Gene Technology
ISSN: 2329-6682

Picimbon, Gene Technol 2015, 4:1
DOI: 10.4172/2329-6682.1000112

Review Article Open Access

# RNA Mutations: Source of Life

**Jean-François Picimbon***

*Biotechnology Research Center (BRC), Shandong Academy of Agricultural Sciences (SAAS), Jinan, China*

***Corresponding author:** Jean-François Picimbon, BRC-SAAS, 202 Industrial North Road, 250100, Jinan, China, Tel: 0(531)-83175350; E-mail: jfpicimbon@gmail.com

**Rec date:** May 22, 2014, **Acc date:** Jun 10, 2014, **Pub date:** Jun 16, 2014



**Abstract**

In a recent study published in PLoS ONE, our group showed that the genes coding for the chemosensory proteins in silkworm moth Bombyx mori were subjected to RNA editing. This post-transcriptional process, described from bacteria to complex organisms as plants and humans, by changing nucleotide sequence enables to increase a repertoire of proteins from a single RNA.

After a short introduction in which I remind the dogma initially established with the discovery of the double helix of DNA where a single gene encodes for a single protein, I present processes, RNA editing and alternative splicing, as two complementary modes for generating different proteins with different functions and regulation. Then, using intronless genes as example, I stress the role of RNA editing over alternative splicing. In the second part of the commentary, I discuss about RNA editing and alternative splicing in the course of evolution and I provide arguments supporting RNA editing as an early mechanism on earth that could have generated various proteins from few RNA. RNA mutations are proposed to be fuel for evolution of pheromone systems and even perhaps the mechanism that brought the original dormant RNA molecule to life. I propose RNA editing contributes source of life in an original RNA world. In the last part, I raise questions for new biotechnological prospects about cell molecular biology, RNA editing, genetic mutation, pathology, therapy, cloning and transgenesis.

**Keywords:** RNA; Editing; Mutation; Translational genetics; Evolution; Pheromone; Biomedicine; Plant transgenesis

IntroductionThe conception that one gene encodes for only one protein has root in the identification of the desoxynucleic acid double helix [1]. As soon as the double helix of what is now called DNA was discovered, scientists have started speculate on the gene-protein relationship. DNA is copied into RNA that is subsequently translated into protein. This can be seen as a dogma concept in molecular biology. The sequence of A-T-G-C bases in the DNA determines the sequence of amino acid residues in the protein product. Translational processes are faithful and pre-determined. RNA is only an intermediary molecule between DNA and protein. Regulatory mechanisms within the cell guarantee no errors in translation during the transfer of sequence information. Mutant proteins result only from gene mutations that are subsequently lethal.

## RNA Editing, an Alternative to Alternative Splicing

A recent study from Xuan et al. [2] in the silkworm moth Bombyx mori clearly demonstrates that the sequence of one single RNA strand does not necessarily matches the gene from which it has been made. A specific mechanism of RNA editing exists in probably all cells to create subtle RNA mutations (A-T-G-C replacement, deletion and/or insertion) and thereby new peptide molecules with potentially new functions. Mutations are not random errors. Most of them are non-synonymous and changed the amino acid motif in key structural regions of the protein. The α-helices α4 and α6 are not affected by RNA mutations in the family of *B. mori* chemosensory proteins (*BmorCSP*s). All the rest of the molecules, and in particular the N-terminal fragment (α1 helix), are subjected to a heavy load of RNA mutations. These mutations most probably change the function of the protein. Not only one or a few amino acids but motifs of up to thirteen amino acid residues are changed by RNA mutation. The N- and C-terminal regions are completely changed in most of all *BmorCSP*s. Most of the proteins produced by mutations are truncated proteins. Protein size changes from 12-14 to 7-9 kDa. Many cysteine residues are added, increasing the possibility of building new disulfide bonds. Many glycine residues are inserted on both sides of cysteine, tending to interrupt the α-helical fold of the protein. Many replacements of Leucine by Proline are found on *BmorCSP*s, suggesting a complete re-configuration of the α-helical profile through RNA mutations in the whole protein family. Mutations change up to 27-44% of the amino acid composition for *BmorCSP2*, *BmorCSP6* and *BmorCSP11* [2]. Investigation of the binding properties and the degree of conformational plasticity in the 3D structure of mutant variants for these proteins will allow us to clearly see whether RNA editing can create a totally different functional protein. In Xuan et al., we report about an enormous variety of point and frame shift mutations that are worm-made and in the silkworm specifically made by the female pheromone gland. Thus, the chance for them to create a totally different function protein is rather high according Darwin's theory of evolution [3].

Using genomic and complementary DNA sequence comparison for moth chemosensory genes in single individuals, we show high divergence within the RNA sequences as reported from other studies in both Human and mice [4,5]. Comparing gDNA and cDNA for six genes (CSP1, CSP2, CSP4, CSP14, PBP1 and Actin4) in five tissues (antennae, legs, wings, head and pheromone gland) at the individual level, we find that chemosensory genes do not produce a single faithful mRNA copy of the gene but rather multiple RNA copies with typo nucleotide changes (RNA editing). We find that RNA editing process in the silkworm *B. mori* concerns about 3 to 7% of RNA clones for a given *BmorCSP* and pheromone-binding protein-1 (PBP1). It is entirely tuned to the coding region. If this concerns other protein gene

families or is restricted to the family of chemosensory proteins needs to be cautiously investigated. RNA deep sequencing approaches have identified thousands of editing sites on RNAs encoding proteins involved in transcription, metabolism and many various neurological functions in human. However, only a few mutations have been found to cause altered amino acid assignment [6].

