MBL | Biological Discovery in Woods Hole

# Mobile Genetic Elements: In Silico, InVitro, In Vivo

September 3-5, 2015
Woods Hole, Massachusetts

Organizers:

Phoebe Rice (University of Chicago)

Irina Arkhipova (Marine Biological Laboratory)

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

All talks are held in the Speck Auditorium (Rowe Building).

### Thursday, September 3, 2015

**2:00 PM** Check-in at Swope Conference Center, registration

**4:30 PM** Pre-meeting tours, poster setup

**5:00 PM** Welcome mixer, posters (Meigs Room)

**6:00 PM** Dinner

Session 1: Phoebe Rice, Chair

**7:00 PM** Welcome: Phoebe Rice (U Chicago), Bill Reznikoff (MBL)

**7:05 PM** *Introns and inteins as adaptive mobile elements*
Keynote speaker: **Marlene Belfort** (SUNY Albany)

**7:50 PM** *Diversity of structure, stability, specificity, and activity of Group I intron-encoded homing endonucleases*
**Barry Stoddard** (Fred Hutchinson Cancer Research Center)

**8:10 PM** *Ty1 retrosomes are asymmetrically associated with spindle pole bodies*
**Joan Curcio** (Wadsworth Center, NY State Dept. of Health)

**8:30 PM** *Transposable elements at the roots of adaptive immunity*
Invited speaker: **Eugene Koonin** (NIH/NLM/NCBI)

**9:00 PM** Posters/Mixer (Meigs Room)

### Friday, September 4, 2015

**7:00 AM** Breakfast

Session 2: Bill Reznikoff, Chair

**8:30 AM** *The floating (pathogenicity) island: a fascinating, if unsavory, genomic dessert*
Invited speaker: **Richard Novick** (New York University Medical Center)

**9:00 AM** *The recombinase locus of SCCmec elements - what are those other genes for?*
Organizer/speaker: **Phoebe Rice** (University of Chicago)

**9:30 AM** *Understanding mobile genetic elements in human microbial ecology*
Invited speaker: **Jack Gilbert** (University of Chicago / Argonne National Lab)

**10:00 AM** *Conformational toggling controls target site choice for the heteromeric transposase element Tn7*
**Joe Peters** (Cornell University)

**10:20 AM** Coffee break

**10:50 AM** *How to build a gene transfer agent: lessons from Rhodobacter capsulatus and RcGTA*
Invited speaker: **Andrew Lang** (Memorial University)

**11:20 AM** *Origin and evolution of gene transfer agents in α-proteobacteria*
**Migun Shakya** (Dartmouth College)

**11:40 AM** *Using a conjugal system to investigate recA-independent horizontal transfer*
**Anthony Kingston** (New England Biolabs)

**12:00 PM** Lunch

Free time or MBL tours:

**12:45 PM** Marine Resources Center tour
**1:30 PM** W. M. Keck Ecological and Evolutionary Genetics Facility tour

Bioinformatics Workshop: Vladimir Kapitonov, Leader

**2:00 PM** *Computational analysis of transposable elements: basic know-how and pitfalls.* Followed by Q&A session
**Vladimir Kapitonov** (NCBI/NLM)

**3:00 PM** *Comparing low-divergent repeats from multiple species through de novo repeat construction from short sequence reads*
**Chong Chu** (University of Connecticut)

**3:20 PM** *Microbial analysis options in the BaseSpace® platform*
**Chandrasen Soans** (Illumina)

**3:40 PM** *Transposon landscapes (TL) diversity in Drosophila melanogaster cell cultures and fly strains*
**Nelson Lau** (Brandeis University)

**4:00 PM** Cold beverages

Session 3: Eugene Koonin, Chair

**4:20 PM** *Giant reverse transcriptase-encoding transposable elements at telomeres*
Organizer/speaker: **Irina Arkhipova** (MBL)

**4:50 PM** *Tyrosons, a novel group of gigantic DNA transposons in fungi*
Invited speaker: **Vladimir Kapitonov** (NIH/NCBI)

**5:20 PM** *The L1 antisense promoter drives transcription of alternative transcripts in many human tissues*
**Steven Criscione** (Brown University)

**5:40 PM** *Intein clustering suggests functional importance in different domains of life*
**Olga Novikova** (SUNY Albany)

**6:00 PM** Dinner

Session 4: Joan Curcio, Chair

**7:00 PM** *Studies of a Human Retrotransposon*
Invited speaker: **John Moran** (University of Michigan)

**7:30 PM** *A multi-scale genome-wide analysis of the determinants of human and mouse endogenous retrovirus distributions*
Invited speaker: **Kateryna Makova** (Penn State University)

**8:00 PM** *Transposons regulate Wnt signaling to control germline stem cell differentiation in Drosophila*
**Prashanth Rangan** (SUNY Albany)

**8:20 PM** Repetitive DNA dynamics and hybrid incompatibilities
Invited speaker: **Daniel Barbash** (Cornell)

**8:50 PM** Recombination-independent recognition of DNA homology
**Eugene Gladyshev** (Harvard)

**9:10 PM** Posters/Mixer (Meigs Room)

**Saturday, September 5, 2015**

**7:00 AM** Breakfast

Session 5: Irina Arkhipova, Chair

**8:30 AM** *IntDOT: an unconventional tyrosine recombinase*
Invited speaker: **Jeff Gardner** (University of Illinois at Urbana-Champaign)

**9:00 AM** *DNA transposition in archaea*
Invited speaker: **Alison Hickman** (NIH/NIDDK)

**9:30 AM** *Elevated Ty1 retrotransposition in aging yeast mother cells*
**Patrick Maxwell** (Rensselaer Polytechnic Institute)

**9:50 AM** *Amplification of LTR-retrotransposons in contagious cancers of bivalves*
**Michael Metzger** (Columbia University)

**10:00 AM** Check-out

**10:10 AM** Coffee break

**10:30 AM** *Programmed elimination of TEs during chromatin diminution in the zooplankton Mesocyclops edax (Crustacea: Copepoda)*
Invited speaker: **Grace Wyngaard** (James Madison University)

**11:00 AM** *Multiple mechanisms contribute to telomere maintenance*
**Tammy Morrish** (University of Toledo)

**11:20 AM** *Activation of retrotransposons contributes to neurodegeneration in a Drosophila TDP-43 model of amyotrophic lateral sclerosis*
**Lisa Krug** (Cold Spring Harbor Laboratory)

**11:40 AM** *A damage independent role for 53BP1 that impacts chromatin organization of Igh and retrotransposons*
**Pedro Rocha** (New York University School of Medicine)

