaminase (ADAR) levels. Altered ADAR expression may lead to global changes in the editing levels, or even tissue-specific control, but do not allow for editing regulation at single-site resolution. If indeed editing regulation does not allow for site-specific control of the editing level, this may constrain the number of functional recoding sites.

other changes in organism-level conditions, such as sleep[102] and circadian changes[103]. However, it is not yet clear to what extent the large numbers of conserved recoding sites in *Drosophila* and cephalopods play a part in adaptation[33,104,105].

Altered RNA editing as RNA mutations

Just as RNA editing events might act like genomic mutations to drive adaptation, deregulated RNA editing might have an effect similar to that of disease-related genomic mutations. Altered editing could manifest as

changes in the level of recoding at known sites; creation of novel disease-specific recoding events; creation of novel editing-assisted splicing events; or modified editing patterns of circRNAs, miRNAs and their targets. These aberrant editing events could be viewed as a new class of non-heritable RNA mutation that can only be detected by sequencing the transcriptome. However, unlike genomic mutations, which are present in every cell and detectable by DNA sequencing (DNA-seq), the phenotypes associated with RNA mutations will be dependent on a fine balance between mutant and non-mutant transcripts, where mutant transcripts may be expressed (and detectable by RNA-seq) only under certain conditions, at a specific stage of a disease or in certain cells or tissues (FIG. 5).

RNA mutations and cancer. The heterogeneity introduced by editing may be exploited by tumour cells to promote cancer progression[106]. For instance, the DNA-editing cytidine deaminase APOBEC3B has recently been shown to increase the rate of tumour genome evolution by introducing large numbers of C-to-U edits into the DNA of cancer cells[107,108]. Compared with DNA editing, A-to-I RNA editing could provide additional variability that might depend on tumour stage and the genomic mutation background. For example, stage-dependent and condition-dependent editing could be utilized to confer resistance to treatment. Indeed, analysis of hundreds of cancer samples has shown that RNA editing levels in most tumour types are significantly higher than those in matched normal tissue[44,109,110]. The number of cancer-specific RNA editing events is much higher than the reported levels of APOBEC3B DNA editing and is comparable to the total number of somatic mutation events in cancer DNA[44]. These editing events are mostly concentrated in the non-coding *Alu* elements, but significant differences between editing levels in cancerous and non-cancerous states have also been demonstrated at dozens of specific recoding sites and miRNA editing sites[44,58,60,109,110].

As is the case with somatic genomic mutations, most newly introduced RNA mutations are likely to be passenger mutations, but a few may serve as driver mutations and represent novel candidates with therapeutic and diagnostic potential. A few examples of specific RNA mutations suggested to drive cancer have been recently studied. One occurs in the recoding site in *AZIN1* transcripts, whose editing is elevated in hepatocellular carcinoma. The edited version of AZIN1 neutralizes the antizyme-mediated degradation of ODC and CCND1, thus promoting cell proliferation and contributing to cancer progression[45]. Driver mutations are not limited to recoding sites. Hypo-editing of a site within miR-455-5p has been shown to enhance melanoma growth and metastasis, as the unedited version (but not the edited one) inhibits the tumour suppressor gene *CPEB1* (REF.[111]). Further evidence of the importance of editing in cancer comes from the observation that there is a significant negative correlation between the global A-to-I RNA editing level and patient survival[44,110]. More examples of driver RNA mutations are starting to emerge[112,113].

RNA mutations and autoimmune disorders. The recent appreciation of the connection between RNA editing and autoimmunity has led to an increased interest in the part editing plays in autoimmune disorders. Maintaining precise editing levels is essential to prevent autoimmune reactions. On one hand, ADAR1 editing is required to prevent endogenous dsRNA from being mistakenly identified as non-self by the immune system[63–65,114]. Accordingly, too little ADAR activity presents a risk of an undesired autoimmune response. In fact, mutations in ADAR1 cause Aicardi–Goutières syndrome[71], an autoimmune disorder. On the other hand, the proteomic diversity created by RNA editing also poses a challenge for the immune system because it introduces novel antigenic determinants. In order to meet this challenge, recoding activity is very high in the thymus, the guardian of central tolerance, where developing T cells are rendered non-reactive to self. This strong recoding is presumably required to train T cells to recognize the edited version as self[115]. Yet elevated or misdirected ADAR activity may result in recoding of additional sites that were not edited in the thymus, leading to unrecognized antigens that may stimulate an undesired autoimmune response. Taken together, these observations suggest that altered editing has a role in autoimmune disorders. We expect the potential link between editing and autoimmune disease to become an active field of study, including screening the editomes of individuals with autoimmune disorders.

RNA mutations and other human disorders. Global alterations in RNA editing might have an effect in pathologies other than cancer and autoimmune disorders. In the few recoding sites that have been extensively studied so far, aberrant editing has been correlated with a number of different human disorders, most of which are neurological or brain-related, including amyotrophic lateral sclerosis[116], epilepsy[117], depression-related suicide[118], schizophrenia and bipolar disorder[119]. Altered editing was also observed in fragile X models, presumably owing to the lack of a functional FMRP protein; the interaction of FMRP protein with ADAR has been shown to affect editing activity[120–122]. We expect that advances in editing screening technologies will help to reveal more associations between the pathological effects observed in a variety of disorders and RNA mutations, which are likely to be mainly caused by misregulation of ADAR or ADAR-interacting proteins. Recently developed computational and experimental approaches[123] are already being used to screen for altered editing in neurological and brain disorders, and initial results show that editing alterations occur in the central nervous system after injuries[124,125], Alzheimer disease[126] and autism[127].

Finally, arteries were recently shown to exhibit the highest editing levels in human[47]. Consistent with this observation, emerging large-scale expression data have revealed that ADAR2 is highly expressed in arteries[128]. A recent study has connected atherosclerosis with aberrant RNA editing of a specific editing site located in the 3′ UTR of the *CATHEPSIN S* mRNA. Editing at this site controls mRNA stability and expression by recruiting a stabilizing RNA-binding protein (RBP)[129]. These findings are likely to lead to an increased interest in the role of RNA editing in vascular diseases.

Somatic mutation
An alteration in DNA that occurs after conception. Somatic mutations are not shared by all cells of the body.

Passenger mutations
Mutations that are caused by genomic instability, which is common in cancer cells, but do not promote malignancy.

