

MEETING REPORT

The non-coding genome: a universe in expansion for fine-tuning the coding world

Reini F Luco

Abstract

A report on the EMBO/EMBL Symposium on The Non-Coding Genome, held in Heidelberg, Germany, 9-12 October, 2013.

We share 98% coding genome similarity with mouse and have about the same number of protein coding genes as worms, yet the differences in complexity are obvious. Where is this complexity encoded? A huge change in our understanding of genome evolution and regulation of gene expression arrived with the development of high-throughput sequencing technologies. It turns out that most of our genome is transcribed, but only a small percentage has coding information imbedded. The rest of the genome, the non-coding genome, mistakenly labeled as 'junk DNA', is where evolutionary complexity resides. In The Non-Coding Genome meeting, several research studies delved deeper into the importance of the non-coding genome, identifying novel classes of non-coding RNAs (ncRNAs) and novel regulatory functions, and expanding our knowledge about this new world, opening more exciting questions to study and answer.

The diversity of non-coding RNAs

Since Phillip Sharp first discovered introns in 1977, increasing evidence has demonstrated the complexity of our genome and has shaken well-established dogmas several times. John Mattick (Garvan Institute of Medical Research, Australia) unraveled a new form of four-dimensional organization of the genome in which alternatively spliced, but not constitutive, exons are preorganized around active promoters and enhancers. Antonio Giraldez (Yale University, USA) combined genome-wide ribosome profiling with mass spectrometry in zebrafish to identify more than 800 micropeptides that are

translated from ncRNAs, conserved in vertebrates and expressed mostly in early development, and that are important for gastrulation and heart development. In addition, Igor Ulitsky (Weizmann Institute, Israel) found common features between the open reading frames (ORFs) in the 5'UTR of ribosome-associated long intergenic non-coding RNAs (lincRNAs) and inefficiently translated mRNAs, which could explain the untranslatable nature of lincRNAs.

Small non-coding RNAs

MicroRNAs (miRNAs) are small (approximately 21 nucleotides in length) regulatory RNAs that control gene expression by either inhibiting protein translation or inducing mRNA decay by imperfectly base-pairing to 3'UTRs of target mRNAs. The biogenesis of miRNAs is under tight temporal and spatial regulation, and misregulation of these processes is often associated with disease. Narry Kim (Institute for Basic Science and Seoul National University, Korea) showed that during early development in flies, most mature miRNAs lose a 3' adenyl tail important for miRNA stability, and suggested that this could be implicated in the maternal-to-zygote transition. Eran Hornstein (Weizmann Institute of Science, Israel) ameliorated the neuromuscular deterioration of a classic mouse model of amyotrophic lateral sclerosis by recovering pre-miRNA processing via a small molecule that potentiates Dicer activity. In addition, Elisa Izauralde (Max Planck Institute for Development Biology, Germany) showed that the enrichment of tryptophan residues along the interacting pockets of GW182 with miRNA-loaded Argonaute and RNA decapping and deadenylase complexes is essential for the miRNA-mediated translational repression of mRNAs. Finally, David Bartel (HHMI/MIT/Whitehead Institute, USA) showed that in early zebrafish development, miRNAs inhibit protein translation by shortening the mRNA polyA tail rather than inducing degradation of the transcript, a mechanism that is predominant later in development.

Correspondence: reini.luco@igh.cnrs.fr
Institut de Genetique Humaine - CNRS UPR1142, Montpellier, France

Regarding other small RNAs (sRNAs), it is well established in plants and invertebrates that RNA interference (RNAi)-dependent silencing pathways regulate various processes, including regulation of gene expression, protection of genome integrity and innate immune responses against viruses. Olivier Voinnet (ETH Zurich, Switzerland) showed that in mammalian stem cells, antiviral small interfering RNAs (siRNAs) are also generated for defense against pathogens. In the animal germline, however, PIWI-interacting RNAs (piRNAs) are in charge of silencing mutagenic transposable elements. Eva-Maria Weick (The Wellcome Trust/Cancer Research UK Gurdon Institute, UK) demonstrated that there are different mechanisms of piRNA biogenesis in *Caenorhabditis elegans* depending on environmental conditions, with the synthesis of a novel population of type 2 piRNAs under stress. Maïke Laussmann (EMBL Heidelberg, Germany) showed by live cell imaging that key proteins in the piRNA pathway are highly dynamic and present different retention times in the cytoplasm granules (also known as nuages) *in vivo*, suggesting different functions in piRNA biogenesis.

Long non-coding RNAs

Since the discovery of Xist and HOTAIR, thousands of long non-coding RNAs (lncRNAs) have been identified and classified depending on their genomic location, whether they are intra- or intergenic, and according to their strand-specificity. lncRNAs have been increasingly shown to be involved in a wide range of regulatory functions, from the regulation of transcription to splicing, and in tethering chromatin-modifying complexes to DNA. One of the best-studied examples is Xist, which is essential for X chromosome inactivation. Jeannie Lee (Massachusetts General Hospital, USA) showed that Xist-mediated recruitment of polycomb repressive complex 2 (PRC2) during X inactivation is non-random with highly localized spreading along gene-rich domains first, whereas during maintenance all regions are covered at once, suggesting the existence of an epigenetic memory. Edith Heard (Institut Curie, France) presented the kinetics of X inactivation showing that Xist coating and loss of euchromatin marks precede jumonji, AT rich interactive domain 2 (JARID2)-dependent recruitment of PRC2 for gene silencing. Roberto Bonasio (HHMI/NYU School of Medicine, USA) confirmed a lncRNA-dependent interaction and combined recruitment of JARID2 and enhancer of zeste homolog 2 (*Drosophila*) (EZH2) to target genes in human induced pluripotent stem cells. Maite Huarte (CIMA/University of Navarra, Spain) described a novel and conserved nuclear lincRNA, PINT (p53-induced noncoding transcript), that is downregulated in tumor cells and induced by the tumor suppressor p53. Huarte demonstrated that PINT

