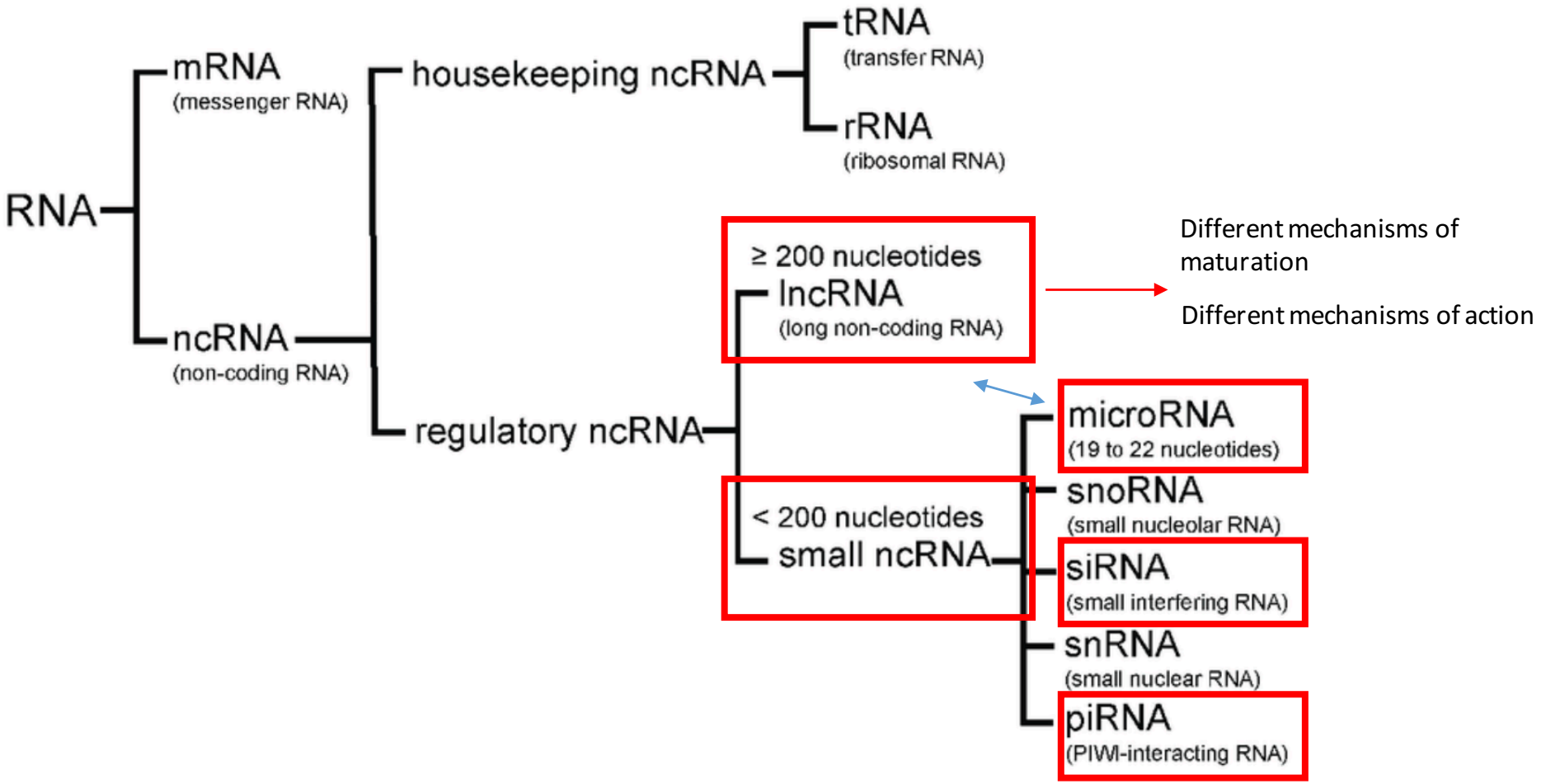
DIFFERENT CATEGORIES OF ncRNAs

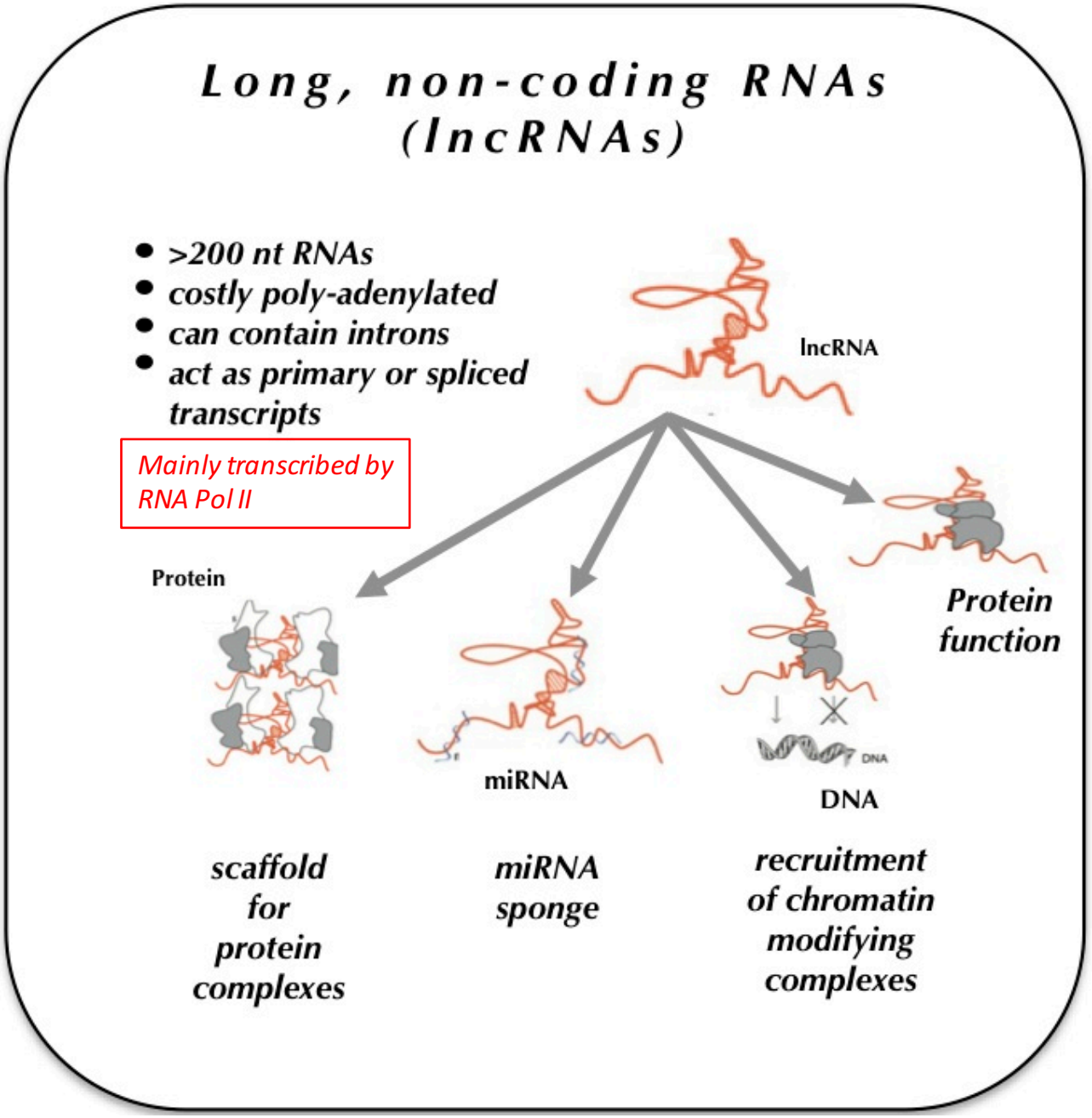


RNA categories. RNAs are divided into two major classes: messenger RNA (mRNA) and non-coding RNA (ncRNA). NcRNAs include housekeeping ncRNA, which consists of transfer RNA (tRNA) and ribosomal RNA (rRNA), and regulatory ncRNA. Regulatory ncRNAs are classified into long ncRNA (lncRNA) and small ncRNA. Small ncRNAs are subclassified into microRNA, small nucleolar RNA (snoRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), and PIWIinteracting RNA (piRNA).

