– Czy może nam pani powiedzieć, na terenie jakiego zakładu się obecnie znajdujemy?
– Tego ni mogę powiedzieć, bo to jest tajemnica państwowa! Mogę tylko powiedzieć, że mam 5 złotych od bombki.
– Czy pani jeszcze może nam powiedzieć, jakie bombki produkujecie?
– Panie, różne, różniste. Tu się robi tak: kuliste, w kształcie grzyba i cygara.
– Cały czas mowa o o bombkach choinkowych.
– Też. A, a wie pan, jaki mamy asortyment? Od A do N...
METHODS TO STUDY TRANSCRIPTOMES

• SAGE - serial analysis of gene expression
  sequencing of small cDNA tags generated by type II restriction enzymes

• CAGE - cap analysis of gene expression
  sequencing of small cDNA tags derived from capped transcripts

• 3’ long SAGE
  identification of SAGE tags that originate from 3’ ends of transcripts

• RNA Seq - high throughput sequencing of cDNAs

• GRO-seq - genomic run-on sequencing
  sequencing of cDNA tags extended from nascent transcripts

• tiling arrays
  microarrays with overlapping probes that cover the complete genome

• ChIP (ChIP-chip, ChIP-Seq) - chromatin immunoprecipitation
  indirectly reveal unknown ncRNAs

• DNA methylation and histone methylation genome-wide maps
METHODS TO STUDY TRANSCRIPTOMES

• NET-Seq - native elongating transcript sequencing
  RNA Seq of 3’ ends of nascent transcripts associated with Pol II

• RIP-Seq - RNA immunoprecipitation-sequencing

• ChIRP – Chromatin isolation by RNA Purification (+RNA-Seq)

• ChART - Capture Hybridization Analysis of RNA targets (+RNASeq)
  biotinylated oligonucleotides used to enrich for DNA sequences associated with RNA

• CRAC - CRosslinking and Analysis of cDNA

• PAR-CLIP - PhotoActivatable ribonucleoside–enhanced
  CrossLinking and ImmunoPrecipitation

• HITS-CLIP - High-Throughput Seq CLIP
METHODS TO DETECT RARE RNAs

• **NORTHERN** (not sensitive, can use LNA probes or RNA enrichment procedures, e.g. Ribo- to remove ribosomes)
• **RNAse protection** (more sensitive but tedious); 3’ / 5’ RACE
• **RT-PCR, qRT-PCR** (sensitive but often not specific for AT-rich)
• **RNA-Seq**
• **Splinted ligation** (only for RNAs with known, specific 3’ ends)

- phosphorylation of DNA oligo with P\(^{32}\)
- annealing in the presence of complementary DNA bridge (splinter)
- ligation of RNA to P\(^{32}\)-oligo with T4 DNA ligase
- phosphatase treatment
- gel electrophoresis and detection

*Maroney et al, Nat. Protocols, 2008*
ncRNA

- **Housekeeping**
  - constitutively expressed
  - required for normal function and cell viability
  - **tRNA** and **rRNA** – translation
  - **snRNA** – splicesosome components, pre-mRNA splicing
  - **snoRNA** – rRNA processing and modification, **scaRNA** (CB specific)
  - RNA components of **RNase P** and **RNase MRP** – endonucleases: tRNA and rRNA processing
  - Signal Recognition Particle **SRP RNA** – protein secretion to ER
  - **tmRNA** tRNA-mRNA hybrid- targeting nascent proteins for degradation
  - **gRNA** – guide RNA in RNA editing
  - **telomerase RNA** – synthesis of telomers

- **Regulatory**
  - expressed temporarily (development, response to stimuli)
  - affect gene expression at the level of transcription or translation
  - **sRNAs**: **siRNA** (exo-siRNAs and endo-siRNAs; ta-siRNA; nat-siRNA; lsiRNAs); **miRNA**; **piRNA** – act in TGS or PTGS
  - **IncRNAs** – much less known, usually act in TGS (chromatin level)
ALL ncRNAs?
FUNCTIONS of LONG ncRNAs