RNA mutations identified by high-throughput sequencing of transcriptome are always matters of debate. The reason is that they are rarely shown to be associated with amino acid replacements at the protein level. RNA sequencing data are often discussed without much reference to a proteomics data set [7,8]. In contrast, our study in the silkworm reports molecular biology, biochemistry and proteomic data demonstrating that a load of proteins with typo amino acid replacement or truncation of full amino acid motifs are made as a result of tissue-specific RNA mutations. This might explain the huge diversity of molecular isoforms observed for a large number of protein families not only in human, mammals and insects but also in fishes, plants and bacteria. RNA editing is here presented as a universal mechanism fundamentally necessary to enlarge the protein repertoire. Various physiological systems such as chemo-sensing, immunity, tissue development and xenobiotic metabolism can surely take a great advantage of such a typo RNA editing mechanism by expanding the number of signaling peptides and receptor molecules from a defined number of genes. In particular, one hypothesis could be that species having for instance a low number of olfactory genes like teleost fishes uses primarily RNA editing to produce a larger repertoire of olfactory receptors, while RNA editing is rather limited in species having a high number of olfactory genes in their genome such as Human and mice [9].

Until now the synthesis of multiple proteins from a single gene has been mainly associated to alternative splicing processes. However, creating new protein cannot be the fruits of alternative splicing alone. Alternative splicing is by definition limited to intron-exon genes. In the alternative splicing process, introns are spliced and exon motifs of a gene are redistributed or removed from the final RNA copy of the gene [10]. Thus, should the ancestral gene necessarily contain an intron for proper function? How could a gene without intron possibly function and evolve? Should protein diversity only limited to intron-containing genes?

The size of protein-coding gene, i.e. intron length, influences the degree of RNA variance indeed. It has been shown in plants that some RNA editing occurs in non-coding regions and that removal of intron in genes such as mitochondrial ribosomal protein *S10* or *Cox2* strongly reduces RNA editing in transcripts [11]. In the silkworm moth B. mori, we reveal a particularly high number of RNA mutations for the chemosensory gene *BmorCSP*4 that retains two retroposons, supporting the notion that insertion of retroposon or intron in the gene drastically boosts mRNA variance. Similarly, transpositions in mice are known to play a regulatory role in genesis of mRNA variants eventually causing increased protein expression [12]. However, in insects, we show that in intronless genes RNA editing can also proceed very efficiently at least in the silkworm moth *B. mori*. About one hundred mutations are detected for intronless chemosensory gene *BmorCSP14*, about two times more than the number of mutations detected on single-intron chemosensory genes (*BmorCSP*1 and *BmorCSP*2). This strongly suggests that RNA editing is not necessarily associated to intron splicing and that mitochondrial (intronless) and nuclear genomes (intron-containing) are most likely subjected to different RNA editing mechanisms. According to the endosymbiontic theory of the origin of eukaryotic cells, mitochondrial and nuclear genomes have separate evolutionary origins, with the mitochondrial genome being derived from the circular genome of bacteria [13]. The mechanisms underlying RNA editing
in mitochondria and bacteria should be then investigated in careful details. RNA editing in the nucleus has been found to occur mainly at the level of premature mRNAs and/or introns. In mammals, RNA editing has been shown to result in both codon changes and formation of splice sites [14]. Thus, alternative splicing and RNA editing are believed to work in pairs to produce a high diversity of proteins from a single gene; splicing will redistribute exons, while editing will modify their base composition subtly (Figure 1).

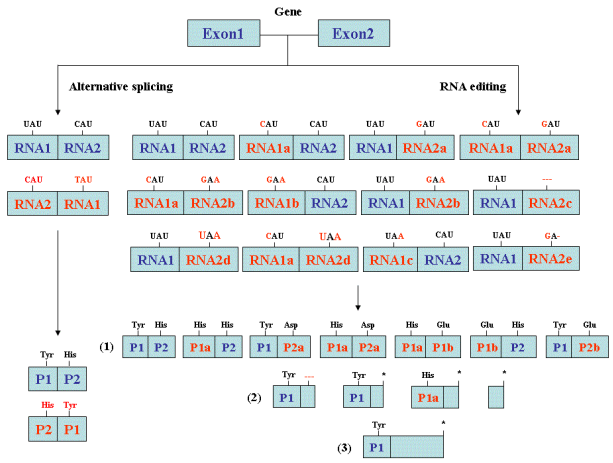


**Figure 1:** Comparison of the efficiency to produce multiple proteins between alternative splicing and RNA editing. A gene made of 2 exons (exon1 and exon2) can only create two proteins (P1-P2 and P2-P1) after intron removal. Gene evolution needs to add a few exons and introns to result in protein diversity. In contrast, RNA editing can create at once an infinite diversity of proteins (P1-P2, P1a-P2, P1-P2a, P1a-P2a, aborted P1, P1a or P1-P2, truncated P2 and elongated P1-P2, P1a-P2, etc). Specific RNA mutations can create proteins with subtle amino acid changes (1), shortened proteins lacking C-terminus or specific internal amino acid motifs following early-stop codon mutation or multi-nucleotide deletion (2) and longer proteins with prominent C-terminal tail (3) following a late-stop codon mutation [2]. RNA mutations are shown in red. -: nucleotide or amino acid deletion, *: early and late stop codon position. The creation of such an infinite diversity of proteins has been certainly crucial for the appearance of life and the evolution of all organisms.