lncRNAs

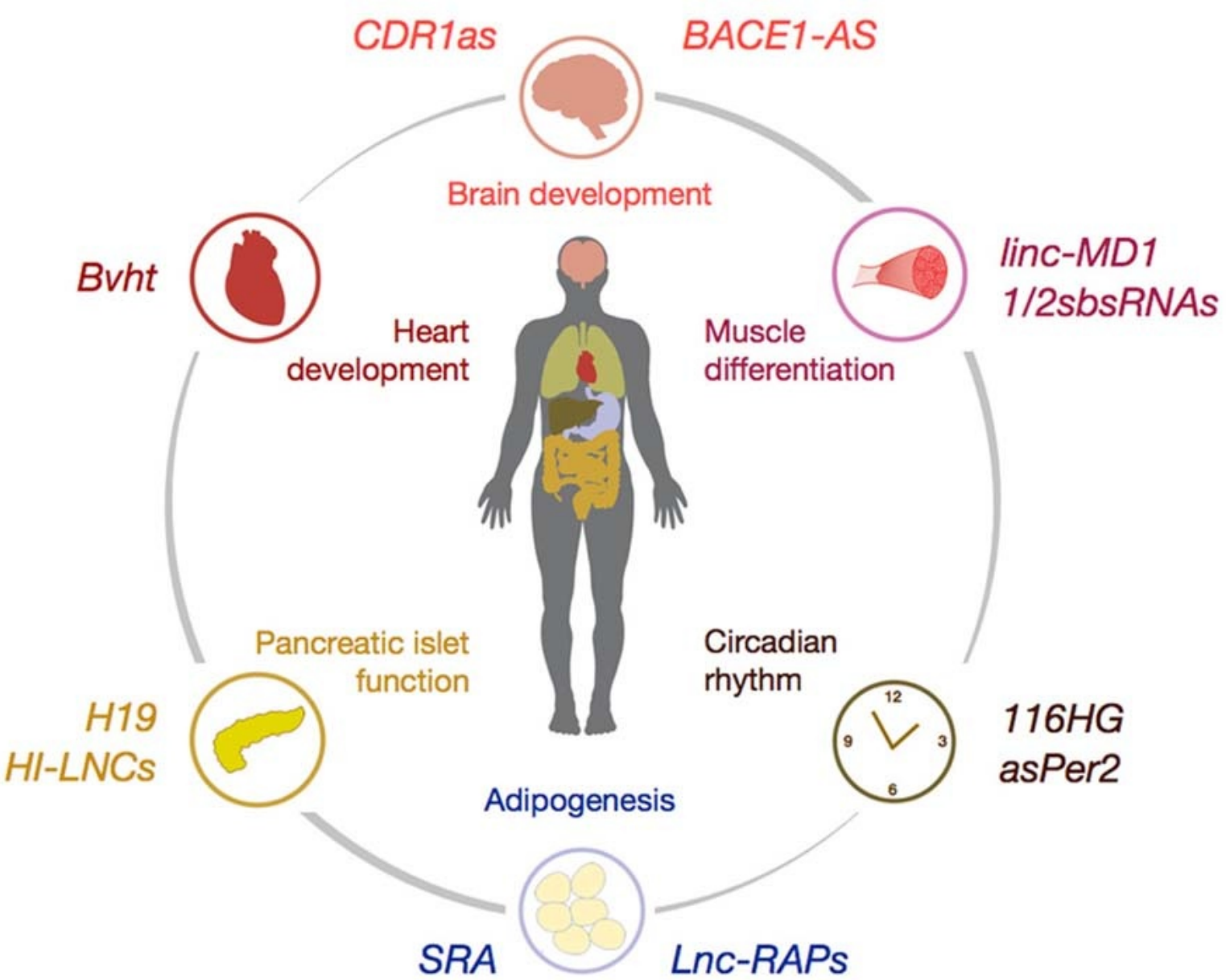
long, non-coding RNAs

Characteristics of lncRNAs



nuclear and/or cytoplasmatic localiztion

lncRNAs in physiology



COMPARING mRNAs - lncRNAs

Resource

The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression

Thomas Derrien,[1,11] Rory Johnson,[1,11] Giovanni Bussotti,[1] Andrea Tanzer,[1] Sarah Djebali,[1] Hagen Tilgner,[1] Gregory Guernec,[2] David Martin,[1] Angelika Merkel,[1] David G. Knowles,[1] Julien Lagarde,[1] Lavanya Veeravalli,[3] Xiaoan Ruan,[3] Yijun Ruan,[3] Timo Lassmann,[4] Piero Carninci,[4] James B. Brown,[5] Leonard Lipovich,[6] Jose M. Gonzalez,[7] Mark Thomas,[7] Carrie A. Davis,[8] Ramin Shiekhattar,[9] Thomas R. Gingeras,[8] Tim J. Hubbard,[7] Cedric Notredame,[1] Jennifer Harrow,[7] and Roderic Guigó[1,10,12]

[1]*Bioinformatics and Genomics, Centre for Genomic Regulation (CRG) and UPF, 08003 Barcelona, Catalonia, Spain;* [2]*INRA, UR1012 SCRIBE, IFR140, GenOuest, 35000 Rennes, France;* [3]*Genome Institute of Singapore, Agency for Science, Technology and Research, Genome 138672, Singapore;* [4]*Riken Omics Science Center, Riken Yokohama Institute, Yokohama, Kanagawa 351-0198, Japan;* [5]*Department of Statistics, University of California, Berkeley, California 94720, USA;* [6]*Center for Molecular Medicine and Genetics, Wayne State University, Detroit, Michigan 48201, USA;* [7]*Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, United Kingdom;* [8]*Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA;* [9]*The Wistar Institute, Philadelphia, Pennsylvania 19104, USA;* [10]*Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, 08002 Barcelona, Catalonia, Spain*

22:1775–1789 © 2012, Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/12; www.genome.org

BASED ON ENCODE DATA → GENCODE ANNOTATION OF lncRNAs

FIRST DETAILLED, GENOME-WIDE ANNOTATION OF lncRNAs

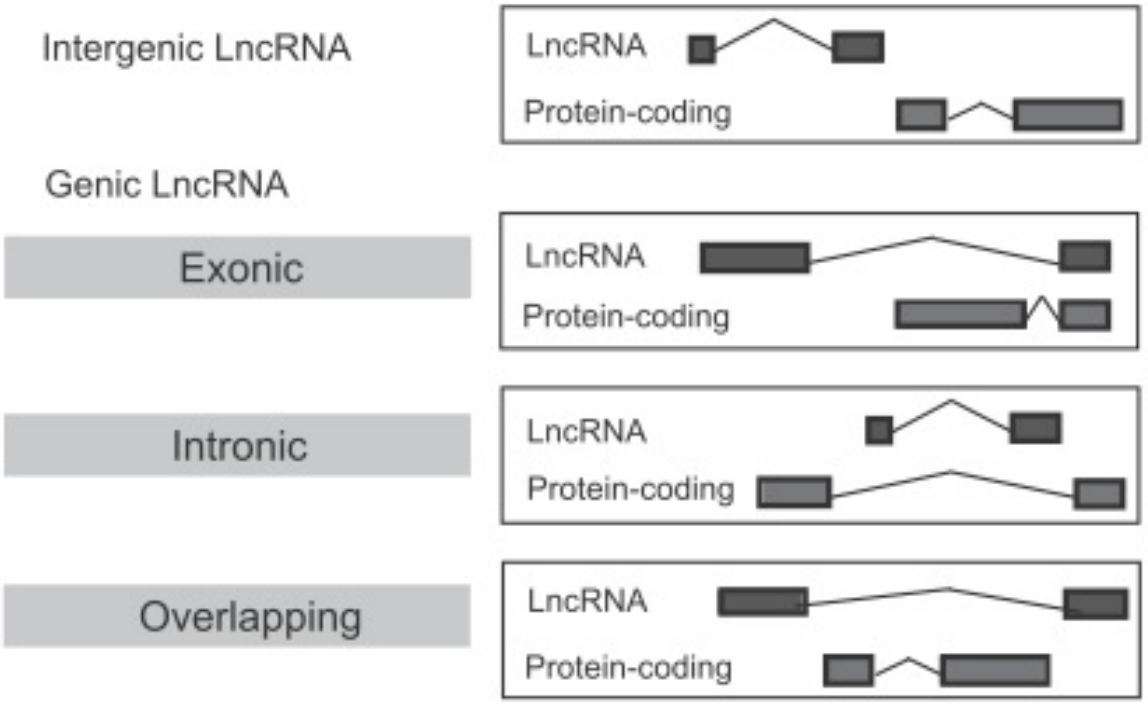
COMPARING mRNAs - lncRNAs

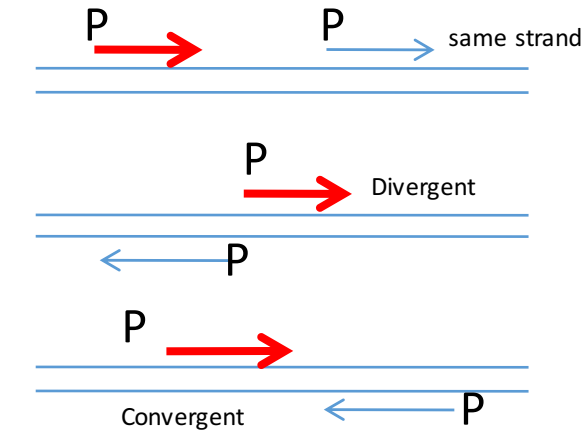
lncRNAs can be

1. **Intergenic (lincRNA)**: Do not intersect with protein coding gene

2. Genic:
Intersect a protein coding gene

- Exonic
- Intronic
- overlapping





Gencode LncRNAs transcripts (14,880)										
Intergenic (9,518)			Genic (5,362)							
Same Strand	Convergent	Divergent	Exonic (2,411)		Intronic (2,784)		Overlapping (167)			
4,165	1,937	3,416	S	AS	S	AS	S	AS		
			NA	2,411	563	2,221	52	115		
				Why?		Why?		Why?		