Driver mutations
Mutations that provide cancer cells with a selective advantage and promote malignancy.

Future perspectives

Recent advances in both sequencing technologies and computational methods for detecting editing have helped to reveal the full extent of the editome, but many questions remain to be answered and many challenges remain to be tackled if we are to gain a fuller understanding of RNA editing and its biological functions.

Identifying and characterizing important editing sites. Detection methods still perform rather poorly for isolated recoding sites (BOX 1), especially where interspecies conservation cannot be utilized. Examples of such sites include species-specific recoding sites and novel recoding sites that act as disease-associated RNA mutations. The latter are anticipated to become a major focus of research in the near future, particularly in cancer and inflammatory and/or autoimmune diseases. Thus, development of better approaches for identifying novel isolated recoding sites is an important open challenge. In addition, emerging computational and experimental methods for the global quantification of editing levels have opened the door to the systematic profiling of RNA editing changes in pathologies and are expected to be widely used for editome screening[44,108,109,117,124–127]. Technology holds the key to further improving the identification and quantification of editing events. For example, the nanopore sequencing platforms[130] could allow direct detection of inosines in full-length RNA molecules, which would overcome many technical issues that currently hinder editing detection[131]. Finally, another challenge will be to determine which ADAR1 target sites are located in dsRNAs that elicit an autoimmune response; these are likely to be the key functional targets of ADAR1, and their defining features are currently poorly understood.

Understanding the mechanisms and functions of editing. An increasing number of recoding sites that are conserved within lineages are being reported. These are likely adaptive; however, our understanding of the implications and functions of editing at these sites is outpaced by their rate of discovery, particularly for non-mammalian species. The phenotype associated with editing is often mild, potentially tissue-specific or condition-specific, and is therefore difficult to pinpoint and characterize[49]. Similarly, despite many recent advances, ADAR specificity and regulation are also only partially understood in terms of what determines which sites are edited, which additional proteins interact with ADAR and affect editing, and how individual recoding sites are regulated. A better understanding of these points would facilitate improved methods for editing detection and control of editing levels at a single-site resolution, as well as utilization of the editing machinery for RNA engineering.

Single-cell analysis of editome heterogeneity. Virtually all editome screens performed so far have provided tissue-averaged information. However, a number of recent works have used cell-type sorted samples to study the heterogeneity of RNA editing within tissues and indicated that editing levels could vary substantially between cell populations[124] and even between single cells of the same subtype[132,133]. These results highlight the potential of editing to contribute to intercellular diversity even within the same tissue. Importantly, sites that appear to be weakly edited when averaged across whole tissues may be strongly edited in some individual cells. Improvements in the achievable coverage of single-cell sequencing methods will allow a better understanding of the intercellular diversity created by editing.

Understanding the evolutionary history and implications of editing. Editing has been observed in all Metazoans that have been tested to date, but species vary considerably in terms of the general scope of editing and the degree to which it is used for functional recoding. Why this should be the case remains unclear and merits further study. As discussed above, we hypothesize that the introduction of new recoding sites is connected to an invasion of repetitive elements into the genome; additional investigations are required to determine if the data support this theory. The role of recoding in adaptation is particularly interesting and requires detailed analyses of the evolutionary route leading to fixation of a new recoding site, the functional impact of recoding on the target proteins and the relationship to external pressures. Another key question is whether human-specific or primate-specific adaptive recoding sites exist and if and how they have contributed to the evolution of our lineage. At the level of the individual organism, an important advantage of recoding over hard-wired genomic mutations is the capacity to respond to external conditions over short timescales; however, it has yet to be shown whether editing is actually used for acclimation. Finally, A-to-I editing was recently reported in species of bacteria and fungi lacking ADAR proteins[134–136]. It is of interest to find out how prevalent A-to-I editing is outside the animal kingdom, identify the responsible enzymes and understand their evolutionary relationship to the Metazoan ADARs. Revealing the mechanism and function of editing in these species might contribute to our understanding of the early evolution of the editing world.

Utilizing ADARs for RNA engineering. Recently, the capacity of ADAR enzymes to alter RNA sequences at specific positions has been harnessed to probe and manipulate RNA. For instance, in targets of RNA-binding proteins identified by editing (TRIBE), the deaminase domain of ADAR is fused to an RBP of interest in order to detect its RNA targets; when the ADAR–RBP fusion protein binds to an RNA molecule, it converts nearby adenosines to inosines, which can be identified using RNA-seq[137] (FIG. 7a). More excitingly, ADAR is now being used for RNA engineering (FIG. 7b). Using a specific guide RNA, ADAR enzymes are targeted to edit a specific adenosine of choice. Successful in vivo targeting of specific locations was demonstrated to correct genetic mutations associated with several diseases[138,139] or to control protein localization by targeting an editing switch introduced into the transcript of choice[140]. Currently, these approaches are not sensitive enough to rely on endogenous ADAR expression

Nanopore sequencing platforms
Emerging sequencing methods by which a single molecule of DNA or RNA can be sequenced without the need for PCR amplification.

