

Received Date : 30-Nov-2015  
Accepted Date : 15-Jan-2016  
Article type : Meeting Review

NEWS AND VIEWS

MEETING REVIEW

### **Mobile Genetic Elements: *In Silico, In Vitro, In Vivo***

Irina R. Arkhipova\* and Phoebe A. Rice†

\*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, USA; †Department of Biochemistry and Molecular Biology, The University of Chicago, 929 E. 57th Street, Chicago, IL 60637, USA

**Mobile genetic elements (MGEs), also called transposable elements (TEs), represent universal components of most genomes and are intimately involved in nearly all aspects of genome organization, function, and evolution. However, there is currently a gap between fast-paced TE discovery *in silico*, stimulated by exponential growth of comparative genomic studies, and a limited number of experimental models amenable to more traditional *in vitro* and *in vivo* studies of structural, mechanistic, and regulatory properties of diverse MGEs. Experimental and computational scientists came together to bridge this gap at a recent conference, “Mobile Genetic Elements: *in silico, in vitro, in vivo*,” held at the Marine Biological Laboratory (MBL) in Woods Hole, MA, USA.**

*Keywords:* transposable elements, retrotransposons, transposition, genomics, bioinformatics

Correspondence: Irina Arkhipova; fax 1-508-457-4727; E-mail: iarkhipova@mbl.edu; Phoebe Rice; fax 1-773-702-0439; E-mail: PRice@uchicago.edu

The existence of MGEs was first described by Barbara McClintock in the 1950s (McClintock 1950, 1953). Although poorly appreciated at first, work on MGEs began to gather steam in the late 1970s following technical revolutions in molecular biology, and McClintock was finally awarded a Nobel Prize for her insightful work in 1983. The following decades saw a steady stream of advances in understanding the details of how such elements can move as well as clever biotechnology applications for them. Most recently, revolutions in DNA sequencing technology have catapulted the field into a new chapter. These new technologies, and the flood of information provided by them, permeated much of the meeting “Mobile genetic elements: *in silico, in vitro, in vivo*” held on September 3-5, 2015 (Fig. 1). The recent affiliation of the MBL with The University of Chicago stimulated the expansion of this initially regional meeting beyond the north-eastern US, to include 78 participants from 14 states and 3 Canadian provinces.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.13543

This article is protected by copyright. All rights reserved.

In keeping with McClintock's original suggestion that MGEs can have regulatory roles, one theme of the meeting was the interactions between MGEs and their hosts. The keynote talk by **Marlene Belfort** (SUNY at Albany) described how mobile introns and inteins (self-splicing protein domains) could provide adaptive features to their hosts. Inteins have generally been thought of as parasites that the host can tolerate because they splice themselves out of their host proteins after translation, thus avoiding interference with host function. However, Dr. Belfort described examples of inteins where splicing is dependent on particular environmental conditions (*e.g.* redox state and temperature), and can thus inhibit their host proteins' function when it would be detrimental to the cell (Callahan *et al.* 2011; Topilina *et al.* 2014, 2015; Novikova *et al.* 2015).

Many mobile introns and inteins spread by exploiting their host organism's double strand break repair process: they have a "homing endonuclease" activity that cleaves copies of (or close relatives of) their host gene that they have not yet invaded, and are then copied into the new locus during the repair process. It is therefore perhaps more important for homing endonucleases to have the ability to recognize the coding region for conserved protein sequences than a particular DNA sequence *per se*. **Barry Stoddard** (Fred Hutchinson Cancer Research Center) examined the comparative target specificity of homing endonucleases with a combination of bioinformatics, biochemistry, and crystallography, showing that homing endonucleases concentrate their recognition efforts on the most conserved nucleotides of each codon, but that when they do evolve to recognize different targets, they can do so with very little structural change to the protein backbone.

