Summary

Antisense long noncoding (aslnc)RNAs are extensively degraded by the nuclear exosome and the cytoplasmic exoribonuclease Xrn1 in the budding yeast, lacking RNAi. Whether the ribonuclease III Dicer affects aslncRNAs in close RNAi-capable relatives remains unknown. Using genome-wide RNA profiling, here we show that aslncRNAs are primarily targeted by the exosome and Xrn1 in the RNAi-capable budding yeast, Dicer only affecting Xrn1-sensitive aslncRNAs levels in Xrn1-deficient cells. The and mutants display synergic growth defects, indicating that Dicer becomes critical in the absence of Xrn1. Small RNA sequencing showed that Dicer processes aslncRNAs into small RNAs, with a preference for Xrn1-sensitive aslncRNAs. Consistently, Dicer localizes into the cytoplasm. Finally, we observed an expansion of the exosome-sensitive antisense transcriptome in compared with, suggesting that the presence of cytoplasmic RNAi has reinforced the nuclear RNA surveillance machinery to temper aslncRNAs expression. Our data provide fundamental insights into aslncRNAs metabolism and open perspectives into the possible evolutionary contribution of RNAi in shaping the aslncRNAs transcriptome.

Antonin Morillon, Daniel Gautheret (2019 Jun 5)

Bridging the gap between reference and real transcriptomes.


Summary

Genetic, transcriptional, and post-transcriptional variations shape the transcriptome of individual cells, rendering establishing an exhaustive set of reference RNAs a complicated matter. Current reference transcriptomes, which are based on carefully curated transcripts, are lagging behind the extensive RNA variation revealed by massively parallel sequencing. Much may be missed by ignoring this unreferenced RNA diversity. There is plentiful evidence for non-reference transcripts with important phenotypic effects. Although reference transcriptomes are inestimable for gene expression analysis, they may turn limiting in important medical applications. We discuss computational strategies for retrieving hidden transcript diversity.

Aria Ronmans, Maxime Wery, Ugo Szachnowski, Camille Gautier, Marc Descrimes, Evelyne Dubois, Antonin Morillon, Isabelle Georis (2019 Mar 1)

Transcription-dependent spreading of the Dal80 yeast GATA factor across the
body of highly expressed genes.

*PLoS genetics*: e1007999 :: [DOI: 10.1371/journal.pgen.1007999](https://doi.org/10.1371/journal.pgen.1007999)

Summary

GATA transcription factors are highly conserved among eukaryotes and play roles in transcription of genes implicated in cancer progression and hematopoiesis. However, although their consensus binding sites have been well defined in vitro, the in vivo selectivity for recognition by GATA factors remains poorly characterized. Using ChIP-Seq, we identified the Dal80 GATA factor targets in yeast. Our data reveal Dal80 binding to a large set of promoters, sometimes independently of GATA sites, correlating with nitrogen- and/or Dal80-sensitive gene expression. Strikingly, Dal80 was also detected across the body of promoter-bound genes, correlating with high expression. Mechanistic single-gene experiments showed that Dal80 spreading across gene bodies requires active transcription. Consistently, Dal80 co-immunoprecipitated with the initiating and post-initiation forms of RNA Polymerase II. Our work suggests that GATA factors could play dual, synergistic roles during transcription initiation and post-initiation steps, promoting efficient remodeling of the gene expression program in response to environmental changes.

Bingning Xie, Emmanuelle Becker, Igor Stuparevic, Maxime Wery, Ugo Szachnowski, Antonin Morillon, Michael Primig (2019 Feb 15)

The anti-cancer drug 5-fluorouracil affects cell cycle regulators and potential regulatory long non-coding RNAs in yeast.


Summary

5-fluorouracil (5-FU) was isolated as an inhibitor of thymidylate synthase, which is important for DNA synthesis. The drug was later found to also affect the conserved 3′-5′ exoribonuclease EXOSC10/Rrp6, a catalytic subunit of the RNA exosome that degrades and processes protein-coding and non-coding transcripts. Work on 5-FU’s cytotoxicity has been focused on mRNAs and non-coding transcripts such as rRNAs, tRNAs and snoRNAs. However, the effect of 5-FU on long non-coding RNAs (IncRNAs), which include regulatory transcripts important for cell growth and differentiation, is poorly understood. RNA profiling of synchronized 5-FU treated yeast cells and protein assays reveal that the drug specifically inhibits a set of cell cycle regulated genes involved in mitotic division, by decreasing levels of the paralogous Swi5 and Ace2 transcriptional activators. We also observe widespread accumulation of different IncRNA types in treated cells, which are typically present at high levels in a strain lacking EXOSC10/Rrp6. 5-FU responsive IncRNAs include potential regulatory antisense transcripts that form double-stranded RNAs (dsRNAs) with overlapping sense mRNAs. Some of these transcripts encode proteins important for cell growth and division, such as the transcription factor Ace2, and the RNA exosome subunit EXOSC6/Mtr3. In addition to revealing a transcriptional effect of 5-FU action via DNA binding regulators involved in cell cycle progression, our results have implications for the function of putative regulatory IncRNAs in 5-FU mediated cytotoxicity. The data raise the intriguing possibility that the drug deregulates IncRNAs/dsRNAs involved in controlling eukaryotic cell division,
thereby highlighting a new class of promising therapeutical targets.