## Why RNA Editing Came First

We think that RNA editing preceded RNA splicing in the course of evolution because it probably preceded the genesis of introns. Our finding about RNA mutations (one gene can produce multiple proteins) gives support to the intron-late theory [15]. Introns are rare in bacteria. They make up about 30-40% of the genome of animals, worms, insects and plants. 98% of human DNA is non-coding. Introns are highly divergent in size and sequence although many similarities can be found in species that diverged long ago [16]. The intron-early theory says that introns were necessary to assemble exons in order to

buildup genes and thereby multiple proteins [17]. We say that there is no need of intron to achieve protein diversity. One single individual gene can produce a huge diversity of RNA strands and protein chains. This has been found in insects but probably exist also in all other organisms including Human, all modern mammals, plants, worms and bacteria. Enzymes specifically involved in the editing of RNAs by conversion of adenosine to inosine, transfer RNA-specific adenosine deaminases acting on RNA (ADARs), have been identified in the most popular model bacterium, *Escherichia coli* [18]. If RNA mutations occur in bacteria such as *E. coli*, RNA mutations could have happened even at a very early stage of pre-cellular life, providing an extraordinary chance for life to start.

In the debate about the beginning of life, our finding about RNA mutations applauds to the RNA world theory [19]. It is known that RNA molecules carry genetic information but can also act as catalyst forms. It is also known that RNA can adopt different conformations, each with a different activity, and even that RNA is capable of self-replication [20]. There is now sufficient evidence to indicate that RNA appeared first on earth since RNA can catalyze electron transfer with iron's help [21]. RNA building blocks (nucleobases) came from space to build cells on earth [22,23]. Everything now in the eukaryotic cell is a piece of RNA. Interestingly, in our study of the moth eukaryotic cell, we clearly show that RNA strands differing by a few bases can lead to an enormous repertoire of proteins. If RNA can cause the reaction to mutate and copy itself (in conditions far from equilibrium such as under thermal shock or irradiation), there is no doubt that this phenomenon played a key role in the start of life. After RNA mutations, enough materials were provided to overcome degradation and eventually produce a diversity of more sophisticated molecules of life such as proteins [24-28]. In the lunar prebiotic RNA world, evolution was probably nothing else than an exponential amplification of relatively short single simple RNA molecules. Darwin's theory in term of information transfer to progeny for survival did not apply. In absence of cells and organisms, there were probably no processes of natural selection other than those driven by the mutation rates of the RNA strands and the kinetics of the chemical reactions they auto-catalyzed. In this bio-molecular magma, RNA-RNA interactions could well have created new molecules and thereby new reactions. RNA molecules and/or RNA-RNA complexes capable of high mutation rates, high levels of self-replication, high degree of conformational changes and high propensity for extra-molecular interactions had to lead the world at the time. Probably only super-replicators were allowed to survive and to evolve fast enough to give a chance to life to start [29-32].

Soon or later, there was a strong need to replace RNA by DNA and viral RNA genomes by more stable, bigger and irreversible DNA genomes [33,34]. This was a mandatory event not necessarily to increase protein diversity but surely to store genetic information and pass it on to its progeny following Darwin's evolution theory [35]. While mutations and RNA editing are required for the organism to evolve efficiently and rapidly, storage and stability of beneficial genetic traits are required to preserve and accumulate the results of successful adaptations.

The propensity for RNA to mutate by itself can be easily tested. It is feasible to perform a Miller's experiment in which a soup of chemicals are replaced by a soup of RNAs and subjected to conditions far from equilibrium before sequencing. A possible scenario for the appearance of life may be that RNA mutates following thermal shock or irradiation leading to the production of multi-functional proteins as shown on Figure 2 ("theory of RNA mutations"). The theory of RNA mutations says "a key element for the appearance of life is that RNA not only produces a large number of 'perfect' copies of itself but also an extremely large number of copies with tiny typo 'mistakes' in the base sequence. RNA concentration is now enough so that replication can take place under any plausible abiotic condition. RNA diversity is now enough so that multiple proteins can be built and eventually with time form membrane or tissue under the same plausible abiotic condition". This theory is not against Darwin's. It tends to explain what happened before in the RNA world, it means before the emergence of the DNA world in which individual differences acquired during the life are now heritable and pass on to next generation. How RNA mutations change the DNA structure to pass on to the next generation? In other words, how RNA mutations contribute to source of life in the DNA world? In the DNA world, there are many epigenetic events such as DNA methylation and histone modification that govern genes expressed in a particular cell. It could be that what it is transferred to progeny is not the RNA mutation but the way to make it. Heritable changes in specific ADAR gene activity could well serve to preserve a beneficial mutation without change in the DNA sequence. Also, DNA repair enzymes are known to be subjected to RNA editing. Thus, specific edited variants of DNA repair enzymes could also well serve the preservation of a beneficial mutation with change in the DNA sequence. The eukaryotic cell is a real mini-brain [36-38].

The theory that the RNA mutations are source of life is based on the multitude of RNA and protein mutants that can be generated from a single gene in a specific tissue such as the pheromone gland of the silkworm moth *B. mori*. On the basis of these results, a similar cloud of multiple diverse mutant RNAs trapped in clay and that are genetically linked through base replacement, heavily interact cooperatively to create even more mutations and insure survival and development of the RNA population in an extremely aggressive environment is proposed to be the key event of the appearance of life on earth.

Experimental support for this theory could be brought by a further in-depth analysis of RNA mutations in the group of *BmorCSPs*. A 7-9 kDa *BmorCSP2*, *BmorCSP6* or *BmorCSP11* having a totally replaced N-terminus, Glycine insertions near Cysteine and completely refolded α-helical pattern most likely has a different function than a native original model 12 kDa AQDKYEPIDDSFDA, IEAKYTDNIDVDEI, EEYYSSQYDNFDVE with a strictly conserved spacing between Cysteines and a clearly defined order of six α-helices. How different these functions are between native and edited *BmorCSPs* need to be studied very carefully. This could bring the first evidence ever that RNA editing creates a totally different functional protein, giving thereby most robust incontestable experimental support to the theory of RNA mutations for the source of life.