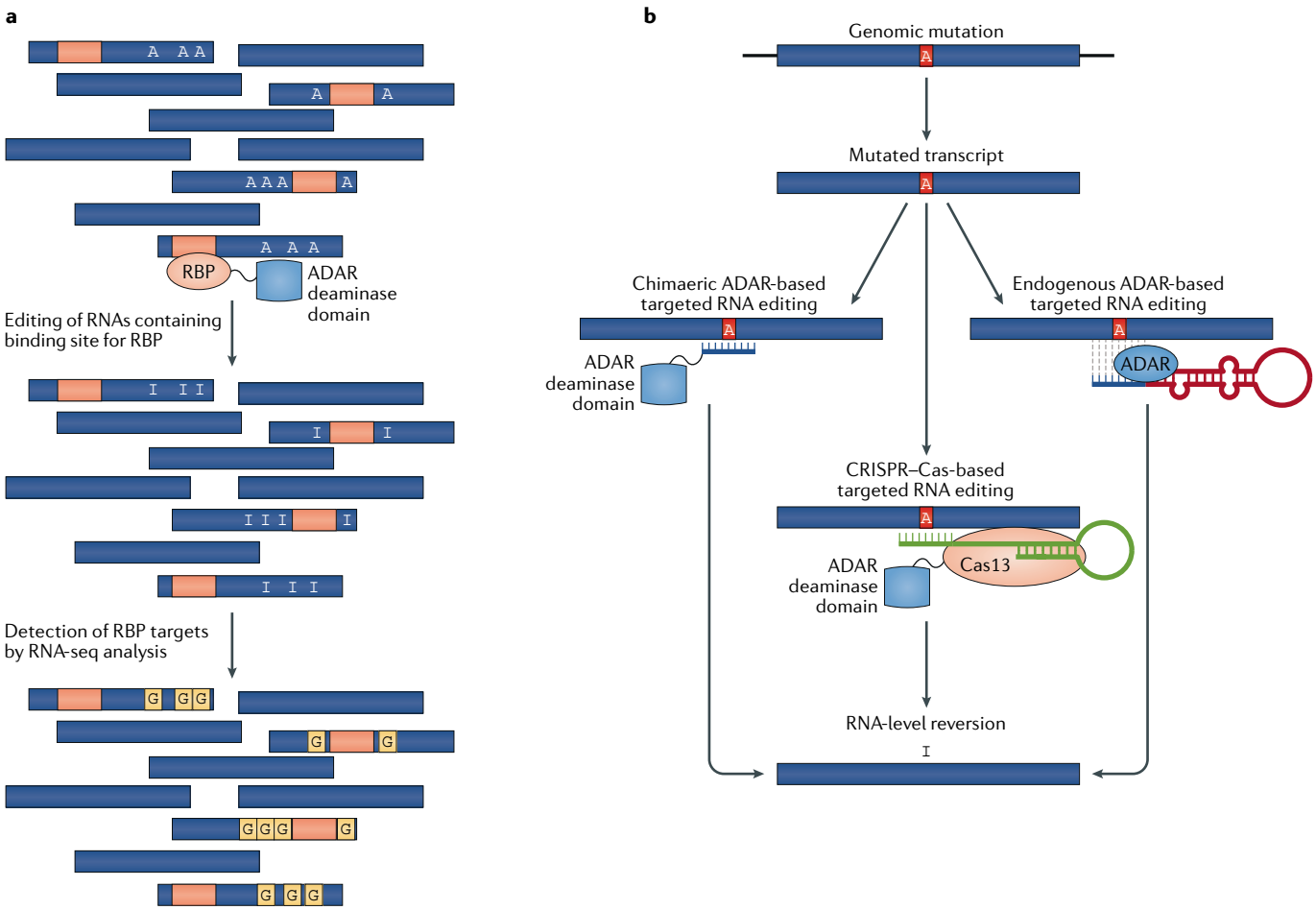


Fig. 7 | **Utilizing ADARs for RNA probing and engineering. a** | Targets of RNA-binding proteins identified by editing (TRIBE)[137] is a method that enables identification of the targets of an RNA-binding protein (RBP) of interest. It involves fusing the RBP of choice with the deaminase domain of double-stranded RNA (dsRNA)-specific adenosine deaminase (ADAR). The fusion protein leaves an inosine trace on the RNA targets of the RBP, which can be readily detected by sequencing the RNA and looking for A → G mismatches. **b** | ADARs can be directed to specific nucleotides within the transcriptome to manipulate the RNA sequence without affecting the genome. Three main approaches for ADAR-based RNA engineering have been considered. The first approach (left) uses an engineered version of ADAR that consists of the deaminase domain fused to an antisense RNA oligonucleotide. The antisense RNA guides the complex to its target and serves as a template for the dsRNA structure required for ADAR function[139,180]. An alternative approach (right) delivers only the guide RNA, which forms a dsRNA that recruits endogenous ADARs. The dsRNA structure was designed to mimic that of the mammalian Q/R site of the glutamate receptor 2 (GRIA2), the classic fully edited ADAR2 target[138,181]. This method might have an advantage in terms of specificity, avoiding promiscuous unintended edits that currently accompany ADAR transfections. However, so far, the expression levels of the endogenous ADAR enzyme are insufficient for effective editing and ectopic expression of ADAR2 is needed. A third approach (middle) increased the efficacy of targeted editing by engineering a catalytically inactive Cas13 protein fused with the deaminase domain of ADAR2 (REF.[182]). RNA-seq, RNA sequencing.

and suffer from extensive off-target editing[141]. Future progress in understanding the rules determining the selectivity of ADARs may make these approaches more attractive and possibly even allow them to be applied in clinical settings.

Conclusions

Three decades ago, ADARs were first reported as enzymes that unwind dsRNA structures[142–144]. Shortly thereafter, their capacity to recode mRNA to create novel protein isoforms was discovered[30], and the handful of identified recoding sites became the main focus of research. Now it is clear that recoding, though important, represents only a tiny fraction of ADAR target RNAs. Furthermore, the recent observation that deletion of *MDA5* rescues the lethal phenotype of mice lacking ADAR1 function[70] has challenged the notion that recoding is the major biological function of ADARs. Instead, it now seems that the primordial function of the widely expressed ADAR1 protein is to unwind endogenous dsRNA structures that would otherwise elicit an innate immune response[63–65] — the very function that was first assigned to the enzyme in the early pioneering works. Recoding events are merely a side effect of this RNA processing activity and are mostly non-adaptive and non-functional[56,70]. New computational methods and the availability of much larger data sets have helped to show that adaptive, conserved recoding occurs in mammals

at only a few dozen sites. However, recent studies in *Drosophila* and cephalopods suggest that mammals are an exceptional lineage in terms of their limited use of recoding.

Over the next 3 decades, we expect to see progress in identifying the important editing sites (recoding and non-coding alike) and an improved understanding of their function and contribution to adaptation across the tree of life. Systematic screening and standardized quantification of the editome will reveal the role played by editing in health and disease. Finally, utilizing the editing machinery for RNA engineering will enable reprogramming of the transcriptome content with a temporal and spatial resolution, with numerous exciting biotechnological and clinical applications.