Blind exploration of the unreferenced transcriptome reveals novel RNAs for prostate cancer diagnosis
bioRxiv: DOI : 10.1101/644104

Summary
The broad use of RNA-sequencing technologies held a promise of improved diagnostic tools based on comprehensive transcript sets. However, mining human transcriptome data for disease biomarkers in clinical specimens is restricted by the limited power of conventional reference-based protocols relying on uniquely mapped reads and transcript annotations. Here, we implemented a blind reference-free computational protocol, DE-kupl, to directly infer RNA variations of any origin, including yet unreferenced RNAs, from high coverage total stranded RNA-sequencing datasets of tissue origin. As a bench test, this protocol was powered for detection of RNA subsequences embedded into unannotated putative long noncoding (Inc)RNAs expressed in prostate cancer tissues. Through filtering and visual inspection of 1,179 candidates, we defined 21 lncRNA probes that were further validated for robust tumor-specific expression by NanoString single molecule-based RNA measurements in 144 tissue specimens. Predictive modeling yielded a restricted probe panel enabling over 90% of true positive detection of cancer in an independent dataset from The Cancer Genome Atlas. Remarkably, this clinical signature made of only 9 unannotated lncRNAs largely outperformed PCA3, the only RNA biomarker approved by the Food and Drug Administration agency, specifically, in detection of high-risk prostate tumors. The proposed reference-free computational workflow is modular, highly sensitive and robust and can be applied to any pathology and any clinical application.

Year of publication 2018

Antoine Hocher, Myriam Ruault, Petra Kaferle, Marc Descrimes, Mickaël Garnier, Antonin Morillon, Angela Taddei (2018 Oct 26)
Expanding heterochromatin reveals discrete subtelomeric domains delimited by chromatin landscape transitions.
Genome research: DOI : gr.236554.118

Summary
The eukaryotic genome is divided into chromosomal domains of heterochromatin and euchromatin. Transcriptionally silent heterochromatin is found at subtelomeric regions, leading to the telomeric position effect (TPE) in yeast fly and human. Heterochromatin generally initiates and spreads from defined loci, and diverse mechanisms prevent the ectopic spread of heterochromatin into euchromatin. Here, we overexpressed the silencing
factor Sir3 at varying levels in yeast and found that Sir3 spreads into Extended Silent Domains (ESDs), eventually reaching saturation at subtelomeres. We observed the spread of Sir3 into subtelomeric domains associated with specific histone marks in wild-type cells and stopping at zones of histone mark transitions including H3K79 tri-methylation levels. Our study shows that the conserved H3K79 methyltransferase Dot1 is essential in restricting Sir3 spread beyond ESDs, thus ensuring viability upon overexpression of Sir3. Lastly, our analyses of published data demonstrate how ESDs unveil uncharacterized discrete domains isolating structural and functional subtelomeric features from the rest of the genome. Our work offers a new approach on how to separate subtelomeres from the core chromosome.


**Histone deacetylation promotes transcriptional silencing at facultative heterochromatin**

*Nucleic Acid Research* : DOI : 10.1093/nar/gky232

**Summary**

It is important to accurately regulate the expression of genes involved in development and environmental response. In the fission yeast Schizosaccharomyces pombe, meiotic genes are tightly repressed during vegetative growth. Despite being embedded in heterochromatin these genes are transcribed and believed to be repressed primarily at the level of RNA. However, the mechanism of facultative heterochromatin formation and the interplay with transcription regulation is not understood. We show genome-wide that HDAC-dependent histone deacetylation is a major determinant in transcriptional silencing of facultative heterochromatin domains. Indeed, mutation of class I/II HDACs leads to increased transcription of meiotic genes and accumulation of their mRNAs. Mechanistic dissection of the pho1 gene where, in response to phosphate, transient facultative heterochromatin is established by overlapping lncRNA transcription shows that the Clr3 HDAC contributes to silencing independently of SHREC, but in an lncRNA-dependent manner. We propose that HDACs promote facultative heterochromatin by establishing alternative transcriptional silencing.