Closest protein coding gene → lncRNA

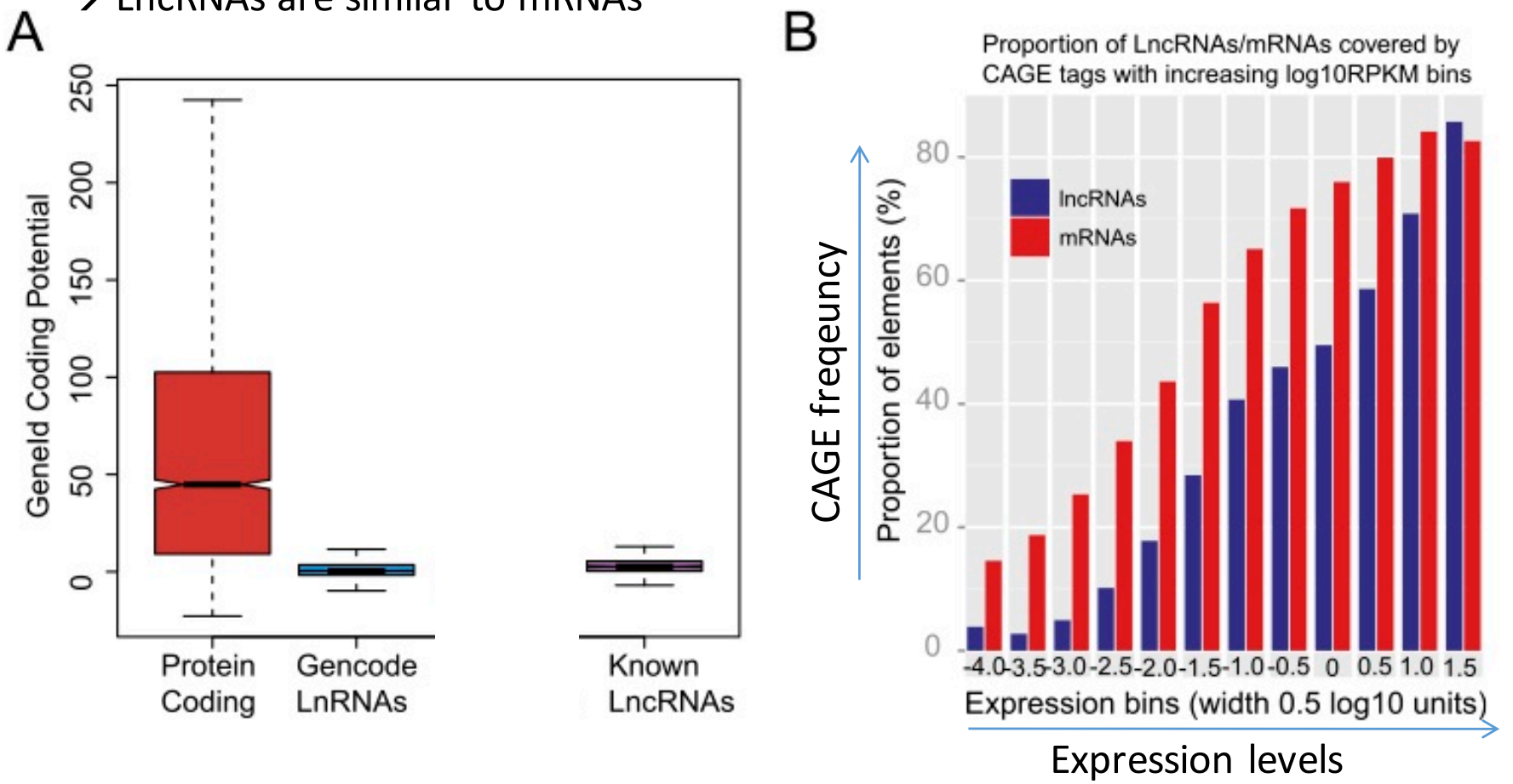
P: promoter

COMPARING mRNAs - lncRNAs

CODING POTENTIAL
INITIATION - EXPRESSION

A. LncRNAs do not have coding potential:
→ Longest possible ORF was searched in mRNA/lncRNAs

B. CAGE tags can be found in lncRNAs → defined transcriptional start site
CAGE tag freqeuncy increases with increased lncRNA expression levels.
mRNAs are characterized by more CAGE tags
→ LncRNAs are similar to mRNAs



TRASCRITTOMICA

Schedule lectures– AA 2020/2021

October (18 hours)
L1: 05.10.2020: 14-16 (2h)
L2: 08.10.2020: 10:45-12:45 (2h)
L3: 09.10.2020: 10:45-12:45 (2h)
L4: 12.10 2020: 14-16 (2h)
L5: 15.10.2020: 10:45-12:45 (2h)
L6: 16.10.2020: 10:45-12:45 (2h)
L7: 19.10.2020: 14-16 (2h)
L8: 22.10.2020: 10:45-12:45 (2h)
L9: 23.10.2020: 10:45-12:45 (2h)
November (16 hours)
L10: 05.11.2020: 10:45-12:45 (2h) Seminar 1
L11: 06.11.2020: 10:45-12:45(2h) Seminar 2
L12: 12.11.2020: 10:45-12:45 (2h) Seminar 3
L13: 13.11.2020: 10:45-12:45 (2h) Seminar 4
L14: 19.11.2020: 10:45-12:45 (2h) Seminar 5
L15: 20.11.2020: 10:45-12:45 (2h) Seminar 6
L16: 26.11.2020: 10:45-12:45 (2h) Seminar 7
L17: 27.11.2020: 10:45-12:45 (2h) Seminar 8
December (12 hours)
L18: 04.12.2020: 14-16 (2h) → EXTRA SEMINAR
L19: 09.12.2020: 08:45- 10:45(2h) Seminar 9
(change with Sistemi Modelli → 03.12.2020 (10:45 – 12:45)
L20: 10.12.2020: 16-18 (2h) Seminar 10
L21: 11.12.2020: 11-13 (2h) Seminar 11
L22: 17.12.2020: 16-18 (2h) Seminar 12
L23: 18.12.2020: 16-18 (2h) Seminar 13
L24: XX.YY,ZZZZPresentation Studenti fuori Trieste: Seminae 14
January
XX: 15.01.2021: 10:45-12:45 (2h)
XX: 21.01.2021: 10:45-12:45 (2h)
48 ore = 6CFU

Edificio C1, Aula L

PPT SLIDES:

MOODLE FEDERALE

PASSWORD: Trascrittomica

Prof. Stefan Schoeftner
E-mail: sschoeftner@units.it

I only reply to official students' emails: @units.it
(no @gmail; @libero….)