Published online: 24 April 2018

1. Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell* **126**, 37–47 (2006).
2. Di Giammartino, D. C., Nishida, K. & Manley, J. L. Mechanisms and consequences of alternative polyadenylation. *Mol. Cell* **43**, 853–866 (2011).
3. Boccaletto, P. et al. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* **46**, D303–D307 (2018).
4. Peer, E., Rechavi, G. & Dominissini, D. Epitranscriptomics: regulation of mRNA metabolism through modifications. *Curr. Opin. Chem. Biol.* **41**, 93–98 (2017).
5. Safra, M. et al. The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* **551**, 251–255 (2017).
6. Benne, R. et al. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**, 819–826 (1986).
7. Takenaka, M. et al. RNA editing in plant mitochondria —connecting RNA target sequences and acting proteins. *Mitochondrion* **19**, 191–197 (2014).
8. Wedekind, J. E., Dance, G. S. C., Sowden, M. P. & Smith, H. C. Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet.* **19**, 207–216 (2003).
9. Bass, B. L. RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **71**, 817–846 (2002).
10. Savva, Y.a, Rieder, L. E. & Reenan, R. A. The ADAR protein family. *Genome Biol.* **13**, 252 (2012).
11. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **79**, 321–349 (2010).
12. Lonsdale, J. et al. The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
13. Thomas, J. M. & Beal, P. A. How do ADARs bind RNA? New protein-RNA structures illuminate substrate recognition by the RNA editing ADARs. *BioEssays* **39**, 1600187 (2017).
14. Kleinberger, Y. & Eisenberg, E. Large-scale analysis of structural, sequence and thermodynamic characteristics of A-to-I RNA editing sites in human Alu repeats. *BMC Genomics* **11**, 453 (2010).
15. Eggington, J. M., Greene, T. & Bass, B. L. Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* **2**, 319 (2011).
16. Levanon, E. Y. et al. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* **22**, 1001–1005 (2004).
17. Athanasiadis, A., Rich, A. & Maas, S. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol.* **2**, e391 (2004).
18. Blow, M., Futreal, A. P., Wooster, R. & Stratton, M. R. A survey of RNA editing in human brain. *Genome Res.* **14**, 2379–2387 (2004).
19. Kim, D. D. Y. et al. Widespread RNA editing of embedded Alu elements in the human transcriptome. *Genome Res.* **14**, 1719–1725 (2004).
20. Ramaswami, G. & Li, J. B. Identification of human RNA editing sites: a historical perspective. *Methods* **107**, 42–47 (2016).
This paper reviews the development of computational approaches for detecting RNA editing sites.
21. Ramaswami, G. & Li, J. B. RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res.* **42**, D109–D113 (2014).
22. Picardi, E., D'Erchia, A. M., Lo Giudice, C. & Pesole, G. REDIportal: a comprehensive database of A-to-I RNA editing events in humans. *Nucleic Acids Res.* **45**, D750–D757 (2017).
23. Rosenthal, J. J. C. & Seeburg, P. H. A-To-I RNA editing: effects on proteins key to neural excitability. *Neuron* **74**, 432–439 (2012).
24. Tomaselli, S., Locatelli, F. & Gallo, A. The RNA editing enzymes ADARs: mechanism of action and human disease. *Cell Tissue Res.* **356**, 527–532 (2014).
25. Rosenthal, J. J. C. The emerging role of RNA editing in plasticity. *J. Exp. Biol.* **218**, 1812–1821 (2015).
This review discusses the contribution of recoding events to proteome diversity.
26. Nishikura, K. A-To-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell. Biol.* **17**, 83–96 (2016).
27. Gallo, A., Vukic, D., Michalík, D., O'Connell, M. A. & Keegan, L. P. ADAR RNA editing in human disease; more to it than meets the I. *Hum. Genet.* **136**, 1265–1278 (2017).
This article summarizes the link between altered RNA editing and human pathologies.
28. Basilio, C., Wahba, A. J., Lengyel, P., Speyer, J. F. & Ochoa, S. Synthetic polynucleotides and the amino acid code, V. *Proc. Natl Acad. Sci. USA* **48**, 613–616 (1962).
29. Hoopengardner, B., Bhalla, T., Staber, C. & Reenan, R. Nervous system targets of RNA editing identified by comparative genomics. *Science* **301**, 832–836 (2003).
30. Sommer, B., Kohler, M., Sprengel, R. & Seeburg, P. H. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* **67**, 11–19 (1991).
31. Burns, C. M. et al. Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* **387**, 303–308 (1997).
32. Zhang, R., Deng, P., Jacobson, D. & Li, J. B. Evolutionary analysis reveals regulatory and functional landscape of coding and non-coding RNA editing. *PLoS Genet.* **13**, e1006563 (2017).
33. Duan, Y., Dou, S., Luo, S., Zhang, H. & Lu, J. Adaptation of A-to-I RNA editing in *Drosophila*. *PLoS Genet.* **13**, e1006648 (2017).
34. Yu, Y. et al. The landscape of A-to-I RNA editome is shaped by both positive and purifying selection. *PLoS Genet.* **12**, e1006191 (2016).