During the course of evolution, some MGEs have become "domesticated" to the point where they are an integral part of an organism. Striking examples of this can be found in adaptive immunity systems that allow the host organisms the flexibility to react to new attacks. **Eugene Koonin** (NCBI/NIH), making splendid use of genomic databases, described how both our own V(D)J recombination system and microbial CRISPR/Cas systems evolved from transposon ancestors (Krupovic and Koonin, 2015). The Cas1 proteins that are required for spacer acquisition in CRISPR systems are also conserved in a set of transposons termed "casposons" (Krupovic *et al.* 2014). **Alison Hickman** (NIDDK/NIH) presented the first evidence that purified casposon Cas1 protein can indeed catalyze integration of paired Casposon ends into a target DNA (Hickman and Dyda, 2015). As Cas1 proteins have a very different catalytic domain structure (Kim *et al.* 2013; Nuñez *et al.* 2014) from any of the other recombinases known to act as transposases, this represents a new family of DNA transposase.

A mother lode of new insights and entirely new MGEs lurks within the ever-expanding sequence databases, but for those outside of the informatics field, exploiting those resources can be a daunting task. To facilitate data processing as well as database utilization, the meeting offered a bioinformatics workshop led by **Vladimir Kapitonov** (NCBI/NIH). The first issue in genome assembly and concomitant TE identification is correctly analyzing the raw sequence data, which can be highly problematic for repetitive sequences. **Chandrasen Soans** offered the BaseSpace cloud computing environment from Illumina, which permits hassle-free workflows for next-generation data assembly and downstream applications. **Chong Chu** (University of Connecticut) presented a computational repeat identification pipeline called REPdenovo, the output of which was validated by repeat alignment to PacBio long reads. The pipeline can assemble highly similar genomic repeats directly from raw shotgun reads, and can be used as a discovery tool for novel repeats which typically fail to assemble from such short reads.

The existence of multiple genomic sequences for the same or similar species provides a new window on the ecology of MGEs. **Nelson Lau** (Brandeis University) described his approach to exploiting MGE ecology utilizing his lab's new database, TIDAL, of transposon insertions and deletions in fruit flies (Rahman *et al.* 2015). The problem of mapping RNA-seq data to repetitive sequences was addressed by **Ying Jin** and **Molly Hammell** (Cold Spring Harbor Laboratory) in a software package called TEtranscripts, which deals with ambiguous short-read mapping to improve recovery of TE-derived transcripts in large transcriptomic datasets generated during differential expression analyses (Jin *et al.* 2015).

A different approach to exploiting the databases, presented by **Vladimir Kapitonov**, is to search for new types of mobile elements beginning with known catalytic modules. Evolution has exploited certain catalytic domains repeatedly in different contexts – for instance, the catalytic domains of “classical” DNA transposases and retroviral integrases are closely related to one another and to RNaseH (Hickman and Dyda 2015). Dr. Kapitonov thus begins with sequence alignments of catalytic domains likely to be exploited by MGEs, searches the databases for proteins with related domains, then looks for patterns in the larger genomic context where those proteins were encoded. In this way, he was able to trace the origins of vertebrate RAG1 and RAG2 V(D)J recombinase subunits to a single transposon found in echinoderms and oysters (Kapitonov and Koonin, 2015).

In looking beyond the most common model organisms, an amazing variety of MGEs and other genomic rearrangements can be found. For example, **Jack Gilbert** (Argonne National Laboratory, The University of Chicago, and the MBL) described the variety of mobile elements within our own microbiome, and **Eugene Koonin** explained that the variety of CRISPR/Cas systems is even greater than previously suspected (Koonin and Krupovic 2015). **Eugene Gladyshev** (Harvard University), working with the filamentous fungus *Neurospora crassa*, described how that organism's unusual ability to generate mutations induced by DNA sequence repeats may lead to understanding a new mechanism for the direct recognition of sequence homology between DNA duplexes (Gladyshev and Kleckner 2014). **Aruna Govindaraju** (University of Texas at Arlington) combined *in silico* modeling with *in vitro* biochemical studies of the R2Bm retrotransposon from *Bombyx mori* to reveal noncanonical catalytic residues in the restriction endonuclease-like domain, which shows an intriguing structural similarity to Holliday junction resolvases (Mukha *et al.* 2013).