Chen and Carmichael, WIRERNA, 2010
FUNCTIONS of LONG ncRNAs

Wapinski and Chang, TiCS, 2011
<table>
<thead>
<tr>
<th>Biological Processes</th>
<th>IncRNA</th>
<th>Characteristics/Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-inactivation</td>
<td>XiST</td>
<td>Spreads on Xi in cis</td>
</tr>
<tr>
<td></td>
<td>Tsix</td>
<td>Prevents XiST stabilization and inhibits the interaction between Rep A and PRC2</td>
</tr>
<tr>
<td></td>
<td>XiTM</td>
<td>Enhancer of Tsix</td>
</tr>
<tr>
<td></td>
<td>Rep A</td>
<td>Binds to the subunit Ezh2 to recruit PRC2 to XiST</td>
</tr>
<tr>
<td></td>
<td>Air</td>
<td>Silences transcription by targeting G9a to Sjic22a3 promoter in Igf2r cluster</td>
</tr>
<tr>
<td>Genomic imprinting</td>
<td>Kcnqiot</td>
<td>Recruits polycomb group proteins to silence transcription at Kcnq1 cluster</td>
</tr>
<tr>
<td></td>
<td>H19b</td>
<td>GS control of Igf2; essential for human tumor growth</td>
</tr>
<tr>
<td></td>
<td>911^a</td>
<td>Trans-repressor of Igf2</td>
</tr>
<tr>
<td></td>
<td>Ube3A-ATS</td>
<td>n.d., but possibly associated with Prader-Willi syndrome (PWS)</td>
</tr>
<tr>
<td>Cell differentiation and</td>
<td>Snurfu</td>
<td>n.d., but possibly associated with Prader-Willi syndrome</td>
</tr>
<tr>
<td>developmental patterning</td>
<td>Nespa</td>
<td>n.d., but possibly associated with Beckwith-Wiedemann syndrome</td>
</tr>
<tr>
<td></td>
<td>Evf-2c</td>
<td>Gs coactivator with Dlx2 of Dlx5/6 enhancer</td>
</tr>
<tr>
<td></td>
<td>HOTAIR</td>
<td>Trans epigenetically represses HoxD locus through recruitment of PRC2 proteins</td>
</tr>
<tr>
<td></td>
<td>BORG</td>
<td>Induced upon mouse myoblast cell differentiated into osteoblastic cells by BMPs</td>
</tr>
<tr>
<td></td>
<td>EGO</td>
<td>Regulates eosinophil granule proteins, such as MBP and EDN transcript expression</td>
</tr>
<tr>
<td></td>
<td>Evf-4</td>
<td>n.d.</td>
</tr>
<tr>
<td></td>
<td>HOTAIRM1</td>
<td>Intergenic; involved in HoxA1, HoxA4 expression during myeloid differentiation</td>
</tr>
<tr>
<td></td>
<td>Nkx2.2AS</td>
<td>Enhances oligodendrocytic differentiation; spliced and polyadenylated</td>
</tr>
<tr>
<td></td>
<td>PINC</td>
<td>Developmentally regulated; involved in cell survival and cell cycle progression</td>
</tr>
<tr>
<td></td>
<td>Sx2ot</td>
<td>Stably transcribed in mESCs, up- or downregulated during ESCs differentiation</td>
</tr>
<tr>
<td>Nuclear foci enriched</td>
<td>TUG1</td>
<td>Required for the proper formation of photoreceptors in developing murine retina</td>
</tr>
<tr>
<td></td>
<td>Zfh-5A</td>
<td>Expressed in particular regions of the developing brain; regulates Zfh-5 (TF)</td>
</tr>
<tr>
<td>Small RNA processing</td>
<td>Menel/BNEAT1</td>
<td>Required for paraspeckle integrity; regulates IRALus-mRNA export</td>
</tr>
<tr>
<td></td>
<td>MALAT1/NEAT2</td>
<td>A precursor of masRNA; overexpressed in non-small cell lung cancer</td>
</tr>
<tr>
<td></td>
<td>TERRA</td>
<td>Contains UUAGG repeats; inhibits telomerase activity</td>
</tr>
<tr>
<td></td>
<td>SatII</td>
<td>Involved in the recruitment of RNA processing factors during stress responses</td>
</tr>
<tr>
<td></td>
<td>Gsmafu/MIA</td>
<td>n.d.</td>
</tr>
<tr>
<td>Disease associated</td>
<td>Fragile X</td>
<td>FMR4 Shares a bidirectional promoter with FMR1; silenced in fragile X patients and upregulated in premutation carriers; affects human cell proliferation</td>
</tr>
<tr>
<td></td>
<td>CD54</td>
<td>Bace1AS Regulates BACE1 expression</td>
</tr>
<tr>
<td></td>
<td>p15SAS</td>
<td>Epigenetically silences the tumor suppressor gene p15</td>
</tr>
<tr>
<td></td>
<td>CUDR</td>
<td>Cellular transformation and apoptosis</td>
</tr>
<tr>
<td></td>
<td>GASS</td>
<td>Growth arrest specific transcript; regulates apoptosis and cell cycle in lymphocytes</td>
</tr>
<tr>
<td></td>
<td>MEG3</td>
<td>Regulates p53</td>
</tr>
<tr>
<td></td>
<td>PCGEMI</td>
<td>Inhibits apoptosis, prostate specific and prostate cancer associated</td>
</tr>
<tr>
<td></td>
<td>PRINS</td>
<td>Tissue specific expression; required for psoriasis susceptibility; protective role in cell stress</td>