Our study about RNA in *Bombyx* shows that the number of RNA mutations is compatible with the emergence of a set of many different multi-functional proteins. This is demonstrated for two families of insect proteins, chemosensory proteins (CSPs) and odorant-binding proteins (OBPs). Both CSPs and OBPs are known as multi-functional protein partners regulating chemical sensing, immunological response, lipid transportation, cellular development, tissue regeneration and xenobiotics metabolism [39]. According to our study in B. mori, a hundred peptides can be made from a single multi-functional CSP or OBP gene. If edited CSP or OBP peptides have a function completely different than the native CSP or OBP peptides needs to be proved. However, it is noteworthy that CSP (and OBP) genes are not only found in insects but are also anchored in the crustacean genome [40].

The appearance of mutant proteins with multi-functional capability could be the key to trigger membrane development, cell formation and organ function. A cell is born from RNA mutations [41]. It remains to be understood how was formed RNA on the surface of quartz crystals in the middle of iron minerals trapped in the clay of earth's crust [42].

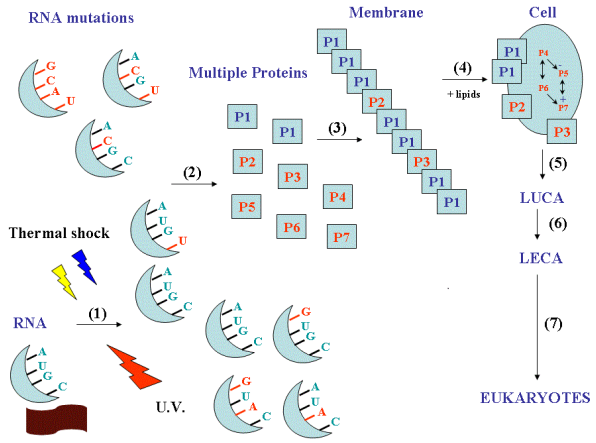


**Figure 2:** Theory of RNA mutations in the beginning of life. Following a thermal shock and/or U.V. irradiation, RNA mutates and replicates on the clay mineral surface of Earth (1). A high diversity of multi-functional proteins is produced (2). A cellular membrane is built with the addition of lipids (3). A cell is born and starts functioning (4), leading first to LUCA (5) then to LECA (6). Multiple LECA cells aggregate in tissues. Multiple tissues aggregate in all eukaryotic organisms of the tree of life (7). From now on, each individual develops following Darwin's theory of evolution (complex creatures will evolve from simplistic ancestors naturally over time). Beneficial DNA and RNA mutations are preserved and passed on to the next generation, leading to a totally new organism by natural selection. LUCA: Last Universal Common Ancestor (bacteria), LECA: Last Eukaryotic Common Ancestor (putative archaezoan species). P1-3: Membrane proteins, P4-P7: Intracellular proteins. +: activation, -: inhibition. The double arrow indicates the beginning of protein interactions (edification of macromolecular complexes). RNA mutations and resulting mutant proteins are shown in red. The birth of a population of RNAs for the synthesis of new mutant multi-functional proteins is proposed to be essential to shape new systems including the first one-cell organism.

To say that the propensity of RNA for base mutation could explain life origin may be too far speculation. However, we can easily speculate that a mechanism such as RNA editing has been crucial in the development of pheromone systems. One key point suggested by Darwin's evolution theory is that the evolution is accumulated. That an arm evolved to a wing has taken many changes and selections. That a protein evolves into a totally different functional protein also has to take many generations of changes and selections. It would not have taken so many generations, changes and selections to switch to different pheromone systems. Pheromones mediate reproductive isolation in closely related species. In moths, pheromones are various volatile carbon desaturated or unsaturated alcohols, aldehydes and acetates differing by carbon chain-length and the position of the double bound. A single change or a few changes during transcription of RNAs encoding pheromone transporters such as CSPs and OBPs are enough to change one single amino acid or a few residues in the binding pocket of the protein (Figure 3) [43]. This might not be enough to create a protein with totally different function. However, this might be enough to create a CSP or an OBP with different binding properties. Creating CSPs and OBPs with novel binding properties might have been essential in the evolution of many various physiological systems including olfaction and pheromone detection.

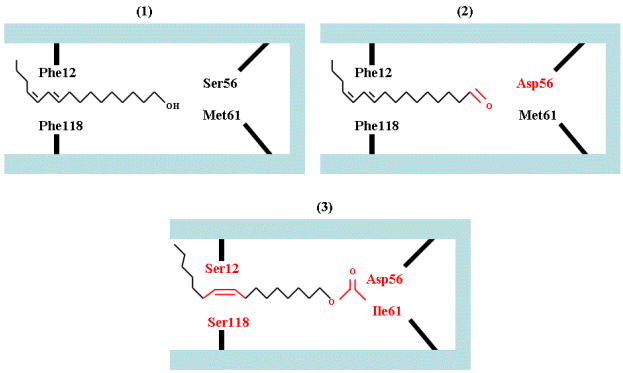


**Figure 3:** Mutations of RNA encoding sensory proteins and evolution of pheromone systems. Following Sandler et al. [43], Bombykol interacts in the binding pocket of protein transporters with phenylalanine, serine and methionine residues at position 12-118, 56 and 61, respectively (1). A switch to Bombykal in the moth pheromone system coordinates with a mutation from Ser56 to Asp56 in the protein binding site (2). A switch to cis-9-hexadecenyl acetate (Z9-16:Ac) during the diversification of pheromone systems coordinates with further mutations in the functional protein binding site (3).