35. Graveley, B. R. et al. The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**, 473–479 (2011).
36. St Laurent, G. et al. Genome-wide analysis of A-to-I RNA editing by single-molecule sequencing in *Drosophila*. *Nat. Struct. Mol. Biol.* **20**, 1333–1339 (2013).
37. Liscovitch-Brauer, N. et al. Trade-off between transcriptome plasticity and genome evolution in cephalopods. *Cell* **169**, 191–202.e11 (2017).
This study demonstrates that recoding in cephalopods is extensive and is maintained by evolution.
38. Alon, S. et al. The majority of transcripts in the squid nervous system are extensively recoded by A-to-I RNA editing. *eLife* **4**, e05198 (2015).
39. Ryan, M. Y., Maloney, R., Reenan, R. & Horn, R. Characterization of five RNA editing sites in shab potassium channels. *Channels* **2**, 202–209.
40. Ingleby, L., Maloney, R., Jepson, J., Horn, R. & Reenan, R. Regulated RNA editing and functional epistasis in *shaker* potassium channels. *J. Gen. Physiol.* **133**, 17–27 (2009).
41. Levanon, E. Y. et al. Evolutionarily conserved human targets of adenosine to inosine RNA editing. *Nucleic Acids Res.* **33**, 1162–1168 (2005).
42. Li, J. B. et al. Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. *Science* **324**, 1210–1213 (2009).
43. Stulić, M. & Jantsch, M. F. Spatio-temporal profiling olamin A RNA-editing reveals ADAR preferences and high editing levels outside neuronal tissues. *RNA Biol.* **10**, 1611–1617 (2013).
44. Paz-Yaacov, N. et al. Elevated RNA editing activity is a major contributor to transcriptomic diversity in tumors. *Cell Rep.* **13**, 267–276 (2015).
45. Chen, L. et al. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. *Nat. Med.* **19**, 209–216 (2013).
46. Yeo, J., Goodman, R.a, Schirle, N. T., David, S. S. & Beal, P. A. RNA editing changes the lesion specificity for the DNA repair enzyme NEIL1. *Proc. Natl Acad. Sci. USA* **107**, 20715–20719 (2010).
47. Tan, M. H. et al. Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249–254 (2017).
This article describes a comprehensive screen of RNA editing across a large number of human and mouse samples.
48. Higuchi, M. et al. Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78–81 (2000).
49. Horsch, M. et al. Requirement of the RNA-editing enzyme ADAR2 for normal physiology in mice. *J. Biol. Chem.* **286**, 18614–18622 (2011).
50. Bhalla, T., Rosenthal, J. J. C., Holmgren, M. & Reenan, R. Control of human potassium channel inactivation by editing of a small mRNA hairpin. *Nat. Struct. Mol. Biol.* **11**, 950–956 (2004).
51. Daniel, C., Wahlstedt, H., Ohlson, J., Björk, P. & Ohman, M. Adenosine-to-inosine RNA editing affects trafficking of the gamma-aminobutyric acid type A (GABA(A)) receptor. *J. Biol. Chem.* **286**, 2031–2040 (2011).
52. Lomeli, H. et al. Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science* **266**, 1709–1713 (1994).
53. Egebjerg, J. & Heinemann, S. F. Ca^{2+} permeability of unedited and edited versions of the kainate selective glutamate receptor GluR6. *Proc. Natl Acad. Sci. USA* **90**, 755–759 (1993).
54. Sailer, A. et al. Generation and analysis of GluR5(Q636R) kainate receptor mutant mice. *J. Neurosci.* **19**, 8757–8764 (1999).
55. Pinto, Y., Cohen, H. Y. & Levanon, E. Y. Mammalian conserved ADAR targets comprise only a small fragment of the human editosome. *Genome Biol.* **15**, R5 (2014).
56. Xu, G. & Zhang, J. Human coding RNA editing is generally nonadaptive. *Proc. Natl Acad. Sci. USA* **111**, 3769–3774 (2014).
This article discusses the general evolutionary role of recoding in humans.
57. Lev-Maor, G. et al. RNA-editing-mediated exon evolution. *Genome Biol.* **8**, R29 (2007).
58. Pinto, Y., Buchumenski, I., Levanon, E. Y. & Eisenberg, E. Human cancer tissues exhibit reduced A-to-I editing of miRNAs coupled with elevated editing of their targets. *Nucleic Acids Res.* **46**, 71–82.
59. Kawahara, Y. et al. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**, 1137–1140 (2007).
60. Wang, Y. et al. Systematic characterization of A-to-I RNA editing hotspots in microRNAs across human cancers. *Genome Res.* **27**, 1112–1125 (2017).
61. Vesely, C., Tauber, S., Sedlazeck, F. J., von Haeseler, A. & Jantsch, M. F. Adenosine deaminases that act on RNA induce reproducible changes in abundance and sequence of embryonic miRNAs. *Genome Res.* **22**, 1468–1476 (2012).
62. Ivanov, A. et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* **10**, 170–177 (2014).
63. Liddicoat, B. J. et al. RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* **349**, 1115–1120 (2015).
64. Pestal, K. et al. Isoforms of RNA-editing enzyme ADAR1 independently control nucleic acid sensor MDA5-driven autoimmunity and multi-organ development. *Immunity* **43**, 933–944 (2015).