In keeping with MBL's mission, several aquatic systems were described, including clams, rotifers, and copepods. **Michael Metzger** (Columbia University) described a contagious cancer affecting clams that is associated with a retrovirus-like mobile genetic element, *Steamer* (Arriagada *et al.* 2014; Metzger *et al.* 2015). **Irina Arkhipova** (MBL) focused her talk on asexual bdelloid rotifers, small aquatic invertebrates in which potent silencing mechanisms apparently shaped an unusual TE landscape with atypically low overall transposon content but a large variety of MGE families and superfamilies, including novel retroelements of exceptionally complex structure. **Grace Wyngaard** (James Madison University) introduced the audience to copepods, freshwater crustaceans with many experimentally convenient features such as small size and external eggs. These undergo a dramatic genomic pruning: at a predetermined stage of early development, billions of base pairs of TEs (>50% of the total genome) are deleted from the genomes of somatic cells and presumably “recycled” as a nutritional source for the developing organism (Sun *et al.* 2015).

Another recurring theme of the meeting was the interplay between the world of MGEs and that of viruses and bacteriophages. In the case of the Staphylococcal pathogenicity islands (SaPIs), bacteriophages are hijacked to package SaPI DNA rather than their own. This trick gets the SaPIs efficiently transferred horizontally to new strains. **Richard Novick** (New York University) described his lab's recent advances in understanding the regulatory interplay between the SaPIs and their helper phages (Chen *et al.* 2015). He also described how the presence of SaPIs can trigger the horizontal transfer of chromosomal genes that are distant from the integrated SaPI or prophage genomes. Horizontal transfer of genetic material among bacteria can also be facilitated by gene transfer agents (Mercer and Lang 2014; Lang *et al.* 2015). Recent advances in understanding the mechanism and evolution of these domesticated phage-like entities were described by **Andrew Lang** (Memorial University) and **Migun Shakya** (Dartmouth College), respectively. Mother Nature's penchant for recycling basic machinery was highlighted by **Phoebe Rice** (University of Chicago), who found (through *in silico* approaches) homologies between the replication initiator / helicase of the SaPIs and a conserved protein encoded by the MGE that carries methicillin resistance in Staphylococci. She also presented the crystal structure of the latter, showing that it forms a hexameric ring as is common for helicases that support DNA replication.

While DNA transposons and phages represent the most prominent forces shaping prokaryotic genomes, in many eukaryotes genome re-shaping and remodeling has been taken over by retroelements. In particular, mammalian genomes are densely populated by LINE elements, which together with non-autonomous Alu repeats make up about one-third of our own genome. **John Moran** (University of Michigan) explained the extraordinary success of the L1-Alu pair by experimentally demonstrating the affinity of L1-encoded ORF2 (reverse transcriptase) to poly(A)-tracts at the 3'-termini of both L1 and Alu elements, which turned out to be critical for their retrotransposition *in cis* and *in trans* (Doucet *et al.* 2015). **Tammy Morrish** (University of Toledo) explored the relative contribution of L1 retrotransposition and other recombination-based mechanisms to chromosome end maintenance in mouse B-cell lymphomas lacking telomerase, finding evidence consistent with multiple mechanisms. **Pedro Rocha** (New York University) reported that, in the absence of the DNA damage sensor protein 53BP1, retrotransposon-containing loci show an increase in the amount of chromosomal interactions that they are engaged in, leading to more frequent involvement in chromosomal translocations.

Computational studies uncovered different complementing aspects of mammalian retrotransposon biology. **Steve Criscione** (Brown University) presented evidence of L1 impact on adjacent genes mediated by its antisense promoter, which gives rise to numerous alternative transcripts in multiple tissues, and effectively drives expression of over a thousand human genes. In addition to LINES, a sizeable fraction of mammalian genomes is derived from endogenous retroviruses (ERVs), which have abandoned their infectious lifestyles upon entering the germ line and now spread intragenomically as retrotransposons. **Kateryna Makova** (Pennsylvania State University) demonstrated the power of sophisticated statistical analysis methods to uncover regional integration and fixation landscapes of ERVs in mouse and human genomes by comparing genomic environments of *ex vivo* integration sites in tissue culture cells to those of naturally occurring fixed and polymorphic insertion sites.