Students' representatives: luca.secco@studenti.units.it

Students' representatives: federico.fierli@.studenti.units.it

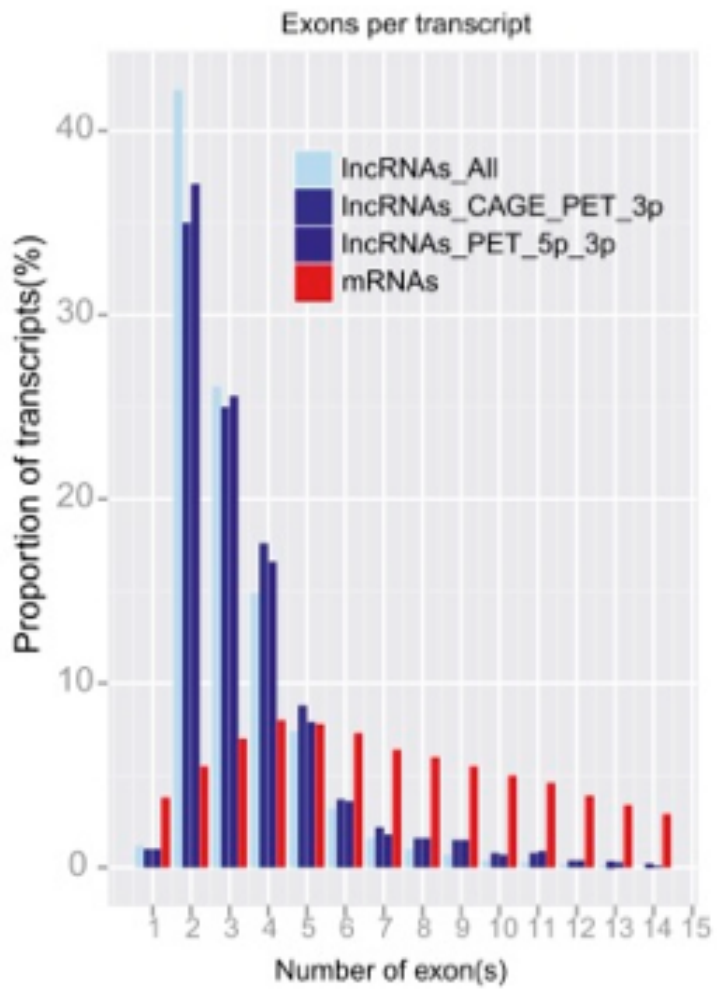
TRASCRITTOMICA
STUDENT SEMINARS

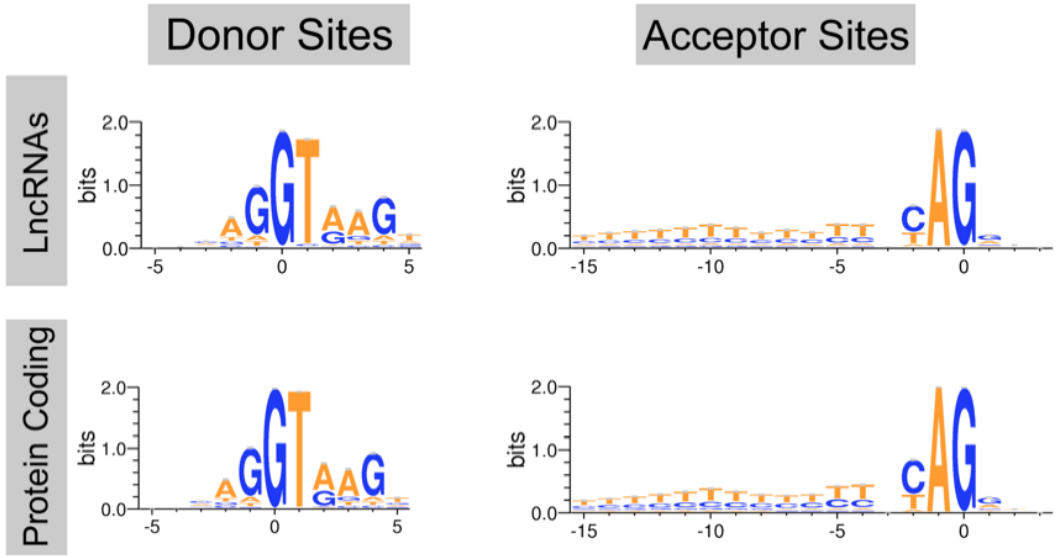
1. Reduce from 17 → 14 groups
2. Form groups only with stundets that do not come to Trieste

Numero gruppo	Studente 1 + email istituzionale	Studente 2 + email istituzionale	Studente 3 + email istituzionale
1	Eleonora Capezzali s266393@ds.units.it	Viviana Ippolito viviana.ippolito@studenti.units.it	Giulia Canarutto s231374@ds.units.it
2	Anna De Lucia ANNA.DELUCIA@studenti.units.it	Giuseppe Molinaro GIUSEPPE.MOLINARIO@studenti.units.it	Domenico De Angelis DOMENICO.DEANGELIS@studenti.units.it
3	Agostino Campanile AGOSTINO.CAMPANILE@studenti.units.it	Alessia Pesaresi alessia.pesaresi@studenti.units.it	Patrik Rossi patrik.rossi@studenti.units.it
4	Antonio Manai s231230@ds.units.it	Beatrice Alessandrini beatrice.alessandrini@studenti.units.it	
5	Sonia Panico s268005@ds.units.it	Fracesca Cuomo s267148@ds.units.it	Lara Favero s239076@ds.units.it
6	Luca Secco LUCA.SECCO@studenti.units.it	Sara Maggiore SARA.MAGGIORE@studenti.units.it	Francesco Furio D'Amico FRANCESCOFURIO.D'AMICO@studenti.units.it
7	Silvia Aldrovandi silvia.aldrovandi@studenti.units.it	Matteo Colombo matteo.colombo@studenti.units.it	Richard Camara s271237@ds.units.it
8	Nicoletta Franco NICOLETTA.FRANCO@studenti.units.it	Camilla D'Angelo CAMILLA.D'ANGELO@studenti.units.it	Mbarsid Racaku s265766@ds.units.it
9	Catello Guida s270324@ds.units.it	Domenico Loperfido domenico.loperfido@studenti.units.it	Angelo Acito angelo.acito@studenti.units.it
10	Karim Rahhali KARIM.RAHHALI@studenti.units.it	Elena Petrini ELENA.PETRINI@studenti.units.it	William Bongiovanni WILLIAM.BONGIOVANNI@studenti.units.it
11	Giulia Visani s269271@ds.units.it	Roberta Noè s270858@ds.units.it	Camilla Volponi s270857@ds.units.it
12	Alessia Ferraro alessia.ferraro@studenti.units.it	Cinzia Bessone cinzia.bessone@studenti.units.it	Elena Di Mattia elena.dimattia@studenti.units.it
13	Federico Fierli federico.fierli@studenti.units.it	Domiziano Dario Tosi domizianodario.tosi@studenti.units.it	Giuseppe Castiglione giuseppe.castiglione@studenti.units.it
14	Rossana Putino rossana.putino@studenti.units.it	Gloria Leva gloria.leva@studenti.units.it	Elena Peresani s230866@ds.units.it
15	Camillo Balanzin s239118@ds.units.it	Michele Esposito s260507@ds.units.it	Martin Belliu MARTIN.BELLIU@studenti.units.it
16	Elisabetta Molteni ELISABETTA.MOLTENI@studenti.units.it	Federica Ruggiero FEDERICA.RUGGIERO@studenti.units.it	Maria Alessia Cucco s260331@ds.units.it
17	Ilaria Frascolla ILARIA.FRASCOLLA@studenti.units.it	Elisabetta Giordano ELISABETTA.GIORDANO@studenti.units.it	Sefora Naomi Agrò SEFORANAOMI.@studenti.units.it

COMPARING mRNAs - lncRNAs

SPLICING





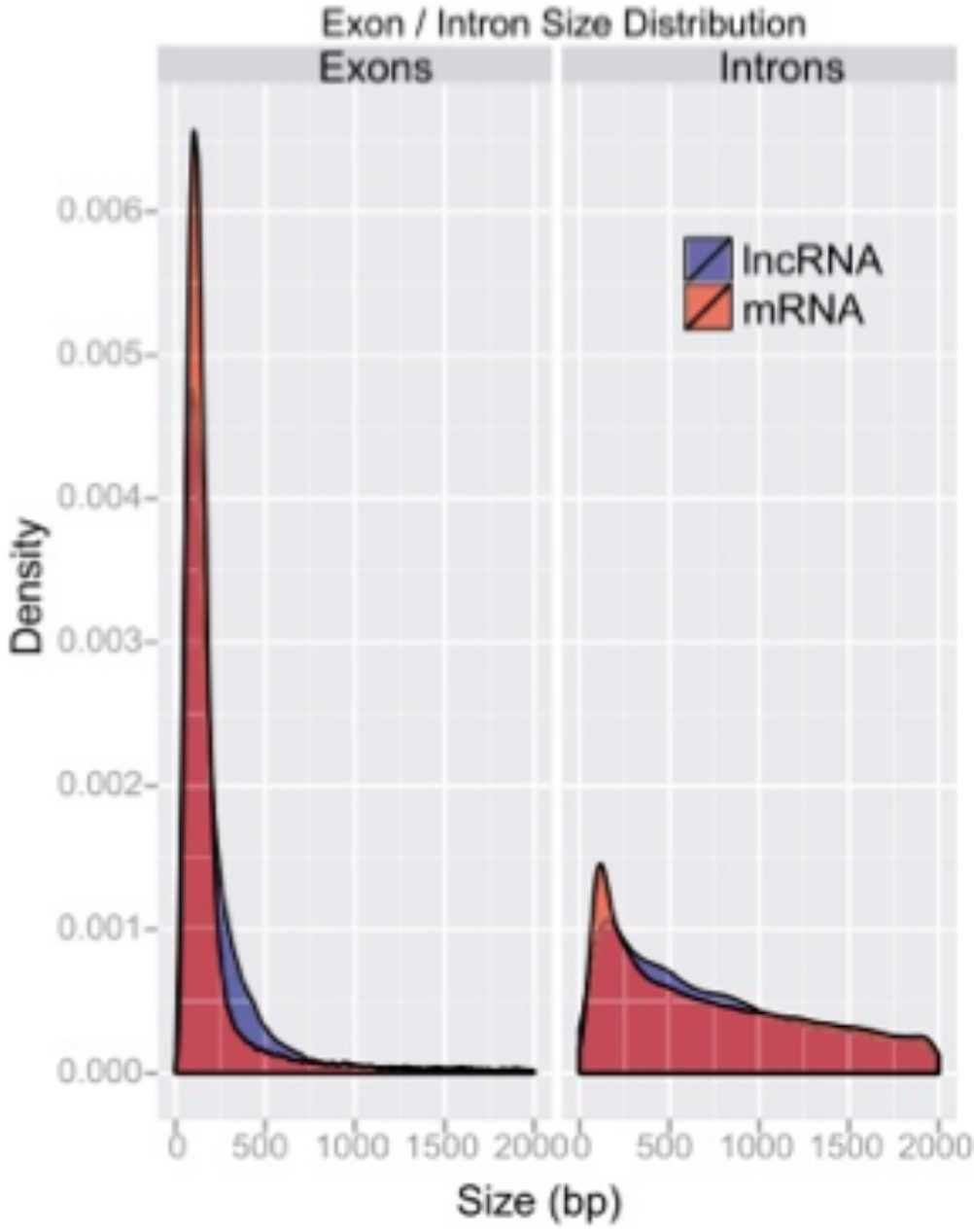
SPLICING:

- **98% of lncRNAs are spliced,**
- **Relevant Splice-site prerequisites at splice donor/acceptor are conserved**
- **Remarkable tendency of lncRNAs to have only 2 exons (42%, mRNAs: 6%)**

BUT: LncRNAs contain fewer INTRONs!! most lncRNA have only 1 intron!!!; mRNAs 3-6

COMPARING mRNAs - lncRNAs

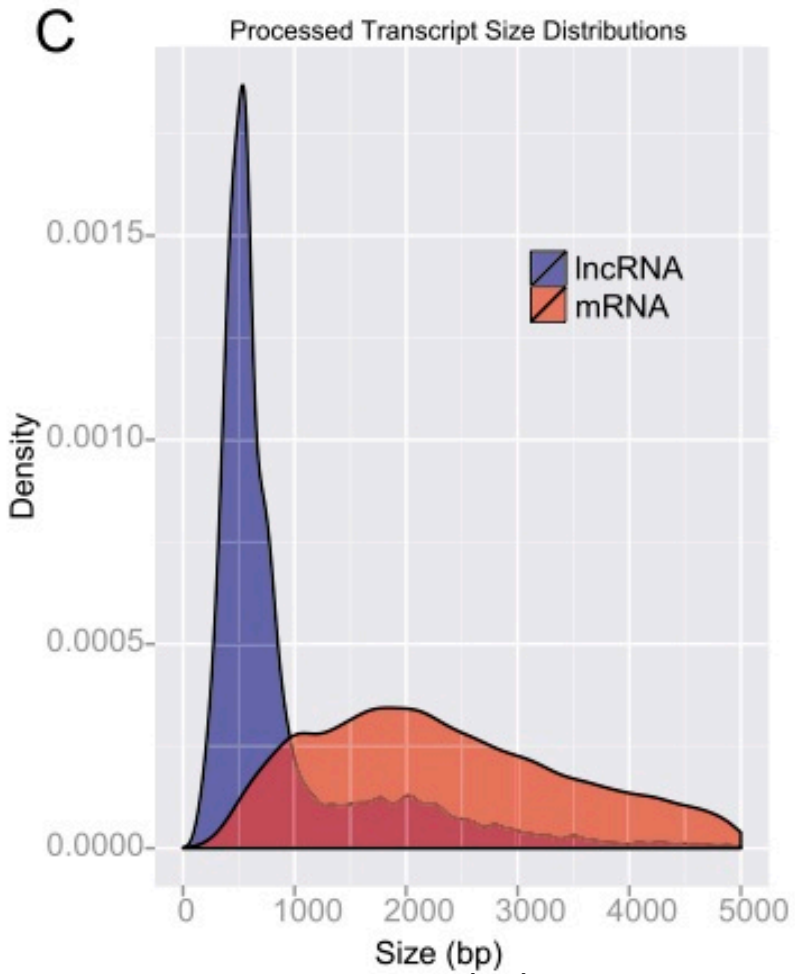
Introns/Exons from lncRNAs are slightly longer



EXON INTRON LENGTH

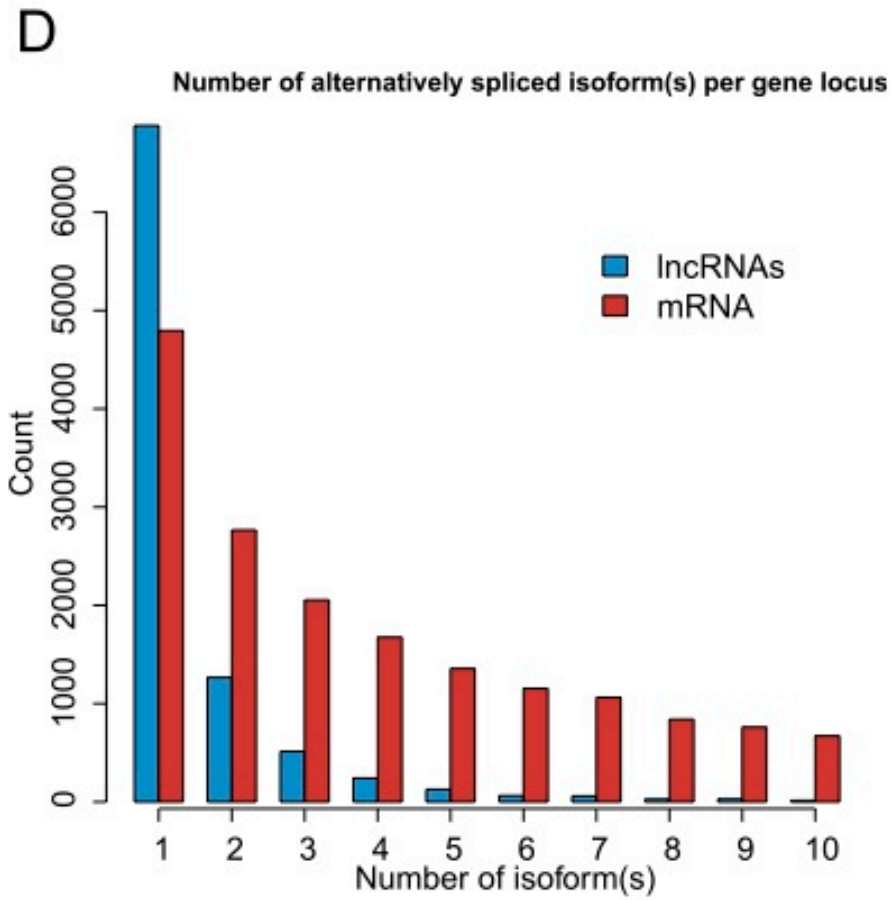
COMPARING mRNAs - lncRNAs

LENGTH



- lncRNAs are on average much shorter: ca. 500nt
- mRNAs are longer and have wider size distribution