</tr>
<tr>
<td></td>
<td>SAF</td>
<td>Transcribed from the opposite strand of HFac; regulates Fas mediated apoptosis</td>
</tr>
<tr>
<td></td>
<td>UCA1</td>
<td>Influences cell growth and promoting invasion; upregulated in bladder carcinoma and embryo</td>
</tr>
<tr>
<td></td>
<td>B2/Alu</td>
<td>Repress transcription in trans during heat shock</td>
</tr>
<tr>
<td></td>
<td>CCND1 ncRNAs</td>
<td>Allosterically modifies the protein translocated in liposarcoma, which subsequently inhibits CREB-binding protein and p300 activities on a repressed gene target, CCND1</td>
</tr>
<tr>
<td></td>
<td>Dihydrofolate reductase (DHFR) upsteam</td>
<td>Regulates DHFR expression by formation of triple helix in the promoter</td>
</tr>
<tr>
<td>Others</td>
<td>NONON</td>
<td>Repressor of NFAT nuclear trafficking</td>
</tr>
<tr>
<td></td>
<td>Gadd7</td>
<td>Regulates lipid induced oxidative and ER stress</td>
</tr>
<tr>
<td></td>
<td>HSR1</td>
<td>Response to heat shock; knock down impairs heat shock response in vivo</td>
</tr>
<tr>
<td></td>
<td>Khps1a</td>
<td>Regulates DNA demethylation of Spht1 CpG island in a tissue specific manner</td>
</tr>
<tr>
<td></td>
<td>Zeb2/ATX</td>
<td>Antisense transcript to Zeb2, a transcriptional repressor of E-cadherin; regulates splicing of Zeb 5'-UTR</td>
</tr>
<tr>
<td>Proteins&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Long noncoding RNAs</td>
<td>Long noncoding RNA functions</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>CTCF</td>
<td>SRA</td>
<td>Enhances insulator function of CTCF</td>
</tr>
<tr>
<td>DNMT3b</td>
<td>pRNA</td>
<td>Targets DNMT3b in <em>cis</em> to the rRNA locus via an RNA:DNA:DNA triplex for cytosine methylation and gene silencing</td>
</tr>
<tr>
<td>G9α</td>
<td>Kcnq1ot1, Air</td>
<td>Targets H3K9 methylase G9α in <em>cis</em> for imprinting</td>
</tr>
<tr>
<td>Glucocorticoid receptor</td>
<td>Gas5</td>
<td>Binds to glucocorticoid receptor as a decoy and titrates GR away from target genes</td>
</tr>
<tr>
<td>hnRNP-K</td>
<td>lineRNA-p21</td>
<td>Targets hnRNP-K in <em>trans</em> to mediate p53-dependent gene repression</td>
</tr>
<tr>
<td>LSD1-CoREST</td>
<td>HOTAIR, many others</td>
<td>Targets the LSD1 complex to demethylate H3K4me2 to enforce gene silencing</td>
</tr>
<tr>
<td>MLL-WDR5</td>
<td>HOTTIP, some cRNAs?</td>
<td>Binds to and localizes the MLL complex and H3K4me3 via chromosomal looping for gene activation</td>
</tr>
<tr>
<td>NF-YA</td>
<td>PANDA</td>
<td>p53 inducible and titrates away NF-YA to favor survival over cell death during DNA damage</td>
</tr>
<tr>
<td>PRC1</td>
<td>ANRIL, Xist</td>
<td>Targets PRC1 in <em>cis</em> for gene silencing. ANRIL influences p16INK4a expression and cell senescence</td>
</tr>
<tr>
<td>PRC2</td>
<td>Xist, HOTAIR, ANRIL, COLDAIR, Gt12, Kcnq1ot1, many others</td>
<td>Targets PRC2 either in <em>cis</em> or <em>trans</em> to mediate H3K27 methylation and gene silencing for dosage compensation, imprinting, and developmental gene expression</td>
</tr>
<tr>
<td>Serine/arginine-rich splicing factors</td>
<td>MALAT1</td>
<td>Sequesters serine/arginine splicing factors to regulate alternative splicing</td>
</tr>
<tr>
<td>Staufen</td>
<td>1/2 SBS RNAs</td>
<td>Pairs with mRNAs via Alu repeats and targets them into a nonsense-mediated decay pathway</td>
</tr>
<tr>
<td>Set1 and Hda1/2/3 HDACs</td>
<td>CUTs, XUTs</td>
<td>Antisense RNAs repress sense transcription via control of H3K4me3 and histone deacetylation</td>
</tr>
<tr>
<td>hnRNP-A</td>
<td>TERRA</td>
<td>Controls telomerase access to telomeres in a cell-cycle phase-specific manner</td>
</tr>
<tr>
<td>TFIIIB</td>
<td>DHFR minor</td>
<td>Titrates away TFIIIB during cell quiescence to decrease DHFR transcription</td>
</tr>
<tr>
<td>TLS</td>
<td>CCND1 promoter ncRNA</td>
<td>Allosterically binds TLS to inhibit CREB binding protein and p300 activity, leading to repression of the CCND1 gene</td>
</tr>
<tr>
<td>YY1</td>
<td>Xist</td>
<td>YY1 binding nucleates Xist on the inactive X chromosome</td>
</tr>
</tbody>
</table>
# ncRNAs