**12:00 PM** Concluding remarks: Irina Arkhipova

**12:05 PM** Lunch

**1:00 PM** End of meeting

# List of Registrants

| Last Name | First Name | Company Name | E-Mail Address | Position |
|---|---|---|---|---|
| Arkhipova | Irina | MBL | iarkhipova@mbl.edu | P.I. |
| Bajracharya | Urshula | Dartmouth College | migun.shakya@dartmouth.edu | Staff/ Other |
| Barbash | Daniel | Cornell University | barbash@cornell.edu | Faculty |
| Belfort | Marlene | UAlbany, SUNY | mbelfort@albany.edu | P.I. |
| Bilto | Iman | University of Manitoba | biltoi@myumanitoba.ca | Grad Student |
| Bradic | Martina | NYU | mb3188@nyu.edu | Post Doc |
| Chang | Yung-Heng | Cold Spring Harbor Laboratory | foreverchang89@gmail.com | Post Doc |
| Chu | Chong | University of Connecticut | chong.chu@engr.uconn.edu | Grad Student |
| Criscione | Steven | Brown University | steven_criscione@brown.edu | Grad Student |
| Curcio | Joan | Wadsworth Center | joan.curcio@health.ny.gov | P.I. |
| Deraspe | Maxime | Université Laval | maxime.deraspe.1@ulaval.ca | Grad Student |
| Dong | Xiaolong | University at Albany-SUNY | xdong@albany.edu | Grad Student |
| Dubnau | Josh | Cold Spring Harbor Lab | dubnau@cshl.edu | Faculty |
| Elliott | Tyler | University of Guelph | telliott@uoguelph.ca | Grad Student |
| Ernst | Susan | Tufts University | susan.ernst@tufts.edu | P.I. |
| Gardner | Jeff | University of Illinois | jeffgard@uiuc.edu | P.I. |
| Gilbert | Jack | Argonne National Laboratory | gilbertjack@gmail.com | P.I. |
| Gladyshev | Eugene | Harvard University | eugene.gladyshev@gmail.com | Post Doc |
| Govindaraju | Aruna | University of Texas Arlington | aruna@uta.edu | Grad Student |
| Hammell | Molly | Cold Spring Harbor Lab | mhammell@cshl.edu | P.I. |
| Hickman | Alison | NIH/NIDDK | alisonh@helix.nih.gov | Faculty |
| Jacobs | Jake | Rutgers University | jjsegs@gmail.com | Grad Student |
| Jin | Ying | Cold Spring Harbor Laboratory | yjin@cshl.edu | Staff/ Other |
| Kapitonov | Vladimir | NIH/NCBI | kvladimirv@gmail.com | Staff/ Other |
| Keeney | Jill | Juniata College | keeney@juniata.edu | Faculty |
| Kelley | Danielle | SUNY Albany | dkelley@albany.edu | Grad Student |
| Kingston | Tony | New England Biolabs | kingston@neb.com | Post Doc |
| Koonin | Eugene | NCBI/NIH | koonin@ncbi.nlm.nih.gov | P.I. |
| Krug | Lisa | Cold Spring Harbor Laboratory | lkrug@cshl.edu | Grad Student |
| Lang | Andrew | Memorial University | aslang@mun.ca | Faculty |
| Lau | Nelson | Brandeis University | nlau@brandeis.edu | P.I. |
| Lennon | Christopher | SUNY-Albany | cwlennon@albany.edu | Post Doc |
| Liao | Wen-Wei | Cold Spring Harbor Laboratory | gattacaliao@gmail.com | Staff/ Other |
| Makova | Kateryna | Penn State University | makovakateryna@gmail.com | Faculty |
| Maxwell | Patrick | Rensselaer Polytechnic Institute | maxwep2@rpi.edu | Faculty |
| McGurk | Michael | Cornell University | mpm289@cornell.edu | Grad Student |
| Meselson | Matthew | MBL/Harvard University | msmeselson@gmail.com | Faculty |
| Metzger | Michael | Columbia University | metzgerm@uw.edu | Post Doc |
| Mika | Katelyn | University of Chicago | kmmika@uchicago.edu | Grad Student |
| Mir Sanchis | Ignacio | University of Chicago | imirsanchis@uchicago.edu | Post Doc |
| Moran | John | University of Michigan | moranj@umich.edu | Faculty |

| | | | | |
|---|---|---|---|---|
| Mozzherin | Dmitry | MBL | dmozzherin@mbl.edu | Staff/ Other |
| Murphy | S. Patrick | NY School of Public Health | ppmurphy@albany.edu | Grad Student |
| Naik | Ankana | University at Albany | anaik@albany.edu | Grad Student |
| Nekrutenko | Anton | Penn State University | anton@nekrut.org | Faculty |
| Neretti | Nicola | Brown University | nicola_neretti@brown.edu | P.I. |
| Novick | Richard | NYU School of Medicine | richard.novick@med.nyu.edu | Faculty |
| Novikova | Olga | University at Albany | onovikova@albany.edu | Post Doc |
| Peifer | Andrew | Rensselaer Polytechnic Institute | peifea@rpi.edu | Grad Student |
| Peters | Claire | Cornell University | joe.ithaca@gmail.com | Staff/ Other |
| Peters | Joseph | Cornell University | jep48@cornell.edu | P.I. |
| Qu | Guosheng | SUNY at Albany | gqu@albany.edu | Post Doc |
| Rai | Sudhir | NICHD | raisk@mail.nih.gov | Post Doc |
| Raleigh | Elisabeth | New England Biolabs | raleigh@neb.com | Staff/ Other |
| Rangan | Prashanth | University at Albany/RNA Inst. | prangan@albany.edu | P.I. |
| Raviram | Ramya | NYU Sackler | ravirr01@nyumc.org | Grad Student |
| Reznikoff | Bill | MBL | breznikoff@mbl.edu | Staff/ Other |
| Rice | Phoebe | University of Chicago | price@uchicago.edu | P.I. |
| Rocha | Pedro | NYU | pereip02@nyumc.org | Post Doc |
| Rodriguez | Fernando | Marine Biological Laboratory | frodriguez@mbl.edu | Post Doc |
| Roman | Christina | University of Chicago | chroman@uchicago.edu | Grad Student |
| Rosett | Jane | MBL | jrosett@mbl.edu | Staff/ Other |
| Roy | Paul | Université Laval | paul.roy@crchul.ulaval.ca | Faculty |
| Rozhkov | Nikolay | Cold Spring Harbor Laboratory | rozhkov@cshl.edu | Post Doc |
| Russell | Shelbi | Harvard University | shelbirussell@fas.harvard.edu | Grad Student |
| Saylor | Brent | University of Guelph | bsaylor@uoguelph.ca | Grad Student |
| Scales | Jessica | Juniata College | Scalejl11@gmail.com | Undergrad Student |
| Shakya | Migun | Dartmouth College | migun.shakya@dartmouth.edu | Post Doc |
| Shih | Meng-Fu | Cold Spring Harbor Laboratory | mshih@cshl.edu | Post Doc |
| Soans | Chandrasen | Illumina, Inc. | csoans@illumina.com | Staff/ Other |
| Soucy | Shannon | University of Connecticut | shannon.soucy@uconn.edu | Grad Student |
| Stoddard | Barry | Fred Hutchinson Cancer Res Ctr | bstoddar@fredhutch.org | Faculty |
| Waldern | Justin | SUNY Albany | jmwaldern@gmail.com | Grad Student |
| Wyngaard | Grace | James Madison University | wyngaaga@jmu.edu | Faculty |
| Yu | Yang | Cold Spring Harbor Laboratory | yyu@cshl.edu | Post Doc |
| Yushenova | Irina | Marine Biological Laboratory | iyushenova@mbl.edu | Post Doc |
| Zagoskin | Maxim | James Madison University | zagoskinmv@gmail.com | Post Doc |

We thank University of Chicago, Illumina Inc., and Integrated DNA Technologies for sponsorship and support.

# INTRONS AND INTEINS AS ADAPTIVE MOBILE ELEMENTS

Marlene Belfort, UAlbany SUNY

Our lab studies introns and inteins that self-splice at the RNA and protein levels, respectively. They are also mobile at the DNA level, exist in all three domains of life, and are widely considered to be selfish genetic elements. I will present several examples of how, contrary to current dogma, mobile introns and inteins can sense stress and be potentially useful to the hosts they parasitize. I will also present a 3.85 Å cryo-EM structure of a bacterial group II intron, which encodes a reverse transcriptase, and functionally resembles both eukaryotic retrotransposons and splicesomal introns. The intron ribonucleoprotein comprises a large ribozyme and intron-encoded protein with maturase, reverse transcriptase and endonuclease activities. Remarkably, the reverse trancriptases's catalytic domain is related to telomerase, which preserves eukaryotic chromosome ends, whereas the maturase domain resembles the spliceosomal Prp8 protein. These remarkable similarities hint at complex ancestral relationships and provide new insights into splicing and retromobility.

# Diversity of structure, stability, specificity, and activity of Group I intron-ended homing endonucleases

Barry L. Stoddard[1,*]

[1] Division of Basic Sciences, Fred Hutchinson Cancer Research Center 1100 Fairview Ave. N. Seattle WA 98109

* Contact Information: bstoddar@fredhutch.org; 206-667-4031

Mobile group I introns, found throughout all forms of microbial life, encode a variety of homing endonucleases, or 'meganucleases', that drive their mobility and persistence via a mechanism that is dependent on their site-specific DNA cleavage properties. Homing endonucleases of the LAGLIDADG protein family, which are used as engineerable gene-targeting nucleases, are encoded within archaeal, mitochondrial and chloroplast genomes, and are notable for exceptional conservation of tertiary structure that nevertheless spans a wide range of folding behaviors and stabilities, DNA targeting specificities, and functional activity. In this presentation the variation in structure/function relationships observed across a panel of nine separate mitochondrial LAGLIDADG homing endonucleases, that target completely different genetic loci in their respective hosts, and the impact of those properties on the engineerability and application of those enzymes, is described. The properties of these nucleases appear to be largely driven by several simultaneous selection forces during their evolution and divergence: the need to tailor their DNA specificity profiles to match the underlying constraints across the coding sequences spanning their target sites, the need to maintain sufficient specificity to avoid toxicity, balanced against the need to exploit opportunities for ectopic transfer to new host DNA sites, and the effect of continuous degradation of form and function that leads to a cycle of invasion, loss of activity, eventual removal from the genome, and subsequent reinvasion.