65. Mannion, N. M. et al. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep.* **9**, 1482–1494 (2014).
References 63, 64 and 65 provide compelling evidence that the main function of ADAR1 editing is to inhibit self-activation of innate immunity by endogenous dsRNA.
66. Weber, F., Wagner, V., Rasmussen, S. B., Hartmann, R. & Paludan, S. R. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J. Virol.* **80**, 5059–5064 (2006).
67. Feng, Q. et al. MDA5 detects the double-stranded RNA replicative form in picornavirus-infected cells. *Cell Rep.* **2**, 1187–1196 (2012).
68. Chung, H. et al. Human ADAR1 prevents endogenous RNA from triggering translational shutdown. *Cell* **172**, 811–824.e14 (2018).
69. Ahmad, S. et al. Breaching self-tolerance to alu duplex RNA underlies MDA5-mediated inflammation. *Cell* **172**, 797–810.e13 (2018).
70. Heraud-Farlow, J. E. et al. Protein recoding by ADAR1-mediated RNA editing is not essential for normal development and homeostasis. *Genome Biol.* **18**, 166 (2017).
71. Rice, G. I. et al. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. *Nat. Genet.* **44**, 1243–1248 (2012).
72. Bazak, L. et al. A-To-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.* **24**, 365–376 (2014).
This study reveals the scope of editing in repetitive elements.
73. Porath, H. T., Knisbacher, B. A., Eisenberg, E. & Levanon, E. Y. Massive A-to-I RNA editing is common across the metazoa and correlates with dsRNA abundancee. *Genome Biol.* **18**, 185 (2017).
74. Porath, H. T. et al. A-To-I RNA editing in the earliest-diverging eumetazoan phyla. *Mol. Biol. Evol.* **34**, 1890–1901 (2017).
75. Neeman, Y., Levanon, E. Y., Jantsch, M. F. & Eisenberg, E. RNA editing level in the mouse is determined by the genomic repeat repertoire. *RNA* **12**, 1802–1809 (2006).
76. Bazak, L., Levanon, E. Y. & Eisenberg, E. Genome-wide analysis of Alu editability. *Nucleic Acids Res.* **42**, 6876–6884 (2014).
77. Feschotte, C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* **9**, 397–405 (2008).
78. Deininger, P. L., Moran, J. V., Batzer, M. A. & Kazazian, H. H. Mobile elements and mammalian genome evolution. *Curr. Opin. Genet. Dev.* **13**, 651–658 (2003).
79. Lowe, C. B., Bejerano, G. & Haussler, D. Thousands of human mobile element fragments undergo strong purifying selection near developmental genes. *Proc. Natl Acad. Sci. USA* **104**, 8005–8010 (2007).
80. Ramaswami, G. et al. Genetic mapping uncovers *cis*-regulatory landscape of RNA editing. *Nat. Commun.* **6**, 8194 (2015).
81. Daniel, C., Widmark, A., Rigardt, D. & Öhman, M. Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome. *Genome Biol.* **18**, 195 (2017).
82. Daniel, C., Venø, M. T., Ekdahl, Y., Kjems, J. & Öhman, M. A distant *cis* acting intronic element induces site-selective RNA editing. *Nucleic Acids Res.* **40**, 9876–9886 (2012).
83. Daniel, C., Silberberg, G., Behm, M. & Ohman, M. Alu elements shape the primate transcriptome by *cis*-regulation of RNA editing. *Genome Biol.* **15**, R28 (2014).
84. Möller-Krull, M., Zemann, A., Roos, C., Brosius, J. & Schmitz, J. Beyond DNA: RNA editing and steps toward alu exonization in primates. *J. Mol. Biol.* **382**, 601–609 (2008).
85. Dagan, T., Sorek, R., Sharon, E., Ast, G. & Graur, D. AluGene: a database of Alu elements incorporated within protein-coding genes. *Nucleic Acids Res.* **32**, D489–D492 (2004).
86. Paz-Yaacov, N. et al. Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. *Proc. Natl Acad. Sci. USA* **107**, 12174–12179 (2010).
87. Rieder, L. E., Staber, C. J., Hoopengardner, B. & Reenan, R. A. Tertiary structural elements determine the extent and specificity of messenger RNA editing. *Nat. Commun.* **4**, 2232 (2013).
88. Sapiro, A. L., Deng, P., Zhang, R. & Li, J. B. Cis regulatory effects on A-to-I RNA editing in related *Drosophila* species. *Cell Rep.* **11**, 697–703 (2015).
89. Wahlstedt, H. et al. Large-scale mRNA sequencing determines global regulation of RNA editing during brain development. *Genome Res.* **19**, 978–986 (2009).
90. Gommans, W. M., Mullen, S. P. & Maas, S. RNA editing: a driving force for adaptive evolution? *BioEssays* **31**, 1137–1145 (2009).
91. Reenan, R. A. Molecular determinants and guided evolution of species-specific RNA editing. *Nature* **434**, 409–413 (2005).
92. Greenberger, S. et al. Consistent levels of A-to-I RNA editing across individuals in coding sequences and non-conserved Alu repeats. *BMC Genomics* **11**, 608 (2010).
93. Picardi, E. et al. Profiling RNA editing in human tissues: towards the inosinome atlas. *Sci. Rep.* **5**, 14941 (2015).
94. Oakes, E., Anderson, A., Cohen-Gadol, A. & Hundley, H. A. Adenosine deaminase that acts on RNA 3 (ADAR3) binding to glutamate receptor subunit B pre-mRNA inhibits RNA editing in glioblastoma. *J. Biol. Chem.* **292**, 4326–4335 (2017).
95. Marcucci, R. et al. Pin1 and WWP2 regulate *GluR2* Q/R site RNA editing by ADAR2 with opposing effects. *EMBO J.* **30**, 4211–4222 (2011).
96. Behm, M., Wahlstedt, H., Widmark, A., Eriksson, M. & Öhman, M. Accumulation of nuclear ADAR2 regulates adenosine-to-inosine RNA editing during neuronal development. *J. Cell Sci.* **130**, 745–753 (2017).
97. Garncarz, W., Tariq, A., Handl, C., Pusch, O. & Jantsch, M. F. A high-throughput screen to identify enhancers of ADAR-mediated RNA-editing. *RNA Biol.* **10**, 192–204 (2013).
98. Garrett, S. & Rosenthal, J. J. C. RNA editing underlies temperature adaptation in K^+ channels from polar octopuses. *Science* **335**, 848–851 (2012).
99. Rieder, L. E. et al. Dynamic response of RNA editing to temperature in *Drosophila*. *BMC Biol.* **13**, 1 (2015).
100. Buchumenski, I. et al. Dynamic hyper-editing underlies temperature adaptation in *Drosophila*. *PLoS Genet.* **13**, e1006931 (2017).
101. Garrett, S. C. & Rosenthal, J. J. C. A role for A-to-I RNA editing in temperature adaptation. *Physiology* **27**, 362–369 (2012).
102. Robinson, J. E., Paluch, J., Dickman, D. K. & Joiner, W. J. ADAR-mediated RNA editing suppresses sleep by acting as a brake on glutamatergic synaptic plasticity. *Nat. Commun.* **7**, 10512 (2016).