ISOFORMS



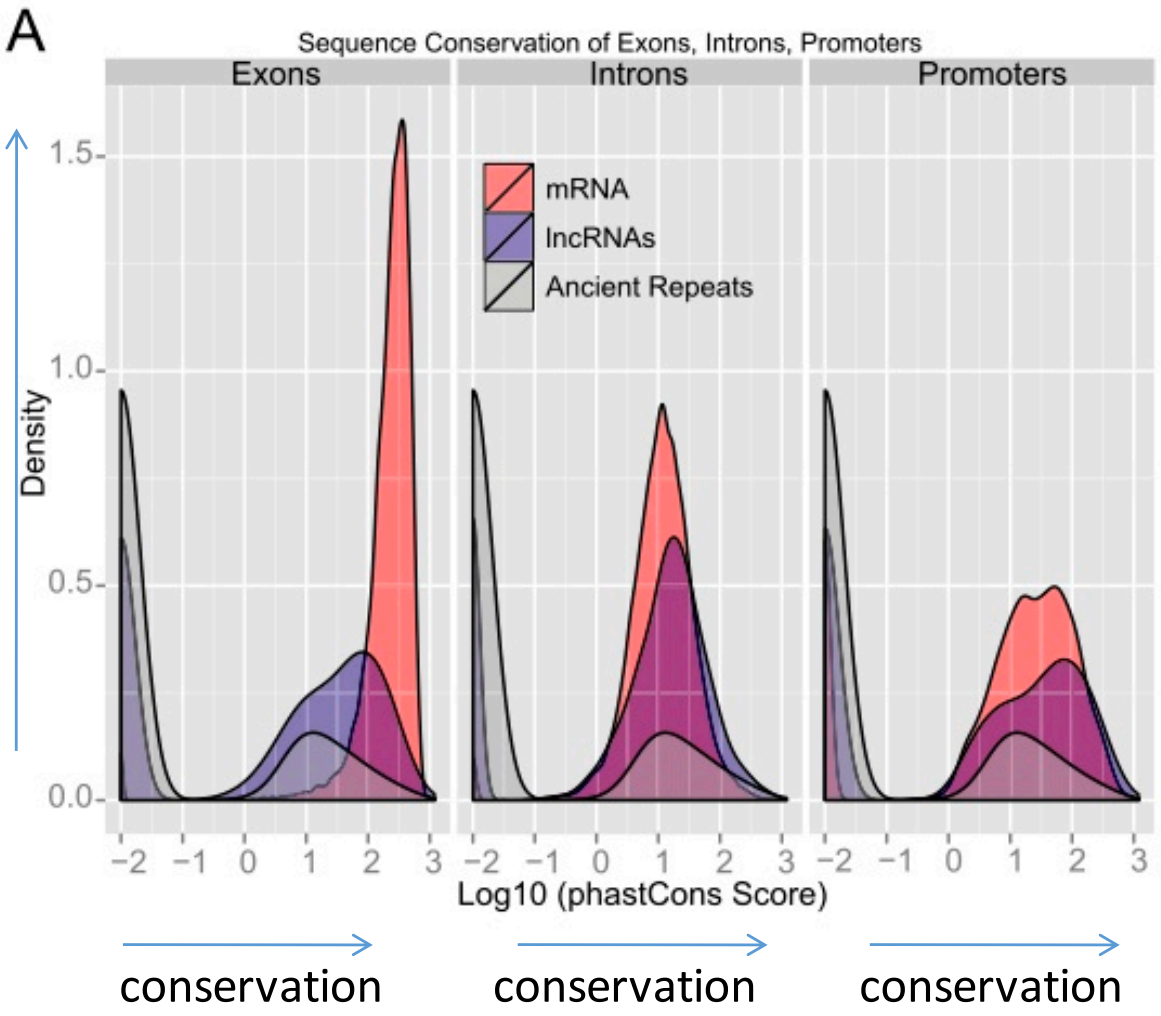
- lncRNAs are uniform → little alternative splicing
- mRNAs: large variety of alternative splicing

COMPARING mRNAs - lncRNAs

CONSERVATION

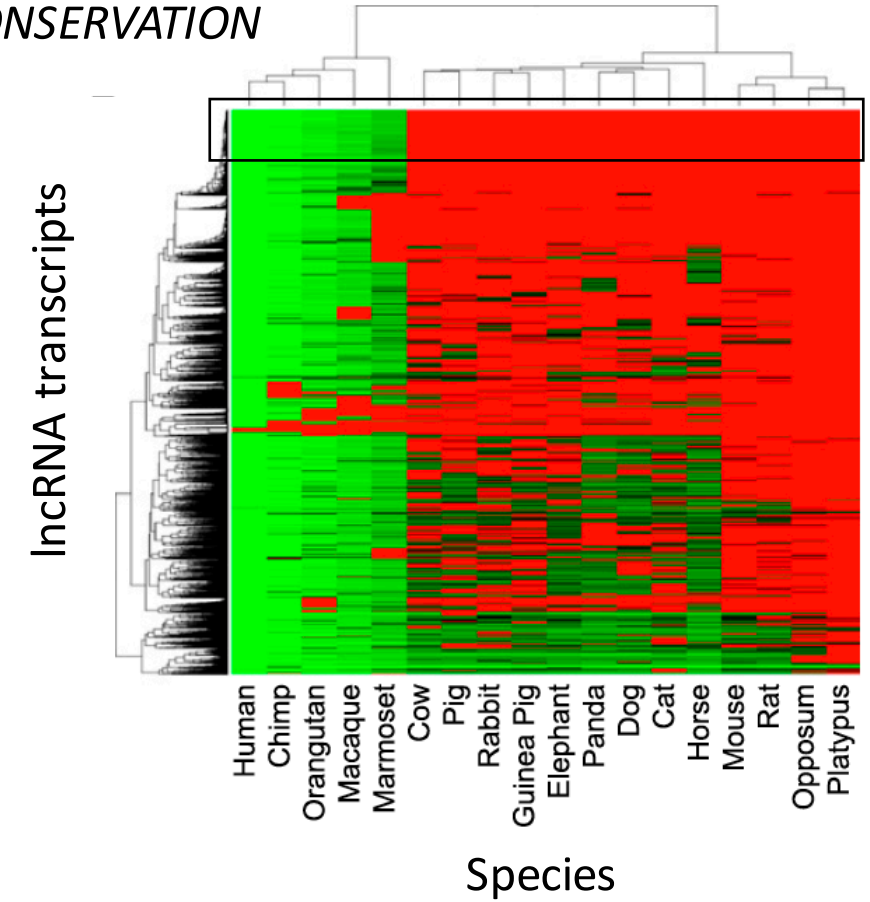
EXONS:
mRNA: high conservation
lncRNA: reduced conservation
But: conservation is higher than mRNA intron conservation

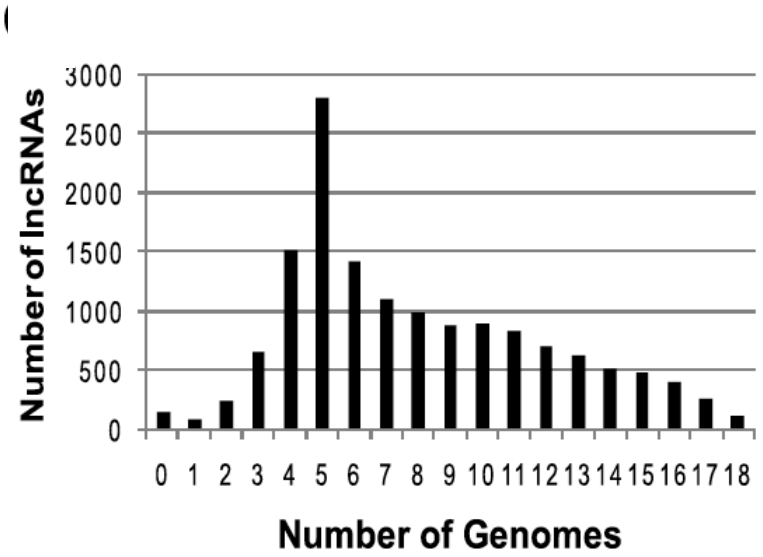
INTRONS:
mRNA: higher conservation than lncRNAs



Evolutional conservation of lncRNAs

CONSERVATION





BLAST of human lncRNAs against all available mammalian genomes and identify homologous genes in other species:
30% of lncRNA transcripts (n = 4546) appear to be primate specific (green=high conservation).