## Classification based on transcript length

<table>
<thead>
<tr>
<th>Category</th>
<th>Abreviation</th>
<th>Specific examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long noncoding RNA</td>
<td>IncRNA</td>
<td></td>
</tr>
<tr>
<td>Long-intergenic noncoding RNA; large Intervening noncoding RNA</td>
<td>lincRNA</td>
<td>ANRIL [117], H19 [147], HOTAIR [148], HOTTIP [149], IncRNA-p21 [145], XIST [150], Paupir [151]</td>
</tr>
<tr>
<td>Very long Intergenic noncoding RNA</td>
<td>vliincRNA</td>
<td>HELLP transcript [42], Vinc_21, vInc_185, vInc_377, vInc_500 [29]</td>
</tr>
<tr>
<td>macroRNA</td>
<td></td>
<td>Aim_23T, KCNOOT1, Lincat, Nespress (reviewed in [152]), STAR1 [28]</td>
</tr>
<tr>
<td>Promoter-associated long RNA</td>
<td>PALR</td>
<td></td>
</tr>
</tbody>
</table>

## Classification based on association with annotated protein-coding genes

| Intronic ncRNA: stable intronic sequence RNA; totally intronic RNA, partially intronic RNA | siRNA, TN, PIN |
| Circular intronic RNAs | circRNAs |
| Sense ncRNA | |
| Natural antisense ncRNA | antRNA, NAT |
| BACE1 AS [163], aHIF [154], Teo [155] |
| Mirror antisense | |
| Exonic circular RNAs | ecircRNAs | cANRIL [118] |
| Chimeric RNAs, trans-spliced RNAs, exon juxtaosition | |
| Stand-alone miRNAs made from 3'UTRs | umiRNA |
| Chromatin-interlinking RNA | cRNA |
| Transcription start site-associated RNAs | TSSas RNAs |

## Classification based on association with other DNA elements of known function

| Enhancer-associated RNA | enhRNA |
| Promoter-associated long RNA | PALR |
| Upstream antisense RNA | umRNA |
| PROMotor uPstream Transcript | PROMPT |
| Telomeric repeat-containing RNA | TERRA |

## Classification based on protein-coding RNA resemblance

| mRNA-like noncoding RNAs | mIncRNAs |
| Long-Intergenic noncoding RNA; large Intervening noncoding RNA, long-Intervening noncoding RNA | lincRNA | ANRIL [117], H19 [147], HOTAIR [148], HOTTIP [149], IncRNA-p21 [145], XIST [150] |

## Classification based on association with repeats

| CUT:1 repeat RNA | |
| Long interspersed nuclear element | LINE 1/2 |
| Transcribed endogenous retroviruses | |
| Expressed Satellite Repeats | |
| Non-coding RNA driven by promoters within repeats | vliIncRNAs, NASTs |
| Polyurine-repeat-containing RNA | GRC-RNA |
| Transcribed pseudogenes | PTENPI and KRASPI [86] |

## Classification based on association with a biochemical pathway or stability

| miRNA primary transcripts | NUT |
| piRNA primary transcripts | |
| Cryptic unstable transcript | |
| PROMotor uPstream Transcript | PROMPT |
| X:1-sensitive unstable transcript | XUT |
| Stable Uncharacterized Transcript, Stable Unannotated Transcript | SUT |

## Classification based on sequence and structure conservation

| Hypoxia-induced noncoding ultraconserved transcript | T-U CR |
| Transcribed ultraconserved regions | UCR105 [95] |
| Long-Intergenic noncoding RNA, long-Intervening noncoding RNA, RNA-Z regions | H19 [166] |