**Ty1 retrosomes are asymmetrically associated with spindle pole bodies**

M. Joan Curcio[1,2], Patrick Murphy[1,2], Alexey Khodjakov[1] and Sheila Lutz[1]

[1]Wadsworth Center, New York State Department of Health, and [2]Department of Biomedical Sciences, University at Albany-SUNY, Albany, NY 12201

The RNA and protein products of long terminal repeat (LTR)-retrotransposons assemble into nucleocapsids, wherein the RNA is reverse-transcribed into cDNA that is transported from the cytoplasm to the nucleus and integrated into the host genome. This intricate intracellular mechanism of replication, also used by endogenous retroviruses to populate mammalian genomes, relies heavily on the host. We are using the Ty1 retrotransposon of budding yeast to characterize nucleocapsid assembly sites and determine how the assembly site locale impacts retrotransposition and host cell function. Ty1 nucleocapsids assemble in a cytoplasmic focus, or retrosome. Retrosomes form as a consequence of Signal Recognition Particle targeting Ty1 RNA translation complexes to the endoplasmic reticulum (ER) and mediating translocation of nascent Gag to the ER lumen. Gag is transported back to the cytoplasm by an unknown mechanism, binds Ty1 RNA translation complexes and multimerizes to nucleate the retrosome. FRAP experiments reveal that Gag in retrosomes is relatively immobile. Using time-lapse video microscopy of cells expressing Gag:GFP, we have shown that the retrosome remains in mother cells throughout multiple cell divisions and is associated with the nuclear membrane-embedded spindle pole body (SPB), a functional equivalent of the centrosome. SPB structural components and duplication factors promote Ty1 retrotransposition. In cells lacking a retrosome, SPBs are asymmetrically segregated such that the new SPB is retained in the mother cell and the old SPB segregates to the bud. In contrast, retrosome association specifically with the old SPB causes the old SPB to be retained in mother cells, thus reversing the asymmetric segregation of SPBs. Since asymmetric centrosome inheritance is a feature of stem cell self-renewal, we suggest that the nucleation of retrotransposon or endogenous retroviral assembly sites at a specific centrosome could negatively impact development and fertility in higher eukaryotes.

# Transposable elements at the roots of adaptive immunity

Eugene V. Koonin

*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA*

Adaptive immunity in both prokaryotes and eukaryotes involves genomic rearrangements that enable imprinting of information on encountered pathogens. Multiple lines of recent evidence point to the key role of transposable elements in the evolution of adaptive immunity. In particular, the adaptation module of the CRISPR-Cas systems of adaptive immunity in archaea and bacteria, that is responsible for the immune memory, apparently evolved from a distinct class of transposons denoted casposons. The predicted recombinase of these transposons gave rise to Cas1, the key endonuclease that is responsible for the integration of foreign DNA spacers into CRISPR arrays. The V(D)J recombination in vertebrates that generates immungolbulin diversity evolved via a parallel route from an unrelated group of transposons that encode the RAG1 recombinase and the ancillary RAG2 protein. A third independent case of evolution of a defense system from transposable elements involves DNA elimination during macronucleus maturation in ciliates that exploits a PiggyMac transposon. The adaptation module is not the only contribution of transposable elements to the evolution of the CRISPR-Cas systems. In addition, in Class 2 CRISPR-Cas systems that have become the basis of the new generation of genome editing methods the effector complexes involved in the destruction of foreign DNA are derivatives of transposon-encoded nucleases. Accordingly, this class of CRISPR-Cas systems evolved entirely from mobile elements, on multiple, independent occasions. Focused search of bacterial and archaeal genomes for transposon domestication led to the discovery of new types of such immune systems. Generally, the same properties of transposable elements that underlie their mobility and hence enable genomic parasitism provide both naturally evolved components of adaptive immunity systems and tools for experimental genome manipulation.

References

Krupovic M, Makarova KS, Forterre P, Prangishvili D, Koonin EV. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. BMC Biol. 2014 May 19;12:36

Koonin, E.V., Krupovic, M. (2015) Evolution of adaptive immunity from transposable elements combined with innate immune systems. Nat Rev Genet. **16**(3):184-92

Koonin, E. V., and Krupovic, M. (2015). A Movable Defense. *The Scientist*(January 1) http://www.the-scientist.com/?articles.view/articleNo/41702/title/A-Movable-Defense/.

# The floating (pathogenicity) island: a fascinating, if unsavory, genomic dessert

Richard P. Novick, MD, Departments of Medicine and Microbiology, NYU School of Medicine, New York, NY, 10016, USA

**Abstract**

Bacteria transfer their genes by an astonishing variety of strategies, many of which involve chromosomal islands – inserted non-essential DNA segments. Although it is vanishingly unlikely for a primordial DNA island to have evolved de novo in its extant host, there are 4 classes of transfer-proficient islands that appear to have evolved from pre-existing mobile elements, 2 from prophages – SaPIs and GTAs - and two from conjugative plasmids – Integrative and conjugative elements (ICEs) and conjugative transposons (CTs). The latter excise and transfer by classical conjugation; the SaPIs are packaged in phage-like infectious particles which are released upon phage induced lysis; the GTAs are not themselves transferred but generate tiny phage-like particles that contain random 4-5 kb fragments of host DNA and require the DNA uptake (*com*) pathway. The ICEs and CTs carry a wide variety of genes that impact the host, especially with reference to pathogenesis; the SaPIs carry fewer, but also participate in transduction of unlinked genes; the GTA particles choose only randomly.

This presentation is focused on the SaPIs, which we identified and characterized some time ago. We propose that they diverged from an ancestral prophage or protophage in the very distant past and established a unique lineage that has evolved to remain distinct from its prophage progenitor. SaPIs have not only developed several characteristic and distinguishing features of their own, but also have acquired genes for superantigens and other host-specific factors, and are responsible for the spread of these.

Biology of the SaPIs is summarized and the SaPIs are compared to the other types of island-based gene transfer systems.

**The recombinase locus of SCC elements - what are those other genes for?**

Phoebe A. Rice, Dept. of Biochemistry & Molecular Biology, the University of Chicago

The SCCs are a family of genomic islands found primarily in staphylococci, that, despite carrying the methicillin resistance that creates MRSA strains, have been subject to surprisingly little mechanistic study. SCC elements are highly mosaic but are defined by a specific insertion site in the host chromosome and a conserved recombinase gene complex composed of 1 or 2 site-specific recombinases from the large serine family and a few surrounding genes whose functions were mysterious.

We began by studying the activity of the recombinases, and found that when two are present, they collaborate to recognize the asymmetric host insertion site. This observation may explain why some other, unrelated mobile elements encoded multiple site-specific recombinases.

We also examined the genes surrounding the recombinases, and found highly conserved ORFs whose sequences suggest that they are involved in DNA replication, even though the SCC elements have been referred to as non-replicative. Finally, we have determined the crystal structure of one, Cch, that is related to replication initiator proteins from the SaPIs. We find that it is a hexameric, ring-shaped helicase that has structural homology to the MCM proteins within its ATPase domain and that unwinds DNA with 3' to 5' polarity.

**Understanding mobile genetic elements in human microbial ecology**

Jack A. Gilbert

Argonne National Laboratory and The University of Chicago

The human body contains 2-3 lbs of bacteria that form complex ecosystems capable of being resilient and at the same time being dynamic and adaptable to the changing environment. In the human gut, this changing environment mainly comprises shifts in diet and medicine. The role of mobile genetic elements, mediated through horizontal gene transfer in supporting the stability and adaptability of the human microbiome remains poorly characterized. However developing models that describe and predict how gene mobility influences the metabolism and ecological response of the gut microbiome will be invaluable for synthetic ecology and engineering of the gut environment. This is especially important to consider when developing new therapies and drugs, as mobile genetic elements allow microbial assemblages to rapidly adapt to new drugs. The development of antibiotic resistance is one example of the role of horizontal gene transfer in enable the microbiome to stabilize in the face of perturbation. I will discuss what is understood about the role of mobile genetic elements within the microbiome and between microbiomes within human populations, and how this is manifest in ecological stability and the potential for new therapies.