103. Terajima, H. et al. ADARB1 catalyzes circadian A-to-I editing and regulates RNA rhythm. *Nat. Genet.* **49**, 146–151 (2016).
104. Yablonovitch, A. L., Deng, P., Jacobson, D. & Li, J. B. The evolution and adaptation of A-to-I RNA editing. *PLoS Genet.* **13**, e1007064 (2017).
105. Yablonovitch, A. L. et al. Regulation of gene expression and RNA editing in *Drosophila* adapting to divergent microclimates. *Nat. Commun.* **8**, 1570 (2017).
106. Galeano, F., Tomaselli, S., Locatelli, F. & Gallo, A. A-To-I RNA editing: the 'ADAR' side of human cancer. *Semin. Cell Dev. Biol.* **23**, 244–250 (2012).
107. Burns, M. B. et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* **494**, 366–370 (2013).
108. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
109. Fumagalli, D. et al. Principles governing A-to-I RNA editing in the breast cancer transcriptome article principles governing A-to-I RNA editing in the breast cancer transcriptome. *Cell Rep.* **13**, 277–289 (2015).
110. Han, L. et al. The genomic landscape and clinical relevance of A-to-I RNA editing in human cancers. *Cancer Cell* **28**, 515–528 (2015).
References 44, 109 and 110 show that editing is elevated in cancer and introduce the concept that so-called RNA mutation putatively helps promote malignancy.
111. Shoshan, E. et al. Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis. *Nat. Cell Biol.* **17**, 311–321 (2015).
112. Shimokawa, T. et al. RNA editing of the GLI1 transcription factor modulates the output of hedgehog signaling. *RNA Biol.* **10**, 321–333 (2013).
113. Cesarini, V. et al. ADAR2/miR-589-3p axis controls glioblastoma cell migration/invasion. *Nucleic Acids Res.* **46**, 2045–2059 (2018).
114. George, C. X., Ramaswami, G., Li, J. B. & Samuel, C. E. Editing of cellular self-RNAs by adenosine deaminase ADAR1 suppresses innate immune stress responses. *J. Biol. Chem.* **291**, 6158–6168 (2016).
115. Danan-Gotthold, M., Guyon, C., Giraud, M., Levanon, E. Y. & Abramson, J. Extensive RNA editing and splicing increase immune self-representation diversity in medullary thymic epithelial cells. *Genome Biol.* **17**, 219 (2016).
116. Kawahara, Y. et al. Glutamate receptors: RNA editing and death of motor neurons. *Nature* **427**, 801 (2004).
117. Srivastava, P. K. et al. Genome-wide analysis of differential RNA editing in epilepsy. *Genome Res.* **27**, 440–450 (2017).
118. Gurevich, I. et al. Altered editing of serotonin 2C receptor pre-mRNA in the prefrontal cortex of depressed suicide victims. *Neuron* **34**, 349–356 (2002).
119. Silberberg, G., Lundin, D., Navon, R. & Öhman, M. Deregulation of the A-to-I RNA editing mechanism in psychiatric disorders. *Hum. Mol. Genet.* **21**, 311–321 (2012).
120. Filippini, A. et al. Absence of the fragile X mental retardation protein results in defects of RNA editing of neuronal mRNAs in mouse. *RNA Biol.* **14**, 1580–1591 (2017).
121. Shamay-Ramot, A. et al. Fmrp interacts with adar and regulates RNA editing, synaptic density and locomotor activity in zebrafish. *PLoS Genet.* **11**, e1005702 (2015).
122. Bhogal, B. et al. Modulation of dADAR-dependent RNA editing by the *Drosophila* fragile X mental retardation protein. *Nat. Neurosci.* **14**, 1517–1524 (2011).
123. Zhang, R. et al. Quantifying RNA allelic ratios by microfluidic multiplex PCR and sequencing. *Nat. Methods* **11**, 51–54 (2014).
124. Gal-Mark, N. et al. Abnormalities in A-to-I RNA editing patterns in CNS injuries correlate with dynamic changes in cell type composition. *Sci. Rep.* **7**, 43421 (2017).
125. Hwang, T. et al. Dynamic regulation of RNA editing in human brain development and disease. *Nat. Neurosci.* **19**, 1093–1099 (2016).
126. Khermesh, K. et al. Reduced levels of protein recoding by A-to-I RNA editing in Alzheimer' s disease. *RNA* **22**, 1–13 (2016).
127. Eran, A. et al. Comparative RNA editing in autistic and neurotypical cerebella. *Mol. Psychiatry* **18**, 1041–1048 (2013).
128. Mele, M. et al. The human transcriptome across tissues and individuals. *Science* **348**, 660–665 (2015).
129. Stellos, K. et al. Adenosine-to-inosine RNA editing controls cathepsin S expression in atherosclerosis by enabling HuR-mediated post-transcriptional regulation. *Nat. Med* **22**, 1140–1150 (2016).
130. Garalde, D. R. et al. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **15**, 201–206 (2018).
131. Novoa, E. M., Mason, C. E. & Mattick, J. S. Charting the unknown epitranscriptome. *Nat. Rev. Mol. Cell. Biol.* **18**, 339–340 (2017).
132. Picardi, E., Horner, D. S. & Pesole, G. Single-cell transcriptomics reveals specific RNA editing signatures in the human brain. *RNA* **23**, 860–865 (2017).
133. Harjanto, D. et al. RNA editing generates cellular subsets with diverse sequence within populations. *Nat. Commun.* **7**, 12145 (2016).
134. Liu, H. et al. Genome-wide A-to-I RNA editing in fungi independent of ADAR enzymes. *Genome Res.* **26**, 499–509 (2016).
135. Liu, H. et al. A-To-I RNA editing is developmentally regulated and generally adaptive for sexual reproduction in *Neurospora crassa*. *Proc. Natl Acad. Sci. USA* **114**, E7756–E7765 (2017).
136. Bar-Yaacov, D. et al. RNA editing in bacteria recodes multiple proteins and regulates an evolutionarily conserved toxin-antitoxin system. *Genome Res.* **27**, 1696–1703 (2017).
137. McMahon, A. C. et al. TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. *Cell* **165**, 742–753 (2016).
138. Wettengel, J., Reautschnig, P., Geisler, S., Kahle, P. J. & Stafforst, T. Harnessing human ADAR2 for RNA repair - recoding a PINK1 mutation rescues mitophagy. *Nucleic Acids Res.* **45**, 2797–2808 (2017).
139. Montiel-Gonzalez, M. F., Vallecillo-Viejo, I., Yudowski, G. A. & Rosenthal, J. J. C. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl Acad. Sci. USA* **110**, 18285–18290 (2013).
140. Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching protein localization by site-directed RNA editing under control of light. *ACS Synth. Biol.* **6**, 1642–1649 (2017).