Laurent et al, TiG 2015
MECHANISM of ACTION of LONG ncRNAs

ncRNAs recruit chromatin modifying complex to genes, resulting in histone modifications (H3meK27) and heterochromatin formation

- ncRNAs act as repressors or enhancers of transcription via binding to protein factors or DNA;
- may act as decoys to titrate trx factors away from genes

ncRNAs mask 5’ splice site resulting in intron retention, recognition of IRE and translation

Mercer et al., Nat. Rev. Genet., 2007
MECHANISM of ACTION of LONG ncRNAs

MODULAR PRINCIPLES of LARGE ncRNAs

1. RNA–Protein
2. DNA–RNA
3. Protein–DNA
4. RNA–RNA

1 + 2 = DHFR
1 + 3 = Hotair and Xist
1 + 4 = Ribosome
1 + 2 + 3 = Telomerase

EPIGENETIC REGULATION by NATs
(= Natural Antisense Transcripts)
MECHANISM of ACTION of LONG ncRNAs

- Cotranscriptional recruitment of chromatin-modifying factors.
- Nucleation of chromatin.
- Dynamic assembly of nuclear structures: paraspecies, nuclear bodies
- Formation of higher-order chromatin loops
  - GUIDES (chromatin modifiers)
  - TRX FACTORS
  - SCAFFOLDS (RNP structures)

Nagano and Fraser, Cell, 2011
**Xist ncRNA – inactivation of X chromosome (XCI)**

Dosage compensation – one copy of X chromosome in females is epigenetically silenced (*mammals*)

- **Xist** (X-inactive specific transcript, 19 kb) expressed from inactive X wrapped around X
- **Tsix** (40 kb) expressed from active X

**Expression of XIST ncRNA** → epigenetic changes → inactive state

- Histone exchange from **H2A** to **macroH2A**
- Histone H3 methylation: positions **H3K9**, **H3K27**
- Histone deacetylation H4 (?)
- DNA methylation following X inactivation (cellular memory)

**RepA** (repeat element) 1.6kb ncRNA (5’ of Xist) directly binds PRC2 (Polycomb)

**Tsix** - does not affect primary choice during XCI but protects active-X from silencing
- Links X reactivation and stem cell reprogramming

**Barr bodies:**
- Heterochromatic condensed
- X chromosome
**MALAT1/mascRNA**

**MALAT1:**
- metastasis-associated lung adenocarcinoma transcript 1 (*NEAT2* in humans)
- enriched in nuclear speckles
- possibly regulates alternative splicing (associates with SR proteins)

**mascRNA:**
- present in the cytoplasm, processed from pre-MALAT1, function unknown

- **Pol II polyadenylated transcript**, a minor form of MALAT1, precursor to mature MALAT1 and mascRNA
- Processing by RNaseP (5’) and RNaseZ (3’) releases 6.7 kb MALAT1 and tRNA-like mascRNA, exported to cytoplasm after addition of the CCA

*Wilusz and Spector, RNA, 2010*
UNUSUAL ncRNAs: NEAT1 and MALAT1

Chen and Carmichael, WIRERNA, 2010
UNUSUAL ncRNAs: MALAT1 FUNCTIONS
transcriptional activation and splicing

Tano and Akimitsu, Frontiers in Genetics, 2012
TERRA – telomeric repeat-containing RNA (yeast and human)

- polyadenylated Pol II transcript
- spans subtelomeric and telomeric regions
- a component of telomeric heterochromatin
- associates with telomeres and telomere proteins (Trf1, Trf2)
- regulated by RNA surveillance (Rat1, Trf4, NMD factors, RNAse H)
- regulates telomerase (telomere shortening) via RNA-DNA hybrids
- acts in chromatin remodelling (development and differentiation)
- affects telomere replication
- upregulated in ICF patients

(Immunodeficiency, Centromeric region instability, Facial anomalies)

rDNA SILENCING by pRNA and NoRC

NoRC – mammalian nucleolar remodeling complex which establishes and maintains heterochromatic state at promoters of silent rDNA repeats (histone modifications and CpG methylation)

- TIP5 TTF-I-interaction protein5
- SnF2 ATP-dependent chromatin remodeler

other
- TTF-1 transcription factor I
- UBF upstream binding factor
- DNMT DNA methyltransferase
- HDAC1 histone deacetylase
rDNA SILENCING by pRNA and NoRC

pRNA binds at $T_0$ to rDNA promoter, independently of TTF-I and other proteins, forming a triplex.

pRNA competes with TTF-I.

rDNA/pRNA triplex recruits methyltransferase DNMT3b.

This results in chromatin hypermethylation and rDNA silencing.