Traditional model organisms were also featured prominently at the meeting. The fruit fly *Drosophila*, in which eukaryotic MGEs were initially discovered, continues to offer rich

opportunities for investigating their roles in fundamental cellular processes. **Daniel Barbash** (Cornell University) emphasized the involvement of TEs and satellite repeats in heterochromatin formation and chromosome segregation, and also introduced a new computational framework (conTE<sub>xt</sub>) for identification and quantitation of tandemly repeated sequences across natural populations. **Prashanth Rangan** (RNA Institute, SUNY at Albany) reported fascinating connections between dSETDB1, a *Drosophila* histone H3K9 methyltransferase required for heterochromatin formation, and its control of Wnt signaling, which is in turn independently down-regulated by other mutations that derepress TEs. **Yang Yu** (Cold Spring Harbor Laboratory) introduced Panoramix, a component of the Piwi complex in *Drosophila* which reinforces transcriptional silencing *via* recruitment of general silencing machinery to nascent transcripts (Yu *et al.* 2015). The fruit fly was also used by **Lisa Krug** (Cold Spring Harbor Laboratory) to express human TDP-43 protein implicated in several neurodegenerative disorders, and to make a critical connection between endogenous retrotransposon activation and neurodegenerative effects of hTDP-43.

Of unicellular eukaryotes, yeasts provide an unmatched advantage in the ease of combining experimental approaches employed by geneticists, biochemists, and cell biologists. **Joan Curcio** (Wadsworth Center) reported the stable association between Ty1 retrotransposon assembly sites and spindle pole bodies (the functional equivalents of the centrosome) in the budding yeast *Saccharomyces cerevisiae*, arguing that Ty1 nucleocapsid assembly may counteract the normally occurring asymmetric centrosome inheritance and could negatively impact reproduction. **Patrick Maxwell** (Rensselaer Polytechnic Institute) uses budding yeast as a model system to study aging, and provided evidence of preferential accumulation of Ty1 retrotransposition intermediates in yeast mother cells with age. The fission yeast, *Schizosaccharomyces pombe*, was used by **Jake Jacobs** (Rutgers University) to demonstrate that Tf1 retrotransposable elements are targeted to arrested replication forks when Sap1 protein is used as polar fork barrier, providing an explanation for universal avoidance of coding sequences by LTR retrotransposons (Jacobs *et al.* 2015).

Replication forks are also targeted by prokaryotic MGEs from several families (Fricker and Peters 2014). Transposon Tn7 is an interesting example that encodes different accessory proteins that steer the transposition apparatus towards different targets, with TnsE being the one that targets replication forks. **Joe Peters** (Cornell University) presented a series of TnsE structures that help explain how it does its job (Shi *et al.* 2015). The value of structural biology in conjunction with biochemistry was also highlighted in other presentations. In particular, **Marlene Belfort** reported a high-resolution cryo-EM structure of a self-splicing bacterial group II intron, which revealed the similarity of its reverse transcriptase catalytic domain to eukaryotic telomerases, and its maturase domain to spliceosomal PRP8 proteins, providing new insights into their catalytic function and highlighting the complexity of ancestral relationships.

Of note, applications of ecological models to TE analysis are currently gaining ground. Molecular ecology of TEs views the genome as a mini-ecosystem, and focuses on their co-existence as members of the genome community, which are occasionally competing for shared resources, e.g. for common enzymes or available integration sites (Venner *et al.* 2009). **Nicola Neretti** (Brown University) used models derived from ecological communities to reveal competition between LINE retrotransposon subfamilies at specific times during the evolution of the mammalian genome, and **Brent Saylor** (University of Guelph) applied ecological methods to analyze spatial distribution of TEs along the chromosomes of different species.

## Summary

The conference demonstrated the value of bringing together researchers studying mobile DNA in different domains of Life – Bacteria, Archaea and Eukarya – as well as integrating experimental and computational approaches to achieve deeper understanding of the complex interplay between TEs and their hosts. Its predecessor, the regional mobile DNA meeting held in Cold Spring Harbor in 2013, drew 75 participants from 11 states and 7 countries, and also served as an excellent venue due to its rich history of TE research. The relatively compact scale of the meeting allowed us to select half of the talks from submitted abstracts, and provided unmatched opportunities for direct interactions between computational and experimental scientists, as well as between faculty, postdocs, and graduate students. Importantly, in a post-meeting survey, one-half of the conference participants responded that the meeting helped them forge new collaborations and /or drive their research in new directions.