**Diversification of human plasmacytoid predendritic cells in response to a single stimulus**

*Nature Immunology* : DOI : 10.1038/s41590-017-0012-z

**Summary**
Innate immune cells adjust to microbial and inflammatory stimuli through a process termed environmental plasticity, which links a given individual stimulus to a unique activated state. Here, we report that activation of human plasmacytoid predendritic cells (pDCs) with a single microbial or cytokine stimulus triggers cell diversification into three stable subpopulations (P1-P3). P1-pDCs (PD-L1+CD80-) displayed a plasmacytoid morphology and specialization for type I interferon production. P3-pDCs (PD-L1-CD80+) adopted a dendritic morphology and adaptive immune functions. P2-pDCs (PD-L1+CD80+) displayed both innate and adaptive functions. Each subpopulation expressed a specific coding- and long-noncoding-RNA signature and was stable after secondary stimulation. P1-pDCs were detected in samples from patients with lupus or psoriasis. pDC diversification was independent of cell divisions or preexisting heterogeneity within steady-state pDCs but was controlled by a TNF autocrine and/or paracrine communication loop. Our findings reveal a novel mechanism for diversity and division of labor in innate immune cells.

Year of publication 2017

Native Elongating Transcript Sequencing reveals global anti-correlation between sense and antisense nascent transcription in fission yeast
RNA : DOI: 10.1261/rna.063446.117

Summary

Antisense transcription can regulate sense gene expression. However, previous annotations of antisense transcription units have been based on detection of mature antisense long non-coding (aslnc)RNAs by RNA-Seq and/or micro-arrays, only giving a partial view of the antisense transcription landscape and incomplete molecular bases for antisense-mediated regulation. Here, we used Native Elongating Transcript sequencing to map genome-wide nascent antisense transcription in fission yeast. Strikingly, antisense transcription was detected for most protein-coding genes, correlating with low sense transcription, especially when overlapping the mRNA start site. RNA profiling revealed that the resulting aslncRNAs mainly correspond to cryptic Xrn1/Exo2-sensitive transcripts (XUTs). ChIP-Seq analyses showed that antisense (as)XUTs expression is associated with specific histone modifications patterns. Finally, we showed that asXUTs are controlled by the histone chaperone Spt6 and respond to meiosis induction, in both cases anti-correlating with levels of the paired-sense mRNAs, supporting physiological significance to antisense-mediated gene attenuation. Our work highlights that antisense transcription is much more extended than anticipated and might constitute an additional non-promoter determinant of gene regulation complexity.

History, Discovery, and Classification of IncRNAs
Summary

The RNA World Hypothesis suggests that prebiotic life revolved around RNA instead of DNA and proteins. Although modern cells have changed significantly in 4 billion years, RNA has maintained its central role in cell biology. Since the discovery of DNA at the end of the nineteenth century, RNA has been extensively studied. Many discoveries such as housekeeping RNAs (rRNA, tRNA, etc.) supported the messenger RNA model that is the pillar of the central dogma of molecular biology, which was first devised in the late 1950s. Thirty years later, the first regulatory non-coding RNAs (ncRNAs) were initially identified in bacteria and then in most eukaryotic organisms. A few long ncRNAs (lncRNAs) such as H19 and Xist were characterized in the pre-genomic era but remained exceptions until the early 2000s. Indeed, when the sequence of the human genome was published in 2001, studies showed that only about 1.2% encodes proteins, the rest being deemed “non-coding.” It was later shown that the genome is pervasively transcribed into many ncRNAs, but their functionality remained controversial. Since then, regulatory lncRNAs have been characterized in many species and were shown to be involved in processes such as development and pathologies, revealing a new layer of regulation in eukaryotic cells. This newly found focus on lncRNAs, together with the advent of high-throughput sequencing, was accompanied by the rapid discovery of many novel transcripts which were further characterized and classified according to specific transcript traits. In this review, we will discuss the many discoveries that led to the study of lncRNAs, from Friedrich Miescher’s “nuclein” in 1869 to the elucidation of the human genome and transcriptome in the early 2000s. We will then focus on the biological relevance during lncRNA evolution and describe their basic features as genes and transcripts. Finally, we will present a non-exhaustive catalogue of lncRNA classes, thus illustrating the vast complexity of eukaryotic transcriptomes.

Year of publication 2016

Alvaro de Andres-Pablo, Antonin Morillon, Maxime Wery (2016 May 28)
LncRNAs, lost in translation or licence to regulate?
Current genetics

Summary

Over the last decade, advances in transcriptomics have revealed that the pervasive transcription of eukaryotic genomes produces plethora of long noncoding RNAs (lncRNAs), which are now recognized as major regulators of multiple cellular processes. Although they have been thought to lack any protein-coding potential, recent ribosome-profiling data indicate that lncRNAs can interact with the translation machinery, leading to the production of functional peptides in some cases. In this perspective, we have explored the idea that translation can be part of the fate of cytoplasmic lncRNAs, raising the possibility for them to work as bifunctional RNAs, endowed with dual coding and regulatory functions.