NoRC function requires TIP5 association with pRNA and promotes the binding of DNA methylation enzymes, leading to heterochromatin formation.

CpG-133 methylation prevents binding of UBF, inhibiting the formation of the transcription complex.

Pol I intergenic transcript is processed to Pol I intergenic transcript.

Stark and Taliansky, Embo Rep., 2008; Mayer et al., Mol. Cell, 2006; Embo Rep., 2008; Schmitz et al., Gene Dev., 2010
Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase

Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae

Bidirectional promoters generate pervasive transcription in yeast

Widespread bidirectional promoters are the major source of cryptic transcripts in yeast
INVISIBLE RNAs

INVISIBLE RNAs

RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters

Pascal Preker, Jesper Nielsen, Susanne Kammler, Søren Lykke-Andersen, Marianne S. Christensen, Christophe K. Mapendano, Mikkel H. Schierup, Torben Heick Jensen

Divergent Transcription from Active Promoters

Amy C. Seila, J. Mauro Calabrese, Stuart S. Levine, Gene W. Yeo, Peter B. Rahl, Ryan A. Flynn, Richard A. Young, Phillip A. Sharp

Post-transcriptional processing generates a diversity of 5' modified long and short RNAs

Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project

Cold Spring Harbor Laboratory Katalin Fejes-Toth, Vihra Sotirova, Ravi Sachidanandam, Gordon Assaf, Gregory J. Hannon; Affymetrix Philipp Kapranov, Sylvain Faisst, Aaron T. Willingham, Radha Duttagupta, Erica Dumais & Thomas R. Gingeras

Genome-Wide High-Resolution Mapping of Exosome Substrates Reveals Hidden Features in the Arabidopsis Transcriptome

Julia A. Chekanova, Brian D. Gregory, Sergei V. Reverdatto, Huaming Chen, Ravi Kumar, Tanya Hooker, Junshi Yazaki, Pinghua Li, Nikolai Skiba, Qian Peng, Jose Alonso, Vladimir Bruskin, Ueli Grossniklaus, Joseph R. Ecker and Dmitry A. Belostotsky
PERVERSIVE TRANSCRIPTION OF THE GENOME

All possible types of RNAs, detected by tiling microarrays and “deep sequencing”, SAGE and GRO, accompany major coding transcripts

(1) protein-coding mRNA; (2) PROMPT - promoter upstream transcripts (short); (3) PASR - promoter-associated sRNAs (< 200 nts); (4) TSSa transcription start site-associated RNAs (20-90 nts); (5) TASR - terminator associated sRNAs (< 200 nts); (6) PARL - promoter-associated long RNAs (> 200 nts); (7) tiRNAs - tiny transcription-initiation RNAs (18 nts)

SAGE, CAGE, GRO tags
antisense RNAs (can be long)
CUTs, SUTs - cryptic unstable or stable unannotated transcripts (200-600 nts)

Jacquier, Nat.Rev.Genet., 2009
PRESENCE of ncRNAs

Mercer et al., Nat.Rev.Genet., 2007

DENSITY of small RNAs

Jacquier, Nat.Rev.Genet., 2009
CUTs, SUTs, XUTs, MUTs and ALL THAT JAZZ

CUT = Cryptic Unstable Transcripts
SUT = Stable Unannotated Transcripts
SAT = Ssu72-associated Transcripts
XUT = Xrn1-dependent Unstable Transcripts
MUT = Meiotic Unstable Transcripts

NO LONGER TRANSCRIPTIONAL NOISE
(yeast, mammals, worms, plants - all organisms?)

- not visible in normal wild-type cells
- accumulate in RNA degradation mutants (EXOSOME, XRN family, TRAMP) or various metabolic conditions (aging, nutrient change, cell cycle etc)
- originate from widespread bidirectional promoters
- “mRNA-like” Pol II transcripts (capped, polyadenylated)

Jacquier, Nat. Rev. Genet., 2009
GENOMIC ORGANIZATION of ncRNA

A short

B long

Tisseur et al., Biochemie, 2011
GENERATION of bidirectional CUTs

Transcription activators (TAs) recruit general transcription factors (TFs)
TFs activate Pre-Initiation Complexes (PIC)
PICs recruit RNA Pol II to strong promoters (e.g. TATA) resulting in mRNA transcription or cryptic sites (both orientations) generating CUT RNAs

alternative models:
- Transcription of CUTs is driven by different PICs
- CUTs may result from „background” transcription due to nucleosome-poor regions
- Some CUTs are by-products of unconventional regulation mechanisms
- 3’ sRNAs (yeast) or TASR (mammals) may originate from gene loops (promoter and terminators regions interact)