# Conformational toggling controls target site choice for the heteromeric transposase element Tn7

Qiaojuan Shi[1], Marco R. Straus[1], Jeremy J. Caron[2], Huasheng Wang[2], Yu Seon Chung[2], Alba Guarné[2] and Joseph E. Peters[1]

[1]Department of Microbiology, Cornell University, Ithaca, New York, 14853, USA
[2]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, L8S 4K1, Canada.

## ABSTRACT

The bacterial transposon Tn7 facilitates horizontal transfer by directing transposition into actively replicating DNA with the element-encoded protein TnsE. Structural analysis of the C-terminal domain of wild type TnsE identified a novel protein fold including a central V-shaped loop that toggles between two distinct conformations. The structure of a robust TnsE gain-of-activity variant has this loop locked in a single conformation, suggesting that conformational flexibility regulates TnsE activity. Structure-based analysis of a series of TnsE mutants relates transposition activity to DNA binding stability. Wild type TnsE appears to naturally form an unstable complex with a target DNA, whereas mutant combinations required for large changes in transposition frequency and targeting stabilized this interaction. Collectively, our work unveils a unique structural proofreading mechanism where toggling between two conformations regulates target commitment by limiting the stability of target DNA engagement until an appropriate insertion site is identified.

How to build a gene transfer agent: lessons from *Rhodobacter capsulatus* and RcGTA

Andrew Lang (Memorial University)

Gene transfer agents (GTAs) are virus-like particles that transfer small pieces of genomic DNA to other cells. A diverse collection of species, including Bacteria and Archaea, are currently known to produce GTAs. *Rhodobacter capsulatus* is a purple non-sulfur alphaproteobacterium that produces the gene transfer agent RcGTA. A 15-kb cluster of genes, which is organized similar to the structural gene regions of many tailed bacteriophages, encodes many of the RcGTA structural proteins. Expression of this gene cluster is affected by multiple cellular regulatory systems, including phosphorelay, quorum sensing, and partner-switching pathways. Comparative transcriptomics was used to study the regulation of expression of the RcGTA gene cluster and to identify co-regulated genes. This identified nine genes at six loci in the *R. capsulatus* genome, outside of the RcGTA gene cluster, as consistently co-regulated with the cluster. The involvement of these genes in RcGTA production was tested using a combination of genetic and biochemical approaches, demonstrating that four additional loci contribute to production of fully functional RcGTA particles and their release from cells. The evolutionary histories and distribution of these additional GTA genes in other species that contain RcGTA gene cluster equivalents provide further insight into the evolution of RcGTA and related elements.

**ORIGIN AND EVOLUTION OF GENE TRANSFER AGENTS in α-PROTEOBACTERIA**

Migun Shakya[1], Daniel P. Birnbaum[1,2], Taylor B. Neely[1,3], and Olga Zhaxybayeva[1,3]
[1]Department of Biological Sciences, Dartmouth College, Hanover, NH 03755
[2]Present address: Broad Institute, Cambridge, MA 02142
[3]Department of Computer Science, Dartmouth College, Hanover, NH 03755
(migun.shakya@dartmouth.edu)

Genomes of most sequenced prokaryotes have at least one integrated phage in their genome (termed *prophage*) [1]. Some of these prophages are active – when triggered, they synthesize viral particles, package their DNA, lyse the host cells, and are released – while others are defunct or "domesticated" by the host [2]. Likely from the latter category, Gene Transfer Agents (GTAs) are phage-like entities encoded within a host genome, which when triggered, synthesize viral particles and are released through lysis of host cells. However, unlike phages, GTAs (*i*) are controlled by the host and environmental factors, (*ii*) package seemingly random pieces of host DNA, (*iii*) are too small to incorporate all of the genes necessary for their production, and (*iv*) are only produced by subset (0.1-3%) of cells in a population (reviewed in [3]). These properties raise many questions about GTA origin and evolution. *When and how did GTAs originate in an evolutionary history of a host lineage? How did they spread across a lineage? Why are GTAs being maintained in a genome?*

To date, genetically unrelated GTAs have been found only in four divergent taxonomic groups of prokaryotes: bacterial classes of α-proteobacteria, δ-proteobacteria and Spirochaetia, and archaeal class Methanococci. Here, we trace origin and evolution of GTAs using *RcGTA* - a GTA system hosted by α-proteobacterium *Rhodobacter capsulatus* - and its homologs in α-proteobacteria

Due to the indisputable homology between *bona fide* GTA genes and their viral counterparts, some of these homologs may simply represent prophages because traditional sequence similarity search methods like BLAST are incapable of differentiating between the two types. To address this problem, we are also developing methods for classification of virus-like sequences in a prokaryotic genome as either "GTA" or "prophage".

1. Casjens, S. (2003) *Mol Microbiol*, *49*, 277-300.
2. Bobay, L. M., Touchon, M., and Rocha, E. P. (2014) *Proc Natl Acad Sci U S A*, *111*, 12127-32.
3. Lang, A. S., Zhaxybayeva, O., and Beatty, J. T. (2012) *Nat Rev Micro*, *10*, 472-482.

Title: Using a conjugal system to investigate recA-independent horizontal transfer

Authors: Anthony Kingston, Christine Ponkratz, and Elisabeth Raleigh

Author Affiliations: New England Biolabs

In bacteria, mechanisms that incorporate DNA into a genome without strand-transfer proteins such as RecA play a major role in generating novelty by horizontal gene transfer. Using conjugal chromosome transfer between *E. coli* K-12 strains, we identified an unexpected RecA-independent recombination mechanism in which a large segment of the recipient chromosome is replaced with genomic DNA from the donor. This process occurs at low frequency ($\sim 10^{-10}$/recipient/h) and a single event can exchange over two megabases of continuous DNA.

The *E. coli* K-12 genome encodes paralogs belonging to the uncharacterized YhgA-like protein family (YjiP, YfcI, YadD, YhgA and YfaD). When overexpressed in the recipient, all but the last of these increased the frequency of RecA-independent recombination events, were toxic to the cell, and induced expression of a reporter of DNA damage. Mutagenesis of conserved residues in YhgA abolished these effects, as did removal of the variable C-terminus.

# Comparing low-divergent repeats from multiple species through de novo repeat construction from short sequence reads

Chong Chu[1], Rasmus Nielsen[2], Yufeng Wu[1]
[1]Department of Computer Science & Engineering
University of Connecticut
[2]Department of Integrative Biology
University of California, Berkeley

Repeat elements are important components of eukaryotic genomes. Sequencing data from many species are now available, providing opportunities for finding and comparing genomic repeat activity among species. One limitation in our understanding of repeat elements is that most analyses rely on reference genomes that are incomplete and often are missing data in highly repetitive regions that are difficult to assembly. To overcome this problem we develop a new method, REPdenovo, which assembles repeat sequences directly from raw shotgun sequencing data. We show that REPdenovo is substantially better than existing methods both in terms of the number and the completeness of the repeat sequences that it recovers. We apply the method to human data and discover a number of new repeats sequences that have been missed by previous repeat annotations. By aligning these discovered repeats to Pacbio long reads, we confirm the existence of these novel repeats. REPdenovo is a new powerful computational tool for annotating genomes and for addressing questions regarding the evolution of repeat families.

**Microbial Analysis Options in the BaseSpace® Platform**

Soans, C., Teiling, C.;  Illumina, Inc., San Diego, CA

**Abstract**: BaseSpace is Illumina's next-generation sequencing cloud computing environment. With cloud computing, users can access a web portal to retrieve, update and analyze their data with a variety of applications. This study is a demonstration of some of the BaseSpace capabilities on a set of data available in NCBI.  A metagenomics data set was imported into BaseSpace from NCBI's Sequence Read Archive (SRA) with a push-button informatics App called "SRA Import." The data set was then analyzed using the 16S Metagenomics App and the Kraken Metagenomics App, both of which are available in BaseSpace. In order to demonstrate the performance of both these applications, we tested them on data described in a recent publication by Jorth P. et al. (2014) in which patient-matched healthy and diseased samples were compared by 16S rRNA sequencing on the Illumina MiSeq instrument. Hierarchical clustering and Shannon species diversity index results indicated a clear difference between healthy and diseased samples, mirroring the results in the publication.  With this demonstrated workflow, researchers have access to a high performing, sensitive, and interactive tool for analyzing their metagenomics data in BaseSpace.  Now, researchers can easily and securely store, analyze, compare, manage, and share their data - all within the BaseSpace Environment.