is important for cell proliferation and survival via direct recruitment of PRC2 to pro-apoptotic genes. Howard Chang (Stanford University, USA) introduced an innovative genome-wide technique for analyzing chromatin accessibility by *in vitro* random insertion of a known transposon sequence into preferentially open chromatin, allowing posterior detection of the integration sites by deep sequencing (assay of transposase accessible chromatin deep sequencing (ATAC-seq)). Finally, John Rinn (Harvard University, USA) introduced the novel X-related lincFIRRE, involved in proximal approximation of nuclear territories. Rinn showed that 6 of the 18 lncRNA knockout strains they had generated exhibit distinct developmental or lethality phenotypes, highlighting the physiological relevance of lncRNAs.

New classes and functions for non-coding RNAs

The first circular RNA (circRNA) was discovered in 1978 from a viroid particle. Since then, many circRNAs have been described in eukaryotes and mammals and several functions have been identified related with this novel class of RNA. Using genome-wide analyses, Nikolaus Rajewsky (Max Delbrück Center for Molecular Medicine, Germany) identified thousands of new circRNAs in human cells, which represent 1% of total RNA. Most of these circRNAs are generated from exonic sequences, are conserved, and are more stable than linear RNAs. The study of 50 of those circRNAs indicated that many are developmentally regulated, and present many miRNA binding sites, suggesting a role as miRNA sponges. Thomas Hansen (Aarhus University, Denmark) delved deeper into circRNA regulation by identifying CiRS-7, which is conserved and highly expressed in the brain and is highly conserved. CiRS-7 inhibits miR-7 function by sequestering the miRNA, and can be repressed selectively by miR-671. Marc Buhler (Friedrich Miescher Institute for Biomedical Research, Switzerland) pointed out that ncRNAs not only act as attractants and guides, but also as evictors of chromatin modifiers. Such as in the case of borderline RNA, a new class of ncRNA that is expressed at insulators to repress the spreading of heterochromatin by evicting Heterochromatin protein 1 (HP1). Finally, Ling-Ling Chen (Shanghai Institutes of Biological Sciences, China) discovered two new classes of polyA(-) excised intron-derived lncRNAs, including linear small nucleolar (sno)-lncRNAs, which are flanked by snoRNA ends, and circular intronic long noncoding RNAs (ciRNAs), which are derived from inefficient debranching after splicing. A sno-lncRNA linked to Prader-Willi syndrome acts as a splicing factor molecular sponge that alters RNA binding protein fox-1 homolog 2 (FOX2)-dependent alternative splicing. The nuclear ciRNA ci-ankrd52 acts as a positive regulator of

transcription, suggesting that unknown types of ncRNAs with novel functions are yet to be discovered.

Non-coding RNAs in prokaryotes

An increasing characterization of the many ncRNAs expressed in bacteria, and other prokaryotes, is highlighting the diversity and unsuspected new roles of these RNAs in genome regulation. Gisela Storz (NIHCD-NIH, USA) has identified small ncRNAs in bacteria that act as transcriptional regulators important for adaptation and survival to host environments. Jennifer Doudna (University of California, USA) introduced an RNA-dependent defense system against foreign nucleic acids based on clustered regularly interspaced short palindromic repeats (CRISPR), which are widely used in bacteria and archaea as an immune system. Conversely, John van der Oost (Wageningen University, The Netherlands) presented a new defensive role for prokaryotic Argonaute, which uses small single-stranded DNA guides complementary to invasive plasmids for cleavage and degradation. Finally, Jorg Vogel (University of Wurzburg, Germany) described an innovative system to investigate the interplay between a pathogen, *Salmonella*, and its host cell during infection by dual high-throughput RNA sequencing (RNA-seq) of both pathogen and host at the same time. They discovered a new small RNA, STnc440, expressed in *Salmonella*, that is important for regulating the pathogen's virulence and its rate of replication for a more efficient infection.

Conclusions

The non-coding genome is intimately related to the coding genome, and a better understanding of this relationship will shed light on important regulatory mechanisms in development, cell commitment and disease. It seems there is a whole universe out there waiting to be discovered.

Abbreviations

circRNA: circular RNA; ciRNA: circular intronic long noncoding RNA; JARID2: jumonji, AT rich interactive domain 2; lincRNA: long intergenic non-coding RNA; lncRNA: long non-coding RNA; miRNA: microRNA; ncRNA: non-coding RNA; PINT: p53-induced noncoding transcript; piRNA: PIWI-interacting RNA; PRC2: polycomb repressive complex 2; siRNA: small interfering RNA; sno: small nucleolar RNA; sRNA: small RNA.

Competing interests

The author declares that she has no competing interests.

Acknowledgements

I thank Nicolás Rascovan (INDEAR, Rosario, Argentina) and Maite Huarte (CIMA, University of Navarra, Spain) for critical reading and comments.

Published: 22 November 2013

doi:10.1186/gb4140

Cite this article as: Luco: The non-coding genome: a universe in expansion for fine-tuning the coding world. *Genome Biology* 2013 **14**:314.