Most lncRNA transcripts are conserved in 4-6 species

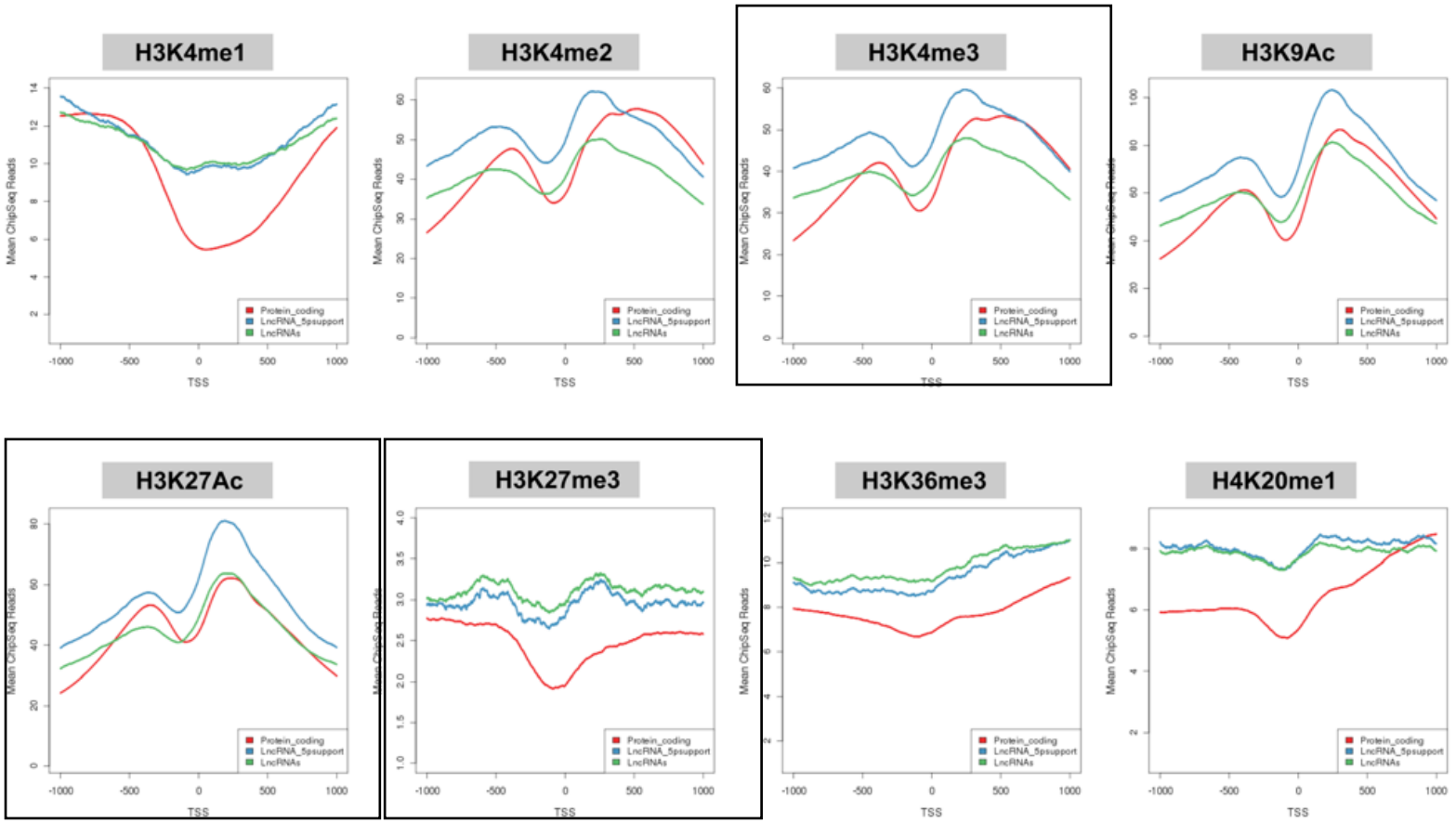
A total of **0.7%** (101) of transcripts appear to be **specific to the human** lineage.
A similar number (134; **1.0%**) is **found in all of the 18 species** analyzed here.

lncRNAs show relevant species specificity

COMPARING mRNAs - lncRNAs

Chromatin signatures at transcriptional start sites are conserved between mRNAs And lncRNAs: MOST lncRNAs ARE TRANSCRIBED BY RNA Pol II

EPIGENTICS



- mRNA
- lncRNA 5' experimental support
- lncRNA

Chromatin signatures around TSS of protein-coding and lncRNA transcripts expressed in the same cell lines where the signatures were monitored by ChipSeq. Shown on the y-axis is the average density of reads covering the TSS of various gene sets, with position plotted on the x-axis (bp relative to positive strand TSS). Protein coding genes are plotted in red, Gencode v7 lncRNAs in green, and lncRNAs with 5′ experimental support (n=2,793) in blue. N.B. A more extensive analysis of histone modifications in multiple cell types is available at http://big.crg.cat/bioinformatics_and_genomics/lncrna_data.

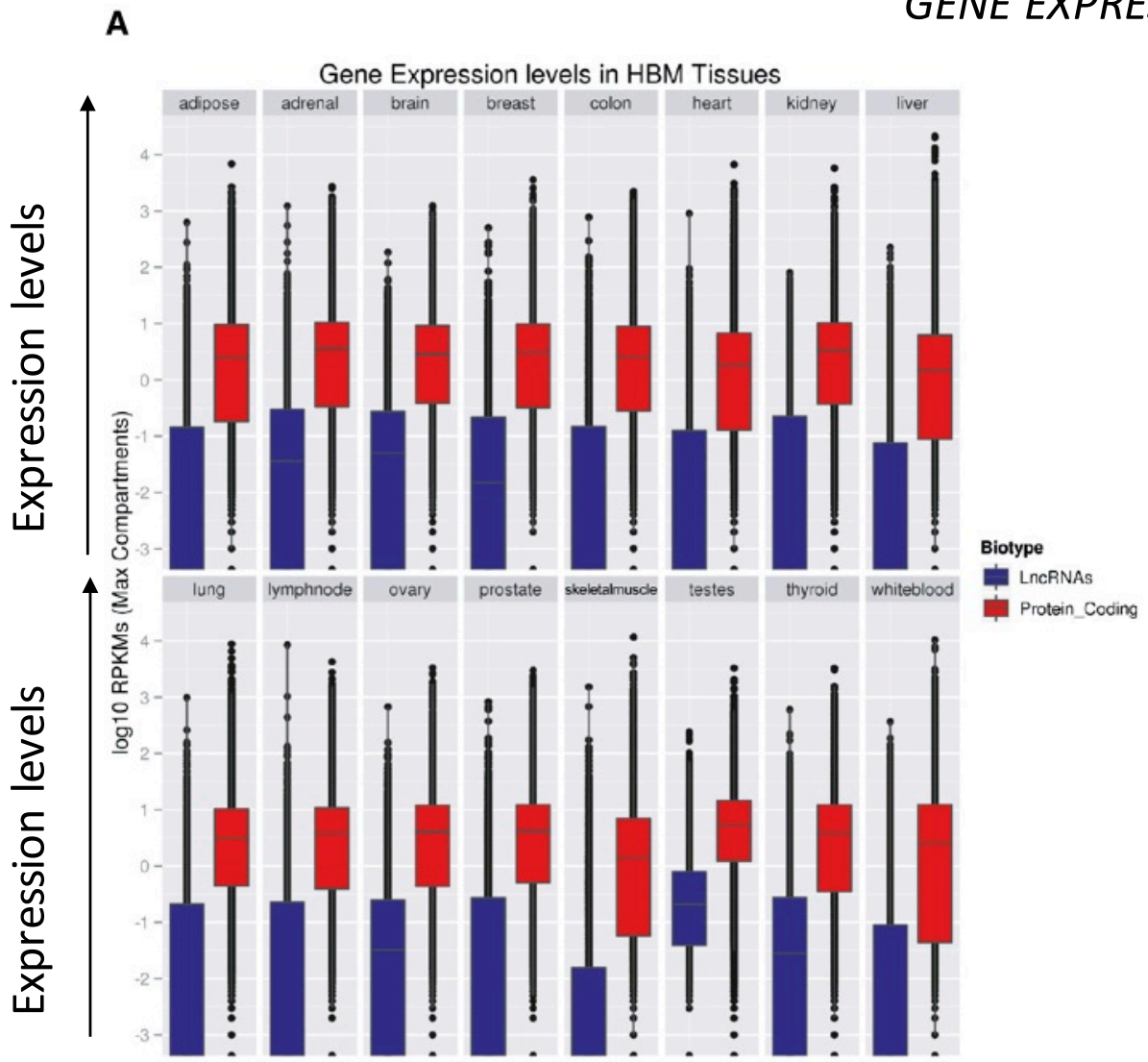
K562 human immortalised myelogenous leukemia cell line

NOTE: Histone exclusion at TSS: (Not shown: RNA Pol II peaks at transcription start site of most lncRNAs)