141. Vallecillo-Viejo, I. C., Liscovitch-Brauer, N., Montiel-Gonzalez, M. F., Eisenberg, E. & Rosenthal, J. J. C. Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme. *RNA Biol.* **15**, 104–114 (2018).
142. Bass, B. L. & Weintraub, H. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **55**, 1089–1098 (1988).
143. Rebagliati, M. R. & Melton, D. A. Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell* **48**, 599–605 (1987).
144. Wagner, R. W., Smith, J. E., Cooperman, B. S. & Nishikura, K. A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and *Xenopus* eggs. *Proc. Natl Acad. Sci. USA* **86**, 2647–2651 (1989).
145. Schrider, D. R., Gout, J.-F. & Hahn, M. W. Very few RNA and DNA sequence differences in the human transcriptome. *PLoS One* **6**, e25842 (2011).
146. Kleinman, C. L. & Majewski, J. Comment on "widespread RNA and DNA sequence differences in the human transcriptome". *Science* **335**, author reply 1302 (2012).
147. Eisenberg, E., Li, J. B. & Levanon, E. Y. Sequence based identification of RNA editing sites. *RNA Biol.* **7**, 248–252 (2010).
148. Pickrell, J. K., Gilad, Y. & Pritchard, J. K. Comment on "widespread RNA and DNA sequence differences in the human transcriptome". *Science* **335**, author reply 1302 (2012).
149. Lin, W., Piskol, R., Tan, M. H. & Li, J. B. Comment on "Widespread RNA and DNA Sequence Differences in the Human Transcriptome". *Science* **335**, author reply 1302 (2012).
150. Piskol, R., Peng, Z., Wang, J. & Li, J. B. Lack of evidence for existence of noncanonical RNA editing. *Nat. Biotechnol.* **31**, 19–20 (2013).
151. Ramaswami, G. et al. Accurate identification of human Alu and non-Alu RNA editing sites. *Nat. Methods* **9**, 579–581 (2012).
152. Picardi, E. & Pesole, G. REDItools: high-throughput RNA editing detection made easy. *Bioinformatics* **29**, 1813–1814 (2013).
153. Diroma, M. A., Ciaccia, L., Pesole, G. & Picardi, E. Elucidating the editome: bioinformatics approaches for RNA editing detection. *Brief. Bioinform*. https://doi.org/10.1093/bib/bbx129 (2017).
154. Bahn, J. H. et al. Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. *Genome Res.* **22**, 142–150 (2012).
155. Park, E., Williams, B., Wold, B. J. & Mortazavi, A. RNA editing in the human encode RNA-seq data. *Genome Res.* **22**, 1626–1633 (2012).
156. Ramaswami, G. et al. Identifying RNA editing sites using RNA sequencing data alone. *Nat. Methods* **10**, 128–132 (2013).
157. Zhang, Q. & Xiao, X. Genome sequence–independent identification of RNA editing sites. *Nat. Methods* **12**, 347–350 (2015).
158. John, D., Weirick, T., Dimmeler, S. & Uchida, S. RNAEditor: easy detection of RNA editing events and the introduction of editing islands. *Brief. Bioinform.* **18**, 993–1001 (2016).
159. Porath, H. T., Carmi, S. & Levanon, E. Y. A genome-wide map of hyper-edited RNA reveals numerous new sites. *Nat. Commun.* **5**, 4726 (2014).
160. Carmi, S., Borukhov, I. & Levanon, E. Y. Identification of widespread ultra-edited human RNAs. *PLoS Genet.* **7**, e1002317 (2011).
161. Sakurai, M. et al. A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. *Genome Res.* **24**, 522–534 (2014).
162. Bahn, J. H. et al. Genomic analysis of ADAR1 binding and its involvement in multiple RNA processing pathways. *Nat. Commun.* **6**, 6355 (2015).
163. Cattenoz, P. B., Taft, R. J., Westhof, E. & Mattick, J. S. Transcriptome-wide identification of A>I RNA editing sites by inosine specific cleavage. *RNA* **19**, 257–270 (2013).
164. Morse, D. P. & Bass, B. L. Detection of inosine in messenger RNA by inosine-specific cleavage. *Biochemistry* **36**, 8429–8434 (1997).
165. Kiran, A. & Baranov, P. V. DARNED: a database of RNA editing in humans. *Bioinformatics* **26**, 1772–1776 (2010).
166. Piechotta, M., Wyler, E., Ohler, U., Landthaler, M. & Dieterich, C. JACUSA: site-specific identification of RNA editing events from replicate sequencing data. *BMC Bioinformatics* **18**, 7 (2017).
167. Wang, Z. et al. RES-scanner: a software package for genome-wide identification of RNA-editing sites. *Gigascience* **5**, 37 (2016).
168. Zhang, F., Lu, Y., Yan, S., Xing, Q. & Tian, W. SPRINT: an SNP-free toolkit for identifying RNA editing sites. *Bioinformatics* **33**, 3538–3548 (2017).
169. Patterson, J. B. & Samuel, C. E. Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. *Mol. Cell Biol.* **15**, 5376–5388 (1995).
170. Cho, D. S. C. et al. Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J. Biol. Chem.* **278**, 17093–17102 (2003).
171. Keegan, L. P. et al. Functional conservation in human and *Drosophila* of metazoan ADAR2 involved in RNA editing: loss of ADAR1 in insects. *Nucleic Acids Res.* **39**, 7249–7262 (2011).
172. Palavicini, J. P., O'Connell, M. A. & Rosenthal, J. J. C. An extra double-stranded RNA binding domain confers high activity to a squid RNA editing enzyme. *RNA* **15**, 1208–1218 (2009).
173. Palavicini, J. P., Correa-Rojas, R. A. & Rosenthal, J. J. C. Extra double-stranded RNA binding domain (dsRBD) in a squid RNA editing enzyme confers resistance to high salt environment. *J. Biol. Chem.* **287**, 17754–17764 (2012).
174. Li, Q. et al. Caste-specific RNA editomes in the leaf-cutting ant Acromyrmex echinatior. *Nat. Commun.* **5**, 4943 (2014).
175. Zaranek, A. W., Levanon, E. Y., Zecharia, T., Clegg, T. & Church, G. M. A survey of genomic traces reveals a common sequencing error, RNA editing, and DNA editing. *PLoS Genet.* **6**, 8 (2010).
176. Chen, J.-Y. et al. RNA Editome in rhesus macaque shaped by purifying selection. *PLoS Genet.* **10**, e1004274 (2014).
177. Li, Z. et al. Evolutionary and ontogenetic changes in RNA editing in human, chimpanzee, and macaque brains. *RNA* **19**, 1693–1702 (2013).
178. Chen, L. Characterization and comparison of human nuclear and cytosolic editomes. *Proc. Natl Acad. Sci. USA* **110**, E2741–E2747 (2013).
179. Goldstein, B. et al. A-To-I RNA editing promotes developmental stage-specific gene and lncRNA expression. *Genome Res.* **27**, 462–470 (2017).
180. Stafforst, T. & Schneider, M. F. An RNA-deaminase conjugate selectively repairs point mutations. *Angew. Chem. Int. Ed. Engl.* **51**, 11166–11169 (2012).
181. Fukuda, M. et al. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Sci. Rep.* **7**, 41478 (2017).
182. Cox, D. B. T. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).

Acknowledgements
The authors thank O. Gabay for the graphical work and B. Knisbacher and the Levanon laboratory members for fruitful discussions. The authors also thank J. Rosenthal and J-B. Li for critical reading of the manuscript. This work was supported by the European Research Council (grant 311257), the Israel Science Foundation (1380/14) and the Minerva Stiftung ARCHES award from the Federal German Ministry for Education and Research (BMBF) to E.Y.L. E.E. was supported by the Israel Science Foundation (2673/17) and the United States-Israel Binational Science Foundation (094/2013).

Author contributions
Both authors contributed to all aspects of this manuscript.

Competing interests
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reviewer information
Nature Reviews Genetics thanks L. Keegan, M. Öhman and the other anonymous reviewer(s) for their contribution to the peer review of this work.