**Transposon landscapes (TL) diversity in *Drosophila melanogaster* cell cultures and fly strains.**

Reazur Rahman, Yuliya Sytnikova, and Nelson Lau

ABSTRACT

Although transposons comprise major proportions of animal genomes, we still do not fully understand how transposons affect genome structure and gene expression. To address this question, we reveal a new tool to capture the diversity of transposon landscapes (TL) in different *Drosophila* strain genomes. This tool is called the Transposon Insertion and Depletion AnaLyzer (TIDAL), and it maps transposon landscapes for the plethora of *Drosophila* genomes available in the Sequencing Read Archives. TIDAL outputs transposon genome charts that map differences between fly strains and the *Drosophila* genome reference, as well as provides rich genomic annotations around each transposon insertion and depletion. We discover wide-spread TL diversity in *Drosophila* cell cultures, common lab fly strains, and isogenic lines in the *Drosophila* Genetic Reference panel, ranging from several hundreds to thousands of transposon InDels per sequenced line. Surprisingly, the majority of the transposon InDels in isogenic fly lines appears to be new somatic mobilization events, and reference wild-type lab strains that are common by name but kept in different labs also display different TLs. Finally, we have built a web-accessible database called TIDAL-Fly v1.0 where transposon landscape maps can be downloaded. TIDAL-Fly will enable geneticists to better understand how transposons impact fly strain genetic backgrounds and demonstrates that even in narrow populations of *Drosophila* there are highly diverse TLs.

**Giant reverse transcriptase-encoding transposable elements at telomeres**

Irina Arkhipova, Fernando Rodriguez, Irina Yushenova

*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, USA*

Transposable elements (TEs) are present in most genomes and have a profound impact on their structure, function and evolution. For many years, our knowledge of the structural and functional TE diversity has remained relatively stable, with the understanding that we have largely grasped the major principles of their structural organization and the underlying basis for their mobility. Here we identify a previously unknown type of eukaryotic retroelements, which do not fit any commonly recognized TE descriptors. As the principal polymerizing component, they contain the *Athena* reverse transcriptases (RT), which were previously identified in bdelloid rotifers and belong to the enigmatic group of intron-containing *Penelope*-like retroelements. These TEs of exceptionally complex structure, which we named *Terminons*, attach to the chromosome termini *via* stretches of telomeric repeats, elicit strong piRNA response, and can reach up to 40 kb in length. In addition to RTs, *Terminons* encode dozens of co-oriented open reading frames (ORFs), some of which code for enzymatic functions that could be assigned a role in transposition, while others resemble structural proteins and could participate in RNP formation. Additionally, some ORFs of enzymatic origin are characterized by precise loss of catalytic residues, while others employ programmed ribosomal frameshifting for expression. A variety of conserved *cis*-acting elements can be involved in different aspects of replication and terminal attachment. The extraordinary length and complexity of these TEs, as well as the high degree of inter-family variability in the ORF content, challenge the current views on structural organization of eukaryotic retroelements, and raise fundamental questions regarding their possible connections with the viral world and their impact on telomere biology.

## *Tyrosons*, a novel group of gigantic DNA transposons in fungi

Vladimir V. Kapitonov

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

Email: kvladimirv@gmail.com

I report here a novel group of transposable or mobile elements, called *Tyrosons*. It appears that *Tyrosons* emerged in a common ancestor of filamentous fungi, over 450 million years ago. Autonomous *Tyrosons* encode a universally conserved tyrosine integrase, which is distantly similar (detectable by PSI-BLAST) to a group of bacterial tyrosine integrases/recombinases and is not similar to tyrosine integrases encoded by known eukaryotic DNA transposons (*Crypton* superfamily) and retrotransposons (DIRS superfamily).

*Tyrosons* are remarkable and unique transposons due to several unusual features: 1) they are unprecedentedly long, some *Tyrosons* are longer than 200 kb; 2) they do not have conserved termini; 3) a typical *Tyroson* family is composed of only one or a few copies; 4) in each family, a *Tyroson* transposon codes for 20-30 genes, mostly acquired from their host genomes and often not similar to proteins encoded by *Tyrosons* from other families.

Intriguingly, *Tyrosons* may serve as engines for dissemination and diversification of proteins involved in self and non-self recognition (heterokaryon incompatibility) and important for adaptation of filamentous ascomycetes to their environment. For instance, numerous genes found in *Tyrosons* code for proteins containing the HET and NACHT domains, known as inducers of heterokaryon incompatibility.

Surprisingly, some of recently discovered islands of secondary metabolism genes, important for niche adaptation of ascomycetous fungi, are in fact *Tyroson* transposons.

**Title:**
The L1 antisense promoter drives transcription of alternative transcripts in many human tissues.

**Authors:**
Steven Criscione[1], Nicholas Theodosakis[2,3], Goran Micevic[2,3], Toby C. Cornish[4], Marcus Bosenberg[2,3], Kathleen H. Burns[4,5,6,7], Nicola Neretti[1], and Nemanja Rodić[2]

**Affiliations:**
[1]Department of Molecular Biology, Cell Biology, and Biochemistry, Brown University, Providence, RI 02912, USA
[2]Department of Pathology, Yale University, New Haven, CT 06510, USA
[3]Department of Dermatology, Division of Dermatopathology, Yale University, New Haven, CT 06510, USA
[4]Department of Pathology, [5]McKusick-Nathans Institute of Genetic Medicine, [6]Department of Molecular Biology & Genetics, [7]High Throughput (HiT) Biology Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA

**Co-corresponding authors:**
Nicola Neretti
Nicola_Neretti@brown.edu
Department of Molecular Biology, Cell Biology, and Biochemistry; Center for Computational Molecular Biology; Brown University, Providence, RI 02912, USA
Nemanja Rodić
Nemanja.Rodic@Yale.edu
Department of Dermatology, Division of Dermatopathology, Yale University, New Haven, CT 06510, USA

**Abstract:**
Long INterspersed Element-1 (LINE-1 or L1), is the only autonomously active transposable element family in humans. L1 sequences comprise ~17% of the human genome, but only the evolutionarily most recent human-specific subfamily is active, with the remainder representing mutated and truncated (fossilized) L1s that have lost the ability to jump. The L1 promoter has a peculiar bidirectional orientation. The L1 sense promoter drives transcription of the two L1-encoded proteins, which are in turn required for retrotransposition. While aberrant L1 protein-1 expression is a feature of nearly a half of all human malignancies, little is known at the genomic scale about the L1 antisense promoter activity on gene regulation. To investigate this question, we analyzed the GenBank ESTs for transcripts that are likely to initiate in the L1 antisense promoter. We identified 1330 putative alternative L1 antisense transcripts, 1262 of which have not been previously reported. We found that most of these are chimeric L1 antisense transcripts that originate in the L1 antisense promoter and are spliced to the nearby exons of cognate neighboring genes. We show the majority of chimeric transcripts are sense-oriented with respect to cognate genes and have a marked tendency to be initiated near the transcriptional start site of the adjoining genes. Enrichment of activation marks and transcription factors at the promoter of many chimeric transcripts gives rise to the intriguing possibility that they are actively transcribed and might have a biological, yet unknown, function. We empirically validate a subset of the chimeric transcripts identified by our EST screen and show they are both readily detectable in many normal human tissues and often expressed at levels comparable to those of the corresponding wild-type transcripts. Our findings reveal emerging new classes of alternative L1 antisense transcripts, which may, in aggregate, impact as many as ~5.7% of all human genes.