Moreover, a general mechanism such as RNA editing has surely served the evolution of the whole organismal system. If every base on the RNA can change, then the gene exists in all possible copies without the genome being infected by multiple genes. This has been certainly crucial at a time when compound genomes reaching a critical size and high level of complexity lost evolvability [44]. RNA organisms such as viruses have a better and faster adaptation for instance by using RNA editing. Very similar to junk DNA, junk RNA might be also reservoir for evolution in general. It is just probably easier for the cell to mutate a base on RNA in the cytosol than to mutate a base on DNA in the nucleus. Then, once a mutation is proved to be beneficial, RNA returns to the nucleus to fix it via reverse transcription into the genome in order to pass it on to the next generation [45-48]. However, in the beginning of the RNA world, there was nothing like individual differences acquired during the life must be heritable or would do nothing with evolution. A kind of evolution existed among molecules differing by mutation and self-replication rates. The fastest were outnumbered and survived. Perhaps evolution in the Darwin's sense started with the compartmentalization of life, i.e. birth of cells. Starting from this moment, beneficial RNA mutations passed successfully from a cell to another cell by simple binary fission, complete fragmentation

or conjugation. This point is suggested in bacteria by the various forms of asexual reproduction. Darwin's theory applies later with the birth of the DNA world and the growth of a highly diverse community of complex organisms in all various ecological niches from every segmented piece of earth.

The more errors within one transcript, the less it happens. The chance for that the same several errors were recreated would be even less. We think that the cell makes as many errors as possible. It simply happens at least in cells from the moth pheromone gland. The sex-pheromone gland cells in *B. mori* invest a lot of energy and materials to create a ton of transcription errors, a load of multiple mutant RNAs and thereby an enormous variety of protein isoforms. Evolution requires specific mechanisms that can add new isoforms, new proteins and eventually with time new cells to provide a new functional system. Our work shows that an overload of RNA mutations has such a function in addition of facilitating genetic diversification and storage.

## Biotechnological Applications in Biomedicine and Transgenesis

This finding raised an interesting question about RNA editing and biotechnology. When producing a recombinant protein in a host cell by biotechnology, does RNA editing process occur and alter the protein function?

Mutations in DNA are well known to be responsible for genetic diseases. RNA editing produces alterations in RNA and cause pathological situations in similar extent. Flies lacking ADAR gene show seriously altered locomotion, paralysis and degeneration of the nervous system [49]. Worms with ADAR gene knocked out show seriously altered chemosensory behaviors [50]. Genetic, metabolic and neurological functions are seriously altered by deficiency in RNA editing in mammals and human. Defective RNA editing has been shown to induce a wide variety of diseases such as epilepsy, sclerosis, glioma, schizophrenia, obesity, diabetes, autism and a long list of cancers [51]. RNA editing can play a key role in the formation of cancerous tumors by either inactivating a cancer tumor suppressor or giving birth to a cancer tumor inducer [52,53]. For instance, over-expression of ADAR-1 is an established marker for diagnosis of esophageal or hepatic cancer [54,55]. Conversely, it has been proposed that RNA editing could help neutralize a pathogenic mutation through injection of a synthetic complementary RNA oligonucleotide, drug and/or regulation of microRNAs [56-58]. Control of RNA editing would be particularly relevant in the immunological system. ADARs are known to act in innate and adaptive immune systems, editing both host and pathogen transcriptomes [59]. It would be possible to combine the catalytic activity of an ADAR and/or another RNA editing enzyme with a sequence-specific RNA recognition site in order to biotechnologically change a mutation before RNA is translated into protein [60]. The success for such a minutious surgery on RNA strongly depends on enzyme and recognition site specificity. If the specificity of enzymes and RNA recognition sites is low, targeted gene mutation will be corrected in the sick tissue but most probably more harmful mutations will be created and broadly spread in healthy tissues.

An important discovery in our study is to show that RNA editing and ADARs play a key role in the olfactory/pheromone systems, particularly in the production of multiple odor/chemosensory binding proteins. This brings new avenues in diagnosis and/or treatment of problems related to anosmia and ageusia. Using moths and insects as study models could be particularly relevant for the development of RNA-directed gene therapies because of the high specificity of the RNA editing mechanisms underlying insect pheromone systems. We prove that a complex RNA editing exists in insect species such as the silkworm moth and that this phenomenon is tightly regulated in a tissue-specific manner. A huge amount of RNA mutations is found in the *Bombyx* pheromone gland compared to the antennae, legs, wings and head. Correlatively, a huge diversity of mutant peptides belonging to the superfamily of chemosensory and odorant binding proteins is found in the moth pheromone gland. Thus, there is a clear relationship between RNA editing process and tissue function. We still do not know if this pleiad of mutant CSP and OBP proteins helps produce different ratios of pheromone compounds synthesizes new components or simply increases pheromone concentration in the gland. However, the realization that a deficiency or aberrant activity of the RNA editing machinery tuned to CSPs or OBPs could lead to altered binding pocket and loss of odor/pheromone recognition in the insect orientates molecular, neurobiological and entomological researches towards the discovery of highly specific RNA editing mechanisms and their substrates. Identification of ADARs and RNA binding domains involved in insect olfaction and in particular in the recognition of sex pheromones could be extremely useful to develop substrate- or site-selective approaches to modify editing not only in human and mammals but also in plants and viruses.