Overall, the meeting highlighted the need for a regular forum which specifically aims to bring together experimental and computational TE researchers in order to accelerate the identification of interesting new targets for experimentation, and to provide designers of automated genome annotation pipelines with essential information on the functionally important sequence elements in known genomic components and on the properties of hitherto unappreciated novel genomic components. For eukaryotic TEs, such a forum was provided in 2006-2012 by the conference “Genomic impact of eukaryotic transposable elements” (Batzer *et al.* 2007; Jurka 2009; Arkhipova *et al.* 2012), which was organized in Asilomar on a triannual basis by the late Dr. Jerzy Jurka, founder of the Genetic Information Research Institute and of Repbase, the database of eukaryotic TEs (Jurka *et al.* 2005). The current meeting made it particularly evident that such a forum needs to include both prokaryotic and eukaryotic TE researchers.

A word cloud representation of aggregate talk and poster abstracts is given in Fig. 2. It clearly illustrates the importance of genomic context in TE studies, and the general emphasis on TE-host cell interactions. Indeed, it is through the advances of genome assembly and analysis methods in comparative genomics that we became aware of the prevalence of large segments of bacterial DNA capable of horizontal gene transfer. The most recent advances in long-read assembly and analysis are now permitting discovery of large movable genetic entities in eukaryotes, which previously escaped detection due to the difficulties in assembly and annotation of large genomes. In addition to providing new insights into genome biology, MGE studies have the potential to bring new tools into the existing genetic toolbox, facilitating new approaches to genome engineering and restructuring -- something that mobile elements do naturally. Thus, investigators working on different systems and with different perspectives, not necessarily being within one another's regular orbits, could benefit most from meeting each other. It is exactly this kind of interaction that this conference intends to promote in the years to come.

## Acknowledgements

We thank The University of Chicago, Illumina Inc., and Integrated DNA Technologies for financial support of the meeting, and Bill Reznikoff for critical reading of the manuscript. Research in the authors' laboratories is supported by NSF MCB-1121334 and NIH R01 GM111917 (I.A.), and by NIH R01 GM101989 and R21 AI117593 (P.R.)

## References

- Arriagada G, Metzger MJ, Muttray AF, Sherry J, Reinisch C, Street C, Lipkin WI, Goff SP (2014) Activation of transcription and retrotransposition of a novel retroelement, *Steamer*, in neoplastic hemocytes of the mollusk *Mya arenaria*. *Proc Natl Acad Sci USA* **111**, 14175-14180.
- Arkhipova IR, Batzer MA, Brosius J, Feschotte C, Moran JV, Schmitz J, Jurka J (2012) Genomic impact of eukaryotic transposable elements. *Mob DNA* **3**, 19.
- Batzer MA, Deininger PL, Jurka J, Moran JV. (2007). Introduction for the Gene special issue dedicated to the meeting "Genomic impact of eukaryotic transposable elements" at Asilomar. *Gene* **390**, 1-2.
- Callahan BP, Topilina NI, Stanger MJ, Van Roey P, Belfort M (2011) Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nat Struct Mol Biol.* **18**, 630-633.
- Chen J, Ram G, Penadés JR, Brown S, Novick RP (2015) Pathogenicity island-directed transfer of unlinked chromosomal virulence genes. *Mol Cell* **57**,138-149.
- Doucet AJ, Wilusz JE, Miyoshi T, Liu Y, Moran JV (2015) A 3' poly(A) tract is required for LINE-1 retrotransposition. *Mol Cell* **60**, 728-741.
- Fricker AD, Peters JE (2014) Vulnerabilities on the lagging-strand template: opportunities for mobile elements. *Annu Rev Genet* **48**,167-186.
- Gladyshev E, Kleckner N (2014) Direct recognition of homology between double helices of DNA in *Neurospora crassa*. *Nat Commun.* **5**, 3509.
- Hickman AB, Dyda F. (2015a) The casposon-encoded Cas1 protein from *Aciduliprofundum boonei* is a DNA integrase that generates target site duplications. *Nucleic Acids Res.* **43**, 10576-10587.
- Hickman AB, Dyda F (2015b) Mechanisms of DNA transposition. *Microbiol Spectr.* **3**, MDNA3-0034-2014.
- Jacobs JZ, Rosado-Lugo JD, Cranz-Mileva S, Ciccaglione KM, Tournier V, Zaratiegui M (2015) Arrested replication forks guide retrotransposon integration. *Science* **349**, 1549-1553.
- Jin Y, Tam OH, Paniagua E, Hammell M (2015) TETranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593-3599.
- Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res.* **110**, 462-467.
- Jurka J (2009) "Genomic Impact of Eukaryotic Transposable Elements", Pacific Grove, California, February 6-10, 2009. *Gene* **448**,103.
- Kapitonov VV, Koonin EV (2015) Evolution of the RAG1-RAG2 locus: both proteins came from the same transposon. *Biol Direct* **10**, 20.
- Kim TY, Shin M, Huynh Thi Yen L, Kim JS (2013) Crystal structure of Cas1 from *Archaeoglobus fulgidus* and characterization of its nucleolytic activity. *Biochem Biophys Res Commun.* **441**, 720-725.
- Koonin EV, Krupovic M (2015) Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nat Rev Genet.* **16**, 184-192.
- Krupovic M1, Makarova KS, Forterre P, Prangishvili D, Koonin EV (2014) Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol.* **12**, 36.