# Intein clustering suggests functional importance in different domains of life

Olga Novikova, Pradeepa Jayachandran, Natalya I. Topilina, Marlene Belfort

Department of Biological Sciences and RNA Institute, University at Albany, 1400 Washington Avenue, Albany, NY 12222

Inteins, also called protein introns, are self-splicing mobile elements found in all domains of life. A bioinformatic survey of genomic data highlights a biased distribution of inteins among functional categories of proteins in both bacteria and archaea, with a strong preference for a single network of functions containing replisome proteins. Many non-orthologous, functionally equivalent replicative proteins in bacteria and archaea carry inteins, indicating a selective retention of inteins in proteins of particular functions across domains of life. Inteins cluster not only in proteins with related roles, but also in specific functional units of those proteins, like ATPase domains. This peculiar bias does not fit the models describing inteins exclusively as parasitic elements. In such models, evolutionary dynamics of inteins is viewed primarily through their mobility with the intein homing endonuclease (HEN) as the major factor of intein acquisition and loss. Although the HEN is essential for intein invasion and spread in populations, HEN dynamics does not explain the observed biased distribution of inteins among proteins in specific functional categories. We propose that the protein splicing domain of the intein can act as an environmental sensor that adapts to a particular niche and could potentially increase the chance of the intein becoming fixed in a population. We argue that selective retention of inteins might be beneficial under certain environmental stresses, to act as panic buttons that reversibly inhibit specific networks, consistent with sporadic intein distribution.

# A 3' poly(A) tract is required for LINE-1 retrotransposition

Aurélien J. Doucet[1], Jeremy E. Wilusz[4], Tomoichiro Miyoshi[1], Ying Liu[3], and John V. Moran[1,2,3]

[1]Department of Human Genetics, [2]Department of Internal Medicine, [3]Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, MI 48109-5618, USA. [4]Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA.

The average human genome contains approximately 80-100 active Long INterspersed Element-1 (LINE-1 or L1) retrotransposons, and the mobility (*i.e.*, retrotransposition) of a cohort of these elements (termed 'hot' L1s) continues to contribute to both inter- and intra-individual genetic diversity. 'Hot' L1s encode two proteins (ORF1p and ORF2p) that are required for retrotransposition. Using genetic, molecular biological, and biochemical approaches, we previously demonstrated that ORF1p and ORF2p preferentially associate with their encoding RNA to mediate its retrotransposition. Here, I will present unpublished data (Doucet *et al.*, 2015, in revision), which demonstrates that a 3' poly(A) tract is a critical determinant in allowing ORF2p to mobilize L1 RNA, as well as RNAs encoded by non-autonomous retrotransposons, to new genomic locations. In sum, our findings indicate that a 3' poly(A) tract is required for the retrotransposition of sequences that account for approximately one billion base pairs of human genomic DNA.

**A multi-scale genome-wide analysis of the determinants of human and mouse endogenous retrovirus distributions**

Kateryna Makova

Department of Biology, Pennsylvania State University, University Park, PA 16802

Endogenous retroviruses (ERV) are remnants of retroviral infections introduced in the germ-line. In most species, these transposable elements are inactive due to sequence degradation or cellular mechanisms that silence them, i.e methylation. HERV-Ks are the only active and polymorphic ERVs among diverse human populations. In contrast, the mouse genome contains two of the most active mammalian ERVs – IAPs and ETns – producing almost 10% of all spontaneous mutations in the germ-line of this species. Here, we investigated integration and fixation preferences of 1866 ETns, 5950 IAPs, 872 fixed HERV-Ks, and integration preferences of 1208 *ex vivo* HERV-Ks, by examining the flanking regions (32 kb upstream and 32 kb downstream) surrounding their integration sites. We fragmented these flanking regions into 1-kb windows and collected data for 45 human and 41 mouse genomic features (e.g. features associated with microsatellites, DNA conformation, genes, other transposable elements, replication, recombination, and epigenetic modifications) for each of them. We contrasted the genomic features in the flanking sequences of ERVs vs. control regions, using Functional Data Analysis (FDA) techniques and in particular, the Interval Testing Procedure (ITP). These depletion and enrichment results reflect the ability to study flanking regions at multiple scales and considering localization – as allowed by the use of FDA methodology.

# Transposons regulate Wnt signaling to control germline stem cell differentiation in *Drosophila*

**Authors:** Maitreyi Upadhyay[1], Yesenia Martino Cortez [1,2], Leticia Tevares[1,3], SiuWah Wong-Deyrup[1], Sean Schowalter[1,4], Pooja Flora[1], Corrine Hill[1,5], Mohammed Ali Nasrallah[1], Sridar Chittur[1,6] and Prashanth Rangan[1,7]

**Affiliations:** [1]Department of Biological Sciences/RNA Institute, University at Albany SUNY, Albany, NY 12222
[2]NYU Langone Medical Center, MSB 495, NY 10016
[3]400 Trabalhador Sao Carlense Avenue, Sao Paulo, Brazil 13566
[4]Boston University School of Medicine, 815 Albany Street, MA 02119, USA
[5]Department of Environmental Health Sciences, University of Massachusetts Amherst, Amherst, MA 01003
[6] CFG Core Facility, University at Albany SUNY, Rensselaer NY 12144

## Abstract

Intrinsic and extrinsic signals regulate germline stem cell (GSC) self-renewal and differentiation for the sustained production of gametes, which is critical for reproductive success. In *Drosophila*, extrinsic Decapentaplegic (Dpp, a TGF-β homolog) signaling from the somatic niche promotes self-renewal, however, the niche signal that induces differentiation remains unknown. dSETDB1, a histone methyltransferase that catalyzes trimethylation of Histone 3 lysine 9 (H3K9me3) and heterochromatin formation, is required non-autonomously for GSC differentiation. dSETDB1 mediated heterochromatin formation represses transposable elements (TEs) by activating the production of piwi interacting RNAs (piRNA). Here, we show that *dSETDB1* is required for the expression of a Wnt ligand, *Drosophila* Wingless type mouse mammary virus integration site number 4 (dWnt4) in the somatic niche cells. This dWnt4 signaling acts on the somatic niche cells to facilitate the formation of cytoplasmic processes that envelop the germ line, which acts as a differentiation cue. Surprisingly, we find that mutations independent of *dSETDB1* that up-regulate TEs, down-regulate *dWnt4.* This suggests that *dWnt4* expression is exquisitely sensitive to the presence of TEs. We propose that dWnt4 signaling in the gonad could act as a checkpoint to prevent TE damaged eggs from being propagated.

**Repetitive DNA dynamics and hybrid incompatibilities**

Daniel A Barbash, P. Satyaki, Tawny N Cuykendall, Kevin H-C Wei and Michael M McGurk.
*Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA*

Repetitive DNAs including transposable elements (TEs) and non-coding satellites colonize eukaryotic genomes. Their ability to rapidly evolve in sequence and copy number makes them a major driver of genome evolution. Genetic studies of hybrid sterility and lethality (hybrid incompatibilities, or HIs) demonstrate that HI-causing genes often localize to repeat-rich heterochromatin and repress repetitive DNA sequences, suggesting that repetitive DNA is an indirect cause of HI. We have also discovered one case where a species-specific satellite DNA directly causes hybrid lethality between Drosophila species.

These findings argue for the importance of a deeper understanding of repetitive DNA variation, but repeat-rich regions of most eukaryotic genome sequences are under-assembled or absent. We have developed new methods to identify and quantitate repetitive DNA from unassembled genome sequences, and are applying them to comprehensively analyze repetitive DNA from hundreds of natural population samples of Drosophila. We have discovered significant levels of polymorphism for different repeat families and are currently assaying samples with divergent repeat structures for effects on chromosome segregation.

Contact: barbash@cornell.edu

## Recombination-independent recognition of DNA homology

**Eugene Gladyshev, Nancy Kleckner**
**Department of Molecular and Cellular Biology, Harvard University**

Repeat-Induced Point mutation (RIP) is a highly efficient genome defense mechanism that exists in many filamentous fungi, whereby repetitive sequences undergo C-to-T transitions, resulting in eventual accumulation of in-frame stop codons and inactivation of transposable elements and duplicated genes. RIP in the model filamentous fungus *Neurospora crassa* provides an attractive model system for investigating recombination-independent recognition of DNA homology. During RIP, duplicated chromosomal DNA undergoes characteristic mutation specifically over the length of the duplication. This process occurs in haploid, mitotically dividing nuclei committed to meiosis, and it can detect repeats as short as 0.3 kbp at any position(s) in the genome.