We are not sure whether RNA mutations only originate from a limited number of ADAR enzymes in the silkworm moth *B. mori* as found for Drosophila melanogaster [61,62]. From our study, it is not very clear between the mutations made by RNA-dependent RNA polymerization and those by DNA-dependent RNA polymerization. We found in the moth transcriptome, a high number of A>I conversions that demand recognition of specific RNA duplexes but also a lot of non-A-to-I conversions, resulting in many different types of changes at the protein level. The diversity of RNA mutations and protein variants is shown to be particularly high in *B. mori*. Thus, it could be that RNA editing in insects involve multiple forms of ADARs and/or DNA-dependent RNA polymerases to create an extremely large repertoire of proteins differing by subtle changes in the amino acid composition. The RNA for each enzyme could be edited too, resulting in even more ADARs or polymerases with specific base editing functions. Alternatively, if RNA can mutate by itself on terrestrial crust, would it be possible that RNA mutates by itself in live cells in vivo? This would be particularly useful in eukaryotes since enzymes such as RNA-dependent RNA transcriptases are restricted to virions of RNA viruses and supposedly do not exist in the DNA world. This strong conventional wisdom, based on the lack of genes coding for RNA-dependent RNA transcriptase in eukaryotes, however, does not exclude the possibility that insect tissues such as the moth pheromone gland retain some RNA-dependent RNA transcription activities via specific mechanisms similar to those described in vertebrates [63-65]. It could also well be that specific infectious viruses and/or symbiotic bacteria take control of RNA editing not only in invertebrate but also in vertebrate host-cells [66,67].

The combination of multiple types of mechanisms (RNA-dependent RNA polymerization and DNA-dependent RNA polymerization) with very versatile DNA/RNA-dependent RNA polymerases such as Quelling Deficient-1 RdRP (QDE-1) from viruses would be extremely powerful to create an extremely large repertoire of proteins differing by subtle changes in the amino acid composition [68,69]. It is not excluded that a QDE-1 type of enzyme exists in some very ancient arthropod species. RdRP genes may well reside in the genome of

Tetraconata and/or some very old aquatic Collembola or when did they disappear? Has RNA-dependent RNA polymerization been totally lost during the transition from the RNA to the DNA world? All mechanisms and enzymes involved in RNA editing of chemosensory genes in insects need to be studied very carefully.

Would a transgene be subjected to RNA editing? If yes, would it be possible to correct the editing of a transgene in host-cells by RNA-directed gene "therapy" as envisioned by biomedical researchers? What would be the advantages to incorporate RNA variance concept into DNA microinjection technology? What would be the effects of inserting transgene along with specific ADAR enzymes? New biotechnologies may rise from creating mutant RNA sequences or inhibiting RNA editing in specific tissues of transgenic organisms.

RNA editing exists in plants, changing C to U in flowers and U to C in ferns and mosses [70]. Therefore, one should not assume that a transgene inserted into plant produces only one protein. On the contrary, our results clearly show that a transgene has a high probability to produce in the host-plant a high diversity of mutant transgenic proteins. Dozens of proteins can be made from a single RNA in moths. RNA editing in plants has been described not only for mitochondrial and chloroplast transcripts but also for nuclear transcripts [71-73]. In *Arabidopsis taliana*, it has even been shown that nuclear RNA editing can be modulated by pathogenic cues [74]. Therefore, it is very likely that a transgene is subjected to RNA editing. Introduction of a foreign gene might convey new functional proteins in the host-plant. However, this might also convey the plant with severe alterations broadly distributed across the entire body structure if the transgene produces toxic truncated isoforms. Transgenic RNA editing changes may occur in specific tissues or compartments of the plant resulting in untoward side effects. Therefore, using ADAR and/or another derivative enzyme coupled with a sequence-specific RNA binding domain to control RNA editing of transgene in the plant needs to be checked very carefully. To take over the RNA editing of the transgene in host plant organisms will be very useful to circumvent the problems related to post-translational modifications of the transgene only if the biotechnologically-manipulated enzyme can repair the transgenic RNA without editing the natural transcripts of the plant (Figure 4).

Different plant or mammal tissues have probably different RNA editing mechanisms as found in insects. Tissue-specific regulation of RNA editing should be therefore carefully investigated in both mammals and plants before to think of using "RNA-directed gene therapy" to develop new transgenic cell lines. This may help foresee potentials for the creation of new model organisms with industrial prospects.

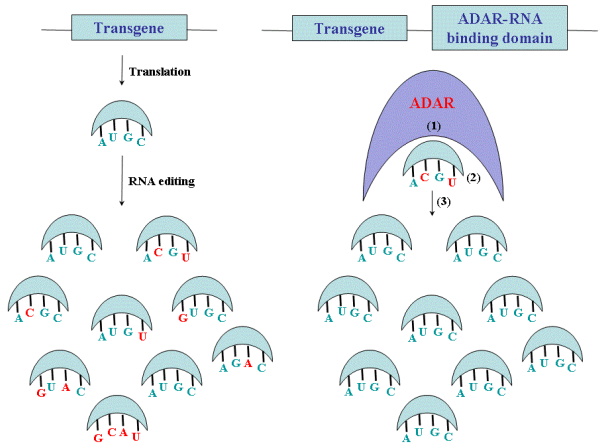


**Figure 4:** Principle of RNA-directed transgene therapy. Induction of transgene expression leads to the accumulation of transgenic mutant RNA transcripts due to the activation of RNA editing enzymes in host-cells (left). Co-expression of a transgene and a specific ADAR and/or another RNA editing enzyme coupled to a suitable RNA binding domain allows re-editing of transgenic mutant RNA transcripts and results inside host-cells in the production of only one single-type of protein, i.e. a perfectly faithful copy to the transgene (right). Specific RNA binding domains recognize specific mutation sites (1). Specific RNA editing enzymes correct specific RNA mutations (2). Only 'perfect' copies to the original transgene are produced in transgenic organisms (3).

This may also help foresee the future of a biotechnologically-manipulated organism after modifying the RNA pool it surely needs for a path of natural evolution. Like a patient, a transgenic plant or a transgenic cow will surely need to be assisted by RNA-directed gene therapy. We can assist them assuming we are capable to find ADARs and consors with ultra-specific enzymatic activities and RNA binding protein domains with recognition sites strictly restricted to mutated codons.