Maxime Wery, Marc Desermes, Nicolas Vogt, Anne-Sophie Dallongeville, Daniel Gautheret,
Antonin Morillon (2016 Jan 26)

**Nonsense-Mediated Decay Restricts LncRNA Levels in Yeast Unless Blocked by Double-Stranded RNA Structure.**


**Summary**

Antisense long non-coding (aslnc)RNAs represent a substantial part of eukaryotic transcriptomes that are, in yeast, controlled by the Xrn1 exonuclease. Nonsense-Mediated Decay (NMD) destabilizes the Xrn1-sensitive aslncRNAs (XUT), but what determines their sensitivity remains unclear. We report that 3′ single-stranded (3′-ss) extension mediates XUTs degradation by NMD, assisted by the Mtr4 and Dbp2 helicases. Single-gene investigation, genome-wide RNA analyses, and double-stranded (ds)RNA mapping revealed that 3′-ss extensions discriminate the NMD-targeted XUTs from stable lncRNAs. Ribosome profiling showed that XUT are translated, locking them for NMD activity. Interestingly, mutants of the Mtr4 and Dbp2 helicases accumulated XUTs, suggesting that dsRNA unwinding is a critical step for degradation. Indeed, expression of anticomplementary transcripts protects cryptic intergenic lncRNAs from NMD. Our results indicate that aslncRNAs form dsRNA that are only translated and targeted to NMD if dissociated by Mtr4 and Dbp2. We propose that NMD buffers genome expression by discarding pervasive regulatory transcripts.

Jia Li, Marie-Anne Poursat, Damien Drubay, Arnaud Motz, Zohra Saci, Antonin Morillon, Stefan Michiels, Daniel Gautheret (2015 Nov 21)

**A Dual Model for Prioritizing Cancer Mutations in the Non-coding Genome Based on Germline and Somatic Events.**

*PLoS computational biology* : e1004583 : [DOI : 10.1371/journal.pcbi.1004583]

**Summary**

We address here the issue of prioritizing non-coding mutations in the tumoral genome. To this aim, we created two independent computational models. The first (germline) model estimates purifying selection based on population SNP data. The second (somatic) model estimates tumor mutation density based on whole genome tumor sequencing. We show that each model reflects a different set of constraints acting either on the normal or tumor genome, and we identify the specific genome features that most contribute to these constraints. Importantly, we show that the somatic mutation model carries independent functional information that can be used to narrow down the non-coding regions that may be relevant to cancer progression. On this basis, we identify positions in non-coding RNAs and the non-coding parts of mRNAs that are both under purifying selection in the germline and protected from mutation in tumors, thus introducing a new strategy for future detection of cancer driver elements in the expressed non-coding genome.
Marc Descrimes, Yousra Ben Zouari, Maxime Wery, Rachel Legendre, Daniel Gautheret, Antonin Morillon (2015 Sep 9)

**VING: a software for visualization of deep sequencing signals.**


**Summary**

Next generation sequencing (NGS) data treatment often requires mapping sequenced reads onto a reference genome for further analysis. Mapped data are commonly visualized using genome browsers. However, such software are not suited for a publication-ready and versatile representation of NGS data coverage, especially when multiple experiments are simultaneously treated.

Flore Sinturel, Albertas Navickas, Maxime Wery, Marc Descrimes, Antonin Morillon, Claire Torchet, Lionel Benard (2015 Sep 8)

**Cytoplasmic Control of Sense-Antisense mRNA Pairs.**

*Cell reports* : 1853-64 : [DOI: 10.1016/j.celrep.2015.08.016]

**Summary**

Transcriptome analyses have revealed that convergent gene transcription can produce many 3′-overlapping mRNAs in diverse organisms. Few studies have examined the fate of 3′-complementary mRNAs in double-stranded RNA-dependent nuclear phenomena, and nothing is known about the cytoplasmic destiny of 3′-overlapping messengers or their impact on gene expression. Here, we demonstrate that the complementary tails of 3′-overlapping mRNAs can interact in the cytoplasm and promote post-transcriptional regulatory events including no-go decay (NGD) in *Saccharomyces cerevisiae*. Genome-wide experiments confirm that these messenger-interacting mRNAs (mimRNAs) form RNA duplexes in wild-type cells and thus have potential roles in modulating the mRNA levels of their convergent gene pattern under different growth conditions. We show that the post-transcriptional fate of hundreds of mimRNAs is controlled by Xrn1, revealing the extent to which this conserved 5′-3′ cytoplasmic exoribonuclease plays an unexpected but key role in the post-transcriptional control of convergent gene expression.