We have recently shown that RIP does not require Rad51 [1]. Further, in a tester system involving partially homologous DNA segments placed next to regions of perfect, nucleating homology, RIP can recognize homology that occurs in units of three base pairs spaced with the periodicity of 11 or 12 base pairs. This pattern, along with other results, suggests that sequence information is communicated by triplets between co-aligned duplexes (presumably DNA/DNA, although RNA/DNA is not excluded). In subsequent work we now find that regions of perfect homology can be replaced with partial homology in which the minimally effective recognition units are pentanucleotides that must occur with the periodicity of 10 (but not 11 or 12) base pairs. Thus, homology recognition for RIP appears to involve at least two types of duplex/duplex interactions.

RIP is known to be strongly dependent on the cytosine methylase RID (RIP Deficient) [2]. Using uniquely sensitive assays, we now find that RIP still occurs at a significant level in the absence of RID, and this residual activity requires DIM-2, the only other cytosine methylase in Neurospora. DIM-2 is normally guided to DNA by Heterochromatin Protein 1 (HP1) that recognizes methylated histone H3-K9, a conserved mark of constitutive heterochromatin. We are therefore currently investigating the possibility that this or other components of the heterochromatin assembly pathway are involved in RID and/or DIM-2-mediated RIP.

[1] Gladyshev E, Kleckner N. 2014. Direct recognition of homology between double helices of DNA in *Neurospora crassa*. Nat Commun. 5: 3509. doi: 10.1038/ncomms4509.
[2] Freitag M, Williams RL, Kothe GO, Selker EU. 2002. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. PNAS 99(13): 8802-7.

**IntDOT: an unconventional tyrosine recombinase**

Jeff Gardner

University of Illinois at Urbana-Champaign

The mobile element CTnDOT encodes a site-selective recombination system which promotes integration into a recipient chromosome and excision from a donor chromosome. The integrase (IntDOT) catalyzes both reactions and is a member of the tyrosine recombinase family. Most tyrosine recombinases require sequence identity between the sites of strand exchange in the partner sites. CTnDOT is unusual because it can catalyze exchanges between sites that have different sequences between the sites of exchange. We have been studying the mechanism of this reaction using genetic and biochemical approaches.

Most tyrosine recombinase systems that show directionality utilize a single element-encoded accessory protein. The CTnDOT excision reaction is unusual because it requires three accessory proteins encoded by the element. Two of the proteins, Xis2C and Xis2D, are small, basic proteins that are required for excision. The third protein, Exc, is a topoisomerase that stimulates the reaction. However the topoisomerase activity is not required for its function in excision. We have been studying the interactions between these proteins and DNA using biochemical approaches.

**DNA Transposition In Archaea**

Alison B. Hickman

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

It has recently become clear that many bacterial and archaeal species possess adaptive immune systems. These are typified by multiple copies of DNA sequences known as clustered regularly interspaced short palindromic repeats (CRISPRs). These CRISPR repeats are the sites at which short spacers containing sequences of previously encountered foreign DNA are integrated, and the spacers serve as the molecular memory of previous invaders. Only two CRISPR-associated (cas) proteins, Cas1 and Cas2, are required for the acquisition stage of adaptation. However, two families of Cas1 proteins have been identified that are not associated with CRISPR loci, and in these, the *cas1* gene is in the genetic neighbourhood of several non-CRISPR genes and is enclosed within inverted repeats. It has been proposed that these DNA regions may be transposons ("casposons")[1]. I will describe the results of our efforts to experimentally characterize the properties of the putative transposases - or "casposases" - of casposons.

1. Krupovic *et al*. BMC Biol. 12:36 (2014).

# Elevated Ty1 retrotransposition in aging yeast mother cells

Melissa Patterson, Alison Scannapieco, Pak Ho Au, Savanna Dorsey, Catherine A. Royer, & Patrick H. Maxwell
Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY

Retrotransposon expression or mobility is elevated with age in multiple organisms, which raises the possibility that retrotransposons could influence aging through increased genome instability and changes in gene expression patterns. Whether activation of retrotransposons with age results from a gradual loss of the general silencing mechanisms that inhibit these elements or from retrotransposon-specific regulatory mechanisms that are age-dependent is unclear. Retromobility of a marked chromosomal Ty1 element was observed to be higher in mother cells than their daughter cells, based on magnetic cell sorting of aging mother cells from their daughter cell populations. Ty1 retromobility in mother cells was higher than predicted for their age in cell divisions, based on the rate of Ty1 retromobility in young cell populations. Mother cells were more likely to have very high concentrations of Ty1 Gag protein, which forms Ty1 virus-like particles. Older mother cell populations also had much more extrachromosomal Ty1 cDNA than their daughter cell populations. Ty1 insertions were more often in non-preferred genomic sites and more often associated with chromosome rearrangements in aging mother cells than in young cells. Overall, these results indicate that Ty1 replication intermediates may preferentially accumulate in mother cells with age and support the possibility that a specific mechanism restricts Ty1 replication to mother cells. Current work is focused on elucidating this mechanism. Many similarities in retrotransposon regulation and consequences exist between yeast and other species, making the work in this yeast model relevant for understanding the contribution of retrotransposons to aging in many organisms.

# Amplification of LTR-retrotransposons in contagious cancers of bivalves

Michael J Metzger[1], Carol Reinisch[2], James Sherry[2], Maria J Carballal[3], David Iglesias[3], Annette Muttray[4], Susan Baldwin[5], Antonio Villalba García[3], Stephen P Goff[1]

[1]Howard Hughes Medical Institute, Biochemistry and Molecular Biophysics, Columbia University, New York, NY; [2]Water Science & Technology Directorate, Environment Canada, Burlington, Canada; [3]Xunta de Galicia, Centro de Investigacións Mariñas, Vilanova de Arousa, Spain; [4]SLR Consulting, Vancouver, Canada; [5]Chemical and Biological Engineering, University of British Columbia, Canada

Fatal leukemia-like cancers have been observed in multiple bivalve species throughout the world. We identified an LTR-retrotransposon, *Steamer*, which is highly expressed and amplified in neoplastic cells of the soft-shell clam (*Mya arenaria*). Through the analysis of *Steamer* integration sites, mitochondrial DNA SNPs, and microsatellite loci, we found that the genotypes of the neoplastic cells do not match the genotypes of the hosts and that the genotypes of the neoplastic cells are nearly identical to each other. This shows that the neoplasia of soft-shell clams is transmitted from animal to animal as a contagious cancer cell, and that cancers from New York to Canada likely derive from a single cancer lineage. The increase in *Steamer* copy number in the neoplastic cells may have been a driver of oncogenesis or may be a passenger in this oncogenic lineage. Interestingly, we have found that *Steamer* expression and genomic copy number also increase with experimental injection of BrdU—known to induce primary leukemia-like disease. We have also found that contagious cancer cells are also responsible for disseminated neoplasia in three other bivalve species (mussels, cockles, and carpet clams). These represent independent examples of contagious cancer lineages, arguing that transmission of cancer cells is widespread in marine invertebrates. Most often the cancer cells found in each species are a clone derived from an individual of that same species. In one case, however, we have identified a cancer clone that is clearly identifiable as derived from one species (*Venerupis corrugata*) but is exclusively spreading in a different species *(V. aurea*), representing the first example of cross-species transmission. Additionally, we have identified *Steamer*-like elements in these and other bivalve species, and we are continuing to investigate their possible association with oncogenic development of these neoplastic lineages.