## Acknowledgment

Shandong Province Overseas High-Level Talents (NO.tshw20091015).

## References

1. Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171: 737-738.
2. Xuan N, Bu X, Liu YY, Yang X, Liu GX, et al. (2014) Molecular evidence of RNA editing in Bombyx chemosensory protein family. PLoS One 9: e86932.
3. Darwin C MA (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, Journal of Researches during H.M.S. Beagle's Voyage Round the World, London: John Murray, Albemarle Street, p. 162.
4. Ju YS, Kim J-II, Kim S, Hong D, Park H et al. (2011) Extensive genomic and transcriptional diversity identified through massively parallel DNA

and RNA sequencing of eighteen Korean individuals. Nature Genetics 43: 745-752.

5. Danecek P, Nellåker C, McIntyre RE, Buendia-Buendia JE, Bumpstead S, et al. (2012) High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. Genome Biol 13: 26.
6. Sakurai M, Ueda H, Yano T, Okada S, Terajima H, et al. (2014) A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. Genome Res 24: 522-534.
7. Hayden EC (2012) RNA studies under fire. Nature 484: 428.
8. Bass B, Hundley H, Li JB, Peng Z, Pickrell J, et al. (2012) The difficult calls in RNA editing. Interviewed by H Craig Mak. Nat Biotechnol 30: 1207-1209.
9. Niimura Y (2009) On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. Genome Biol Evol 1: 34-44.
10. Matlin AJ, Clark F, Smith CW (2005) Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 6: 386-398.
11. Castandet B, Choury D, Bégu D, Jordana X, Araya A (2010) Intron RNA editing is essential for splicing in plant mitochondria. Nucleic Acids Res 38: 7112-7121.
12. Li Z, Mulligan MK, Wang X, Miles MF, Lu L, et al. (2010) A transposon in Comt generates mRNA variants and causes widespread expression and behavioral differences among mice. PLoS One 5: e12181.
13. Thrash JC, Boyd A, Huggett MJ, Grote J, Carini P, et al. (2011) Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. Sci Rep 1: 13.
14. Laurencikiene J, Källman AM, Fong N, Bentley DL, Ohman M (2006) RNA editing and alternative splicing: the importance of co-transcriptional coordination. EMBO Rep 7: 303-307.
15. Logsdon JM Jr, Tyshenko MG, Dixon C, D-Jafari J, Walker VK, et al. (1995) Seven newly discovered intron positions in the triose-phosphate isomerase gene: evidence for the introns-late theory. Proc Natl Acad Sci U S A 92: 8507-8511.
16. Dulbecco R (1987) The Design of Life, Yale University Press, 89.
17. Gilbert W (1987) The exon theory of genes. Cold Spring Harb Symp Quant Biol 52: 901-905.
18. Wolf J, Gerber AP, Keller W (2002) tadA, an essential tRNA-specific adenosine deaminase from Escherichia coli. EMBO J 21: 3841-3851.
19. Atkins JF, Gesterland RF, Cech T (2006) The RNA world: the nature of modern RNA suggests a prebiotic RNA world. Plainview, N.Y: Cold Spring Harbor Laboratory Press, pp. 137-156.
20. Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. Science 323: 1229-1232.
21. Hsiao C, Chou IC, Okafor CD, Bowman JC, O'Neill EB, et al. (2013) RNA with iron(II) as a cofactor catalyses electron transfer. Nat Chem 5: 525-528.
22. Callahan MP, Smith KE, Cleaves HJ 2nd, Ruzicka J, Stern JC, et al. (2011) Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. Proc Natl Acad Sci U S A 108: 13995-13998.
23. Black RA, Blosser MC, Stottrup BL, Tavakley R, Deamer DW, et al. (2013) Nucleobases bind to and stabilize aggregates of a prebiotic amphiphile, providing a viable mechanism for the emergence of protocells. Proceedings of the National Academy of Sciences USA 110: 13272-13276.
24. Poole AM, Jeffares DC, Penny D (1998) The path from the RNA world. J Mol Evol 46: 1-17.
25. Alberts B, Johnson A, Lewis J et al. (2002) The RNA world and the origins of life. Molecular Biology of the Cell. 4th edition. New York, Garland Science.
26. Robinson R (2005) Jump-starting a cellular world: investigating the origin of life, from soup to networks. PLoS Biol 3: e396.
27. Vlassov AV, Kazakov SA, Johnston BH, Landweber LF (2005) The RNA world on ice: a new scenario for the emergence of RNA information. J Mol Evol 61: 264-273.
28. Yarus M (2010) Life from an RNA World: the ancestor within. Harvard University Press, p 198.
29. Wright MC, Joyce GF (1997) Continuous in vitro evolution of catalytic function. Science 276: 614-617.
30. Schmitt T, Lehman N (1999) Non-unity molecular heritability demonstrated by continuous evolution in vitro. Chem Biol 6: 857-869.
31. Ellington AD, Chen X, Robertson M, Syrett A (2009) Evolutionary origins and directed evolution of RNA. Int J Biochem Cell Biol 41: 254-265.
32. Robertson MP, Joyce GF2 (2014) Highly efficient self-replicating RNA enzymes. Chem Biol 21: 238-245.
33. Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. Annu Rev Microbiol 51: 151-178.
34. Smith O, Clapham A, Rose P, Liu Y, Wang J, et al. (2014) A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. Scientific Reports 4: 4003.
35. Leu K, Obermayer B, Rajamani S, Gerland U, Chen IA (2011) The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. Nucleic Acids Res 39: 8135-8147.
36. Rando OJ, Verstrepen KJ (2007) Timescales of genetic and epigenetic inheritance. Cell 128: 655-668.
37. Mattick JS, Mehler MF (2008) RNA editing, DNA recoding and the evolution of human cognition. Trends Neurosci 31: 227-233.
38. Mattick JS, Amaral PP, Dinger ME, Mercer TR, Mehler MF (2009) RNA regulation of epigenetic processes. Bioessays 31: 51-59.
39. Picimbon JF (2003) Biochemistry and evolution of CSP and OBP proteins. London, San Diego: Elsevier Academic Press, 539-566.
40. Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, et al. (2011) The ecoresponsive genome of Daphnia pulex. Science 331: 555-561.
41. Adamala K, Szostak JW (2013) Nonenzymatic template-directed RNA synthesis inside model protocells. Science 342: 1098-1100.
42. Cairns-Smith AG (1982) Genetic takeover: and the mineral origins of life. Cambridge University Press.
43. Sandler BH, Nikonova L, Leal WS, Clardy J (2000) Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chem Biol 7: 143-151.
44. Vasas V, Szathmáry E, Santos M (2010) Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. Proc Natl Acad Sci U S A 107: 1470-1475.
45. Lewin R (1983) How mammalian RNA returns to its genome. Science 219: 1052-1054.
46. Herbert A, Rich A (1999) RNA processing in evolution. The logic of soft-wired genomes. Ann N Y Acad Sci 870: 119-132.
47. Brosius J (2003) The contribution of RNAs and retroposition to evolutionary novelties. Genetica 118: 99-116.
48. Gommans WM, Mullen SP, Maas S (2009) RNA editing: a driving force for adaptive evolution? Bioessays 31: 1137-1145.
49. Palladino MJ, Keegan LP, O'Connell MA, Reenan RA (2000) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. Cell 102: 437-449.
50. Tonkin LA, Saccomanno L, Morse DP, Brodigan T, Krause M, et al. (2002) RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans. EMBO J 21: 6025-6035.
51. Maas S, Kawahara Y, Tamburro KM, Nishikura K (2006) A-to-I RNA editing and human disease. RNA Biol 3: 1-9.
52. Slotkin W, Nishikura K (2013) Adenosine-to-inosine RNA editing and human disease. Genome Med 5: 105.
53. Li JB, Church GM (2013) Deciphering the functions and regulation of brain-enriched A-to-I RNA editing. Nat Neurosci 16: 1518-1522.
54. Qin YR, Qiao JJ, Chan TH, Zhu YH, Li FF, et al. (2014) Adenosine-to-inosine RNA editing mediated by ADARs in esophageal squamous cell carcinoma. Cancer Res 74: 840-851.