- Lang AS, Zhaxybayeva O, Beatty JT (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nat Rev Microbiol.* **10**, 472-482.
- McClintock B (1950) The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA* **36**, 344-355.
- McClintock B (1953) Induction of instability at selected loci in maize. *Genetics* **38**, 579-599.
- Mercer RG, Lang AS (2014) Identification of a predicted partner-switching system that affects production of the gene transfer agent RcGTA and stationary phase viability in *Rhodobacter capsulatus*. *BMC Microbiol.* **14**, 71.
- Metzger MJ, Reinisch C, Sherry J, Goff SP (2015) Horizontal transmission of clonal cancer cells causes leukemia in soft-shell clams. *Cell* **161**, 255-263.
- Mukha DV, Pasyukova EG, Kapelinskaya TV, Kagramanova AS (2013) Endonuclease domain of the *Drosophila melanogaster* R2 non-LTR retrotransposon and related retroelements: a new model for transposition. *Front Genet.* **4**, 63.
- Novikova O, Jayachandran P, Kelley DS, Morton Z, Merwin S, Topilina NI, Belfort M. (2015) Intein clustering suggests functional importance in different domains of life. *Mol Biol Evol.* Nov 25. pii: msv271. [Epub ahead of print]
- Nuñez JK, Kranzusch PJ, Noeske J, Wright AV, Davies CW, Doudna JA (2014) Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat Struct Mol Biol.* **21**, 528-534.
- Rahman R, Chirn GW, Kanodia A, Sytnikova YA, Brembs B, Bergman CM, Lau NC (2015) Unique transposon landscapes are pervasive across *Drosophila melanogaster* genomes. *Nucleic Acids Res.* **43**, 10655-10672.
- Shi Q, Straus MR, Caron JJ, Wang H, Chung YS, Guarné A, Peters JE (2015) Conformational toggling controls target site choice for the heteromeric transposase element Tn7. *Nucleic Acids Res.* **43**, 10734-10745.
- Sun C, Wyngaard G, Walton DB, Wichman HA, Mueller RL (2014) Billions of basepairs of recently expanded, repetitive sequences are eliminated from the somatic genome during copepod development. *BMC Genomics* **15**, 186.
- Topilina NI, Novikova O, Stanger M, Banavali NK, Belfort M (2015) Post-translational environmental switch of RadA activity by extein-intein interactions in protein splicing. *Nucleic Acids Res.* **43**, 6631-6648.
- Topilina NI, Green CM, Jayachandran P, Kelley DS, Stanger MJ, Piazza CL, Nayak S, Belfort M (2015) SufB intein of *Mycobacterium tuberculosis* as a sensor for oxidative and nitrosative stresses. *Proc Natl Acad Sci USA* **112**, 10348-10353.
- Venner S, Feschotte C, Biémont C (2009) Dynamics of transposable elements: towards a community ecology of the genome. *Trends Genet.* **25**, 317-323.
- Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ (2015) Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**, 339-342.