# Programmed Elimination of TEs during Chromatin Diminution in the Zooplankton *Mesocyclops edax* (Crustacea: Copepoda)

Authors: G Wyngaard[1], C Sun[2], B Walton[1], R Lockridge Mueller[2]
[1]James Madison University, [2]Colorado State University

Chromatin diminution in copepods is the programmed deletion of DNA from presomatic cell lineages during embryogenesis, This process marks the timing of germline-soma differentiation and reorganization of the somatic genome. In the freshwater *Mesocyclops edax* this excision results in a diploid somatic genome of 3 Gb which is derived from its diploid germline genome of 15 Gb. A previous study suggested that most of the eliminated DNA is comprised of transposable elements, of which many are active. To test this hypothesis, we assembled high throughput Illumina reads of germline (embryos) and somatic (antennae) DNA. The eliminated DNA was classifed into categories: (1-3) TEs that were preferentially deleted, preferentially retained, or unbiased in their deletion, (4) short simple repeats, and (5) genes. LTR Gypsy (~ 130,000 copies) and LINE L1 and LTR Pao (~ 80,000 copies each) were among the most abundant superfamilies excised. Class I TEs were typically both present in the germline and eliminated from the somatic line in lower copy numbers than class II TEs, with the exception of hAT-Ac for which ~ 60,000 copies were eliminated. Similar to the parasitic nematodes that possess chromatin diminution, a substantial number of genes (~ 1800) are completely eliminated from the somatic genome. Still, chromatin diminution in copepods remains unique among the phylogenetically widespread taxa that possess this trait in that the excised DNA is dominated by rapidly expanding populations of TEs, rather than degenerate TEs, short simple repeats and genes. Thus, in addition to gene regulation, chromatin diminution may serve as a mechanism to ameliorate deleterious effects of TEs and increased genome size in the soma.

# MULTIPLE MECHANISMS CONTRIBUTE TO TELOMERE MAINTNENANCE

Archie Call[1], Jessica McQuigg[1], David Velliquette[1], Hilda Ghadieh[1], Matt Legg[1], Benjamin Kaumeyer[1], Bryan Ross[1], Michelle Morgan[1], Leama Ajaka[1], Dulat Bekbolysnov[1], Vivek Behera[2*], Joshua Budman[2] **, Stephen Dria[2***], Ira Maine[1], Andrei Malykh[3]. Jose Luis-Garcia Perez[4], Tammy A. Morrish

1. University of Toledo, Department of Biochemistry & Cancer Biology, 2. Johns Hopkins University, Department of Biomedical Engineering, 3. Capital Biosciences, Inc 4. GENYO. Centre for Genomics and Oncological Research: Pfizer/ University of Granada/ Andalusian Regional Government

*Current Addresses: Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA and Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA
** Tissue Analytics, Baltimore, Maryland 21230
***Adient Medical, Pearland, TX 77584

Some human tumors lack telomerase and maintain telomeres using recombination-based mechanisms, typically called ALT(alternative lengthening of telomeres), which likely depend on Rad50 or Rad51 and may involve break-induced replication (BIR), extrachromosomal DNA, or LINE-1 retrotransposons. To mechanistically understand telomere recombination, we are using a B-cell lymphoma (Eμmyc+mTR-/-) model lacking telomerase and a shRNA approach. Lymphomas and primary cells were assayed for subtelomere copy number changes, a characteristic of BIR, using a custom Agilent eArray (aCGH) and analyzed using circular binary segmentation. Subtelomere copy number changes were frequent in Eμmyc+mTR-/-G4 lymphomas(>200breakpoints) and detected at every chromosome, compared to an isolated subset of chromosomes in Eμmyc+mTR+/+ lymphomas(<50breakpoints). In primary B-cells, negatively enriched from splenocytes and bone marrow, a small subset of chromosomes also had subtelomere recombination (<25 breakpoints). We next assayed the lymphomas and primary cells for the presence of extrachromosmal DNA, an additional characteristic of ALT+ tumors. Using this quantifiable C-circle assay, extrachromosomal DNA was highest in lymphomas lacking telomerase at levels similar to the human ALT+, Daoy cells(cc=5.30, and 3.40, respectively). Finally, to evaluate if genes implicated in telomere recombination-based mechanisms are altered in this model, Rad50 protein levels were assayed using a LI-COR Western protocol, and was elevated in splenocytes and lymphomas from both Eμmyc+mTR+/+ and Eμmyc+mTR-/-G3/4 compared to primary cells from mTR+/+ and mTR-/-G4 mice. However, knockdown of Rad50 (Rad50kd) only changed telomere lengths in Eμmyc+mTR-/-G4 tumors compared to Eμmyc+mTR+/+ (2 $x10^6$ bp vs 1x $10^5$ bp, respectively). We are currently investigating if Rad50kd impacts retrotransposition in human ALT+ and immortalized mTR+/+ and mTR-/- cells. Overall, these findings indicate that multiple mechanisms are likely contributing to telomere maintenance including subtelomere recombination and extrachromosomal DNA and are most prevalent in tumors lacking telomerase. Furthermore, the overexpression of Rad50 in lymphomas may be contributing to these mechanisms of telomere recombination.

**Activation of retrotransposons contributes to neurodegeneration in a Drosophila TDP-43 model of amyotrophic lateral sclerosis**

Krug, L.[1,2], Chatterje, N.[1], Borges-Monroy, R.[1,3], Julien, A.[1.4], Dubnau, J[1,2]

1. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA.
2. Watson School of Biological Sciences, Cold Spring Harbor Laboratory
3. The Undergraduate Program on Genomic Sciences of the National Autonomous University of Mexico
4. Magistère de Génétique Graduate Program at Université Paris Diderot, Sorbonne Paris Cité

Functional aberration of the aggregation-prone RNA binding protein, TDP-43, is central to the etiology of both familial and sporadic amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). Stereotyped protein pathology encompassing accumulation in cytoplasmic inclusions and associated nuclear clearance is observed in both neurons and glial cells in the vast majority of ALS patients and ~40% of FTLD patients. Such protein pathology and accompanying neurodegenerative phenotypes likewise occur in animal models in which human TDP-43 (hTDP-43) is expressed. Here we demonstrate that pathological expression of hTDP-43 in *Drosophila* neurons and glia results in activation of both a LINE-like element and an endogenous retrovirus (ERV) in the central nervous system. We further show that activation of retrotransposons is induced by TDP-43-specific erosion of small interfering RNA (siRNA) mediated gene silencing. We establish that loss of repression of a specific ERV, *Gyspy*, causally contributes to the neurodegenerative effects of hTDP-43 manipulation, as blocking *Gypsy* expression partially rescues hTDP-43 induced toxicity. Finally, we provide evidence that degeneration is caused by activation of DNA damage-mediated Chk2 signaling, resulting in apoptosis. These unpublished findings dovetail with our prior report that TDP-43 protein normally binds directly to retrotransposon-derived RNA sequences in human and rodent brain tissue. Our findings directly implicate retrotransposons and ERVs in the pathogenesis of TDP-43 mediated neurodegenerative disorders, including ALS and FTLD.

# A Damage Independent Role for 53BP1 that Impacts Chromatin Organization of *Igh* and Retrotransposons.

Pedro Rocha, Ramya Raviram, Yi Fu, JungHyun Kim, Arafat Aljoufi, Emily Swanzey, Vincent Luo, Alessandra Pasquarella, Alessia Balestrini, John Petrini, Gunnar Schotta and Jane Skok.

Class Switch Recombination (CSR), the process by which B cells generate antibody molecules with different effector functions requires the introduction of two double-strand breaks (DSBs) in the constant region of the immunoglobulin gene (*Igh)*, followed by ligation of the two free DNA ends. If not properly processed CSR has the potential to develop chromosomal translocations and initiate cancer.

We recently discoverd that the DNA damage sensor protein 53BP1 is responsible to ensure that the most upstream site for introduction of DSBs is always targeted first. Surprisingly, none of the proteins responsible for recruitment of 53BP1 to sites of DNA breaks or factors downstream of 53BP1 recapitulates this phenotype. This suggests that 53BP1 plays additional roles in CSR beyond its traditional role in DNA repair and that it is able to impact the choice of break location prior to AID-mediated DNA damage. 4C-seq experiments indicate that the mechanism underlying these alterations involves a 53BP1-mediated change in *Igh* chromatin architecture that impacts the spatiotemporal order in which breaks are introduced.

These studies support a model in which 53BP1, recruited to chromatin prior to damage, alters chromosome conformation and chromatin accessibility in a manner that impacts DNA damage. This effect also extends to other heterochromatic enriched sites such as transposable elements at which 53BP1 might be enriched. In the absence of 53BP1, retrotransposon-containing loci have an increase in the amount of chromosomal interactions that they are engaged in, which is suggestive of a more active epigenetic state. This could explain why in the absence of 53BP1, retrotransposons are more frequently involved in chromosomal translocations. Our work describes an unanticipated role for 53BP1 upstream of DNA damage that might help susceptible loci to deal with dangerous DSBs.