55. Chen L, Li Y, Lin CH, Chan TH, Chow RK, et al. (2013) Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. Nat Med 19: 209-216.
56. Woolf TM, Chase JM, Stinchcomb DT (1995) Toward the therapeutic editing of mutated RNA sequences. Proc Natl Acad Sci U S A 92: 8298-8302.
57. Pray LA (2008) Gleevec: The breakthrough in cancer treatment. Nature Education 1: 37.
58. Wan G, Mathur R, Hu X, Zhang X, Lu X (2011) miRNA response to DNA damage. Trends Biochem Sci 36: 478-484.
59. Hamilton CE, Papavasiliou FN, Rosenberg BR (2010) Diverse functions for DNA and RNA editing in the immune system. RNA Biol 7: 220-228.
60. Maas S (2010) Gene regulation through RNA editing. Discov Med 10: 379-386.
61. Keegan LP, Leroy A, Sproul D, O'Connell MA (2004) Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. Genome Biol 5: 209.
62. Keegan LP, McGurk L, Palavicini JP, Brindle J, Paro S, et al. (2011) Functional conservation in human and Drosophila of Metazoan ADAR2 involved in RNA editing: loss of ADAR1 in insects. Nucleic Acids Res 39: 7249-7262.
63. Butcher SJ, Grimes JM, Makeyev EV, Bamford DH, Stuart DI (2001) A mechanism for initiating RNA-dependent RNA polymerization. Nature 410: 235-240.
64. Lipardi C, Wei Q, Paterson BM (2001) RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. Cell 107: 297-307.
65. Birchler JA (2009) Ubiquitous RNA-dependent RNA polymerase and gene silencing. Genome Biol 10: 243.
66. Castelló A, Quintas A, Sánchez EG, Sabina P, Nogal M, et al. (2009) Regulation of host translational machinery by African swine fever virus. PLoS Pathog 5: e1000562.
67. Lutay N, Ambite I, Grönberg-Hernandez J, Rydström G, Ragnarsdóttir B, et al. (2013) Bacterial control of host gene expression through RNA polymerase II. The Journal of Clinical Investigations 123: 2366-2379.
68. Iyer LM, Koonin EV, Aravind L (2003) Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. BMC Struct Biol 3: 1.
69. Lee HC, Aalto AP, Yang Q, Chang SS, Huang G, et al. (2010) The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. PLoS Biol 8.
70. Takenaka M, Zehrmann A, Verbitskiy D, Härtel B, Brennicke A (2013) RNA editing in plants and its evolution. Annu Rev Genet 47: 335-352.
71. Tillich M, Lehwark P, Morton BR, Maier UG (2006) The evolution of chloroplast RNA editing. Mol Biol Evol 23: 1912-1921.
72. Yura K, Go M (2008) Correlation between amino acid residues converted by RNA editing and functional residues in protein three-dimensional structures in plant organelles. BMC Plant Biol 8: 79.
73. Meng Y, Chen D, Jin Y, Mao C, Wu P, et al. (2010) RNA editing of nuclear transcripts in Arabidopsis thaliana. BMC Genomics 11 Suppl 4: S12.
74. García-Andrade J, Ramírez V, López A, Vera P (2013) Mediated plastid RNA editing in plant immunity. PLoS Pathog 9: e1003713.