Jacquier, Nat.Rev.Genet., 2009
ncRNA instability and their termination mode

3’ end CLEAVAGE and POLYADENYLATION (CP)

Nrd1/Nab3/Sen1-dependent termination

Jacquier, Nat. Rev. Genet 2009
Unstable CUTs (versus more stable SUTs)
- are detected in TRAMP or exosome mutants
- are terminated by Nrd1/Nab3-dependent mechanism and polyadenylated by Trf4/TRAMP
- Nrd1/Nab3, TRAMP and exosome complexes interact
- some CUTs (SRG1, IGS1-R) are polyadenylated by Pap1
- some CUTs are exported to the cytoplasm (XUTs) and degraded by Xrn1
- ncRNP composition is largely unknown
CUTs POSSIBLE MODES of ACTION

different PICs at CUT and mRNA promoters compete for TFs

transcription interference: CUT transcription displaces TFs from the mRNA promoter

TSS selection: CUT and mRNA have the same promoter but different TSS and compete for TFs

transcription-induced chromatin modification: CUT modifies chromatin at the mRNA promoter to silence transcription

U1 and non-coding transcription
PHYSIOLOGICAL FUNCTIONS of CUTs

Regulation of gene expression via CUT transcription and TSS selection: nucleotide shortage

CUT level unchanged (-/+ uracil)
CUT is terminated at the R-box
+ uracil -> no initiation at mRNA TSS -> no URA2 transcription
- uracil -> internal PolIII initiation at mRNA TSS

CUT readthrough produced
+ uracil -> scanning PolIII does not recognize mRNA TSS -> no URA2 transcription
- uracil -> scanning PolIII re-initiate at mRNA TSS

PHYSIOLOGICAL FUNCTIONS of CUTs

Regulation of gene expression via antisense RNA and epigenetic modification: 

**PHO84** (inorganic phosphate transporter)

Stabilization of as CUT leads to H3K18 deacetylation by Hda1 at **PHO84** promoter

**Camblong et al., Cell, 2007; Wery et al., WIREsSMB’11**
Regulation of gene expression via antisense RNA and epigenetic modification: GAL10-GAL1 locus

**Induction (galactose)** – full transcription of GAL1/GAL10 mRNAs

Repression (glucose) – Gal80/4 inhibitor binding at UAS inhibits transcription of GAL1/GAL10 mRNAs and allows Reb1 binding within GAL10 gene. This induces transcription of CUT RNA, which in turn leads to H3K36 histone methylation by HTM Set1 and Set2, histone deacetylation via recruitment of histone deacetylase complex Rpd3S, and further inhibition of mRNA transcription.
PHYSIOLOGICAL FUNCTIONS of CUTs

Intergenic Pol II ncRNAs downstream of Pol I rDNA units

IGS1-R ncRNAs affect rDNA recombination and regulate rDNA copy number

Loss of Nrd1 increases histone trimethylation and acetylation over IGS1

PHYSIOLOGICAL FUNCTIONS of XUTs

Transcriptional silencing of the Ty1 transposon

- Directly or indirectly controlled by Set1
- Polyadenylated Pol II transcript
- Antisense to TY1 promoter
- Degraded by cytoplasmic Xrn1
- Silences TY1 expression by promoting histone deacetylation and trimethylation (by Set1)
- Can act in-trans

Berretta et al., Gene Dev, 2008; Wery et al., WIREsSMB’11
PHYSIOLOGICAL FUNCTIONS of CUTs

**PHO84 silencing in-trans** (co-suppression)

- silencing *in-cis*
  - requires Hda1/2/3 histone deacetylases

- silencing *in-trans*
  - occurs at transcription initiation
  - is Hda1/2/3-independent
  - possibly requires additional silencing factors
  - **PHO84** antisense transcription depends on histone methyltransferase Set1

Camblong et al., Gene Dev., 2009
CUT ACTION *in-cis* or *in-trans*

CUT transcribed *in-cis*, when stabilized, recruits chromatin modification enzymes (HDAC) to gene promoter

CUT transcribed from a distant locus, when stabilized, recruits chromatin modification enzymes (HTM) to inhibit transcription

### ncRNA ACTION *in-cis* or *in-trans*

<table>
<thead>
<tr>
<th>Regulatory model</th>
<th>Expression correlation</th>
<th>Perturbation effect</th>
<th>Allele-specific regulation</th>
<th>Known ncRNA examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>trans</em></td>
<td></td>
<td>x</td>
<td></td>
<td>HOTAIR HoxC, HoxD</td>
</tr>
<tr>
<td><em>trans</em></td>
<td>✓</td>
<td>x</td>
<td></td>
<td>P53 lincRNA-P21 Cdkn1a/P21 Unknown</td>
</tr>
<tr>
<td><em>trans</em></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>trans</em></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Allele 1, Allele 2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><em>cis</em></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