COMPARING mRNAs - lncRNAs

GENE EXPRESSION

Absolute expression of lncRNAs is much lower than mRNA expression

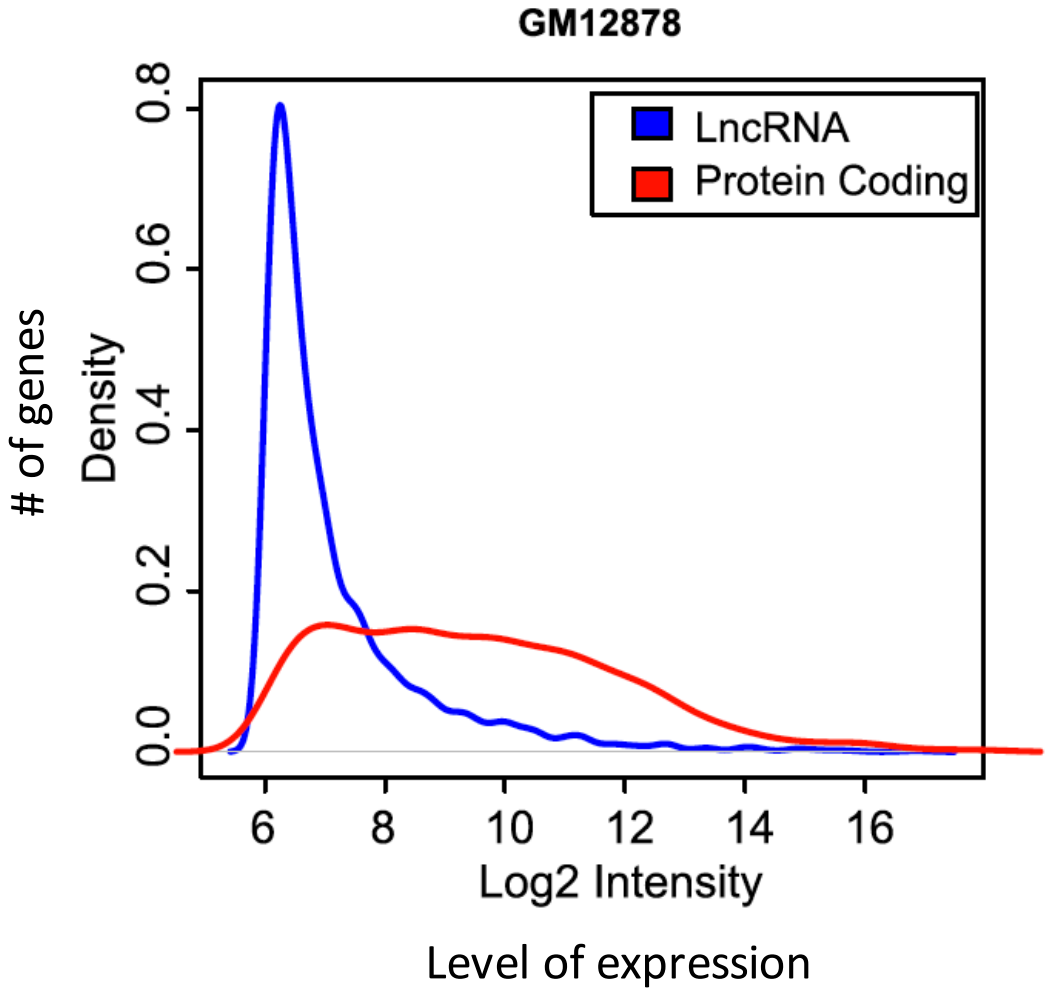


RNA-seq

COMPARING mRNAs - lncRNAs

GENE EXPRESSION

Absolute expression of lncRNAs is much lower than mRNA expression



Level of expression

Microarray data

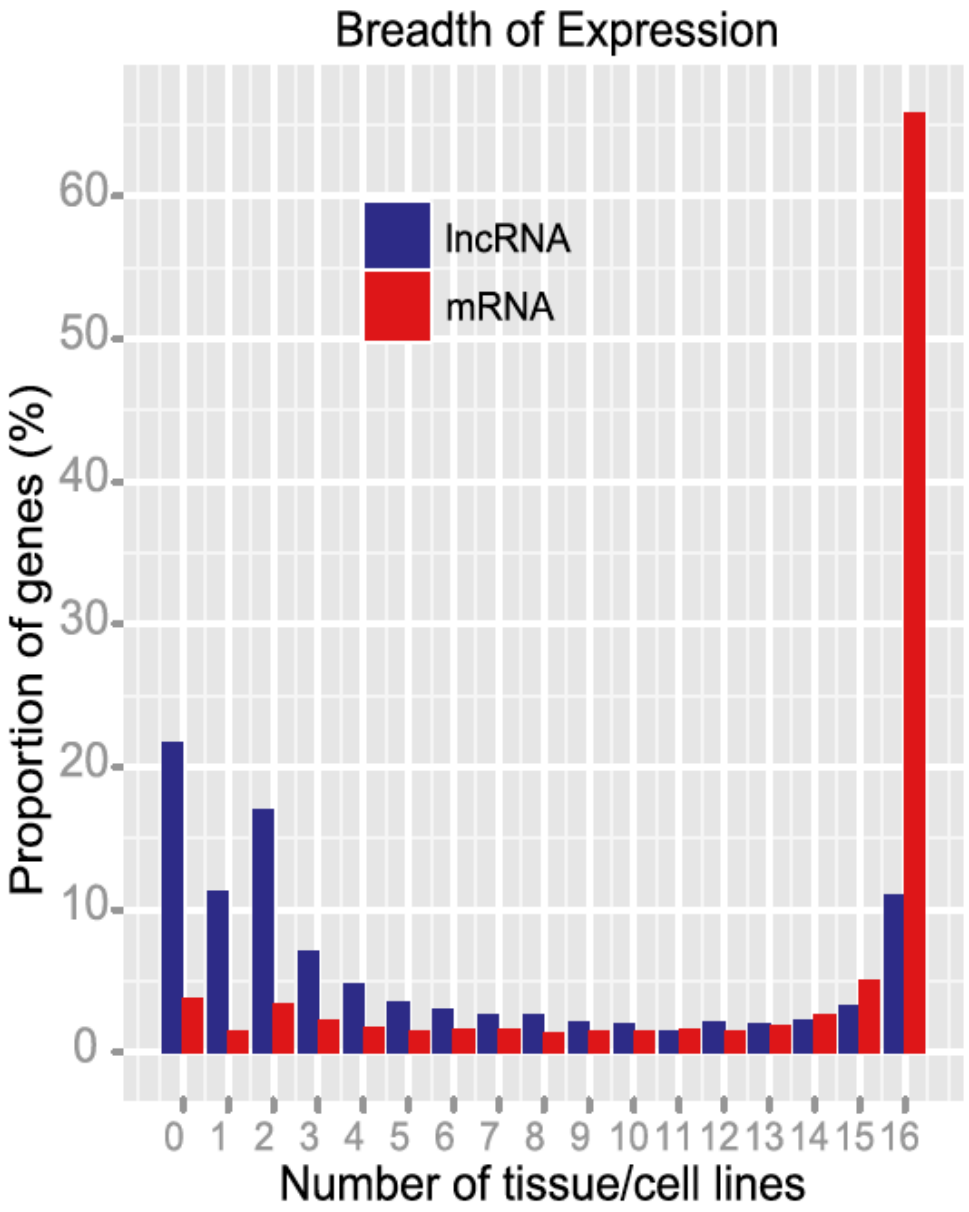
COMPARING mRNAs - lncRNAs

SPECIFICITY OF GENE EXPRESSION

mRNAs show broad expression in different tissues –
65% of mRNAs are expressed in 16 tissues

lncRNAs show tissue/cell specific expression –
50% of lncRNAs expressed in 4-5 tissues.

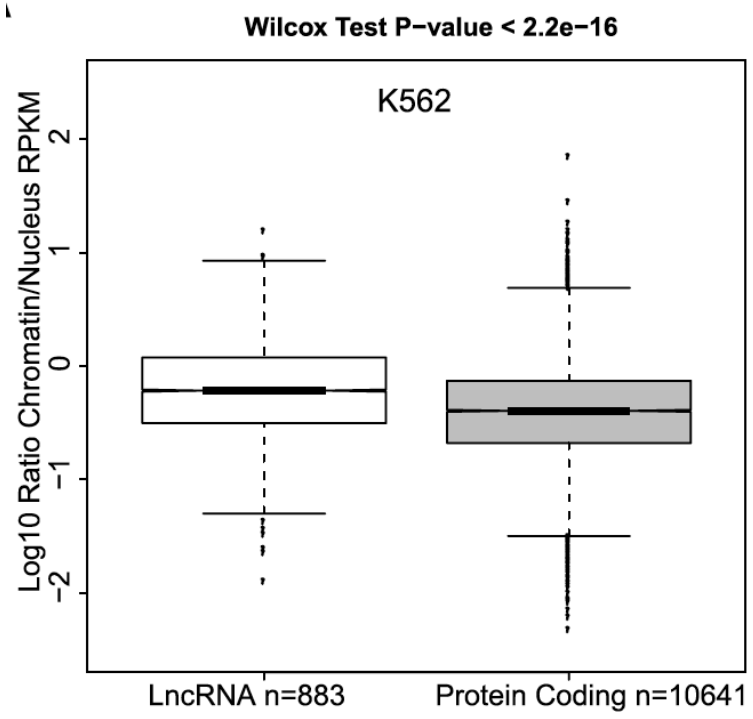
HIGH VALUE OF lncRNAs AS BIOMARKER – EXPRESSION SIGNATURES IN CANCER



COMPARING mRNAs - lncRNAs

LOCALIZATION