- ✓ Neighbour affected
- x Neighbour unaffected

*Guttman and Rinn, Nature, 2012*
NOVEL ncRNAs: ceRNAs vs circRNAs

**ceRNAs**: competing endogenous RNAs, often antisense regulatory RNAs

**circRNAs**: circular RNAs, bind miRNAs and act as their antagonists, enhance cross-talk between ceRNAs

- ceRNA asRNA stabilizes mRNA by sequestering miRNAs that target mRNA

- circRNA antisense RNAs arise by head-to-tail splicing, contain miR-responsive elements and sequester miRNAs; often regulated via miRNAs and degraded by Ago2 Slicer

- circRNAs with distinct MREs may sequester different miRNA families

Taulliet al., Nat Str Mol Biol., 2013
ceRNAs

circRNAs

Guil and Esteller, TiBS 2015
circRNAs regulate transcription

exon-intron circRNAs (EIciRNAs)
- localize in the nucleus
- associate with U1 snRNP
- enhance the expression of their parental gene in trans

Li at al., Nat Struct Mol Biol, 2015
NOVEL ncRNAs: eRNAs

**eRNAs**: enhancer RNAs, short (not always, up to 2 kb) ncRNAs transcribed from enhancer regions (RNA-Seq, ChIP-Seq)

2d-eRNAs: bidirectional, comparatively short, nonpolyadenylated

1d-eRNAs: unidirectional, long, polyadenylated

*Natoli and Andrau, Annu Rev Genet., 2012*
NOVEL ncRNAs: eRNAs, functions

Noise:
Transcription of enhancers reflects random collisions of Pol II with accessible genomic regions.

Transcription-dependent effects:
Movement of Pol II molecules causes changes to the chromatin template affecting its accessibility.

RNA-dependent effects:
Transcripts generated at enhancers cause functional effects either in cis or in trans collaborating with transcriptional activators or evicting repressors.

Natoli and Andrau, Annu Rev Genet., 2012
NOVEL ncRNAs: eRNAs, functions

Chromosome looping

Quinn and Chang, Nat Rev Genet 2015
NOVEL ncRNAs: ciRNAs

ciRNAs: circular intronic IncRNAs, accumulate in human cells due to lariat debranching defect

- processing depends on GU-rich motive near 5’ splice site and branchpoint
- regulate parent gene expression by modulating elongation Pol II activity

Zhang et al, Mol Cell., 2013
UNUSUAL ncRNAs: tRFs tRNA-derived RNA fragments

Stress-induced enzymatic tRNA cleavage
(S. cerevisiae, D. melanogaster, A. thaliana, A. nidulans, human cell lines)

Thompson and Parker, Cell, 2009
UNUSUAL ncRNAs: tRFs  tRNA-derived RNA fragments

- > 17 short abundant tRFs (13-26 nts), generated by RNaseZ from mature (5’ and 3’ ends) and precursor (3’ trailer) tRNAs identified in the cytoplasm in prostate cancer cells. Lack of tRF1001 impairs cell proliferation.
- Abundant Dicer-dependent tRFs (class I, from mature 3’ and 5’ ends) in HeLa moderately downregulate target genes.
- Class II tRFs (from RNaseZ 3’ cleavage to Pol III termination, cytoplasmic) associate with Ago2-3. Function- regulation of silencing via differential association with Ago proteins?
Angiogenin-derived 5’-tiRNAs with terminal 5’-oligoG
- repress translation in vitro and in vivo
- displace eIF4G/eIF4A from uncapped transcripts and eIF4F from m7G cap
- trigger formation of stress granules (SGs)
- translational repressor YB-1 contributes to tiRNA-mediated repression

18-mer ncRNA derived from TRM10 mRNA during salt stress in yeast
- associates with polysomes
- inhibits general translation

Gebetsberger and Polacek, RNA Biol., 2013
Non-canonical miRNAs

Maute et al, WIREs RNA., 2014
Unusual ways of ncRNAs

Quinn and Chang, Nat Rev Genet 2015
TAKE-HOME MESSAGE

• The majority of eukaryotic genomes are transcribed giving rise to a variety of RNAs

• At least some of the “invisible” transcripts in some conditions form functional ncRNAs

• These usually act in transcriptional silencing *in-cis* or *in-trans* by recruiting modifying enzymes (DNA, histones) to promoters or interacting with DNA (pRNA)

• Defects in ncRNA level or activity correlate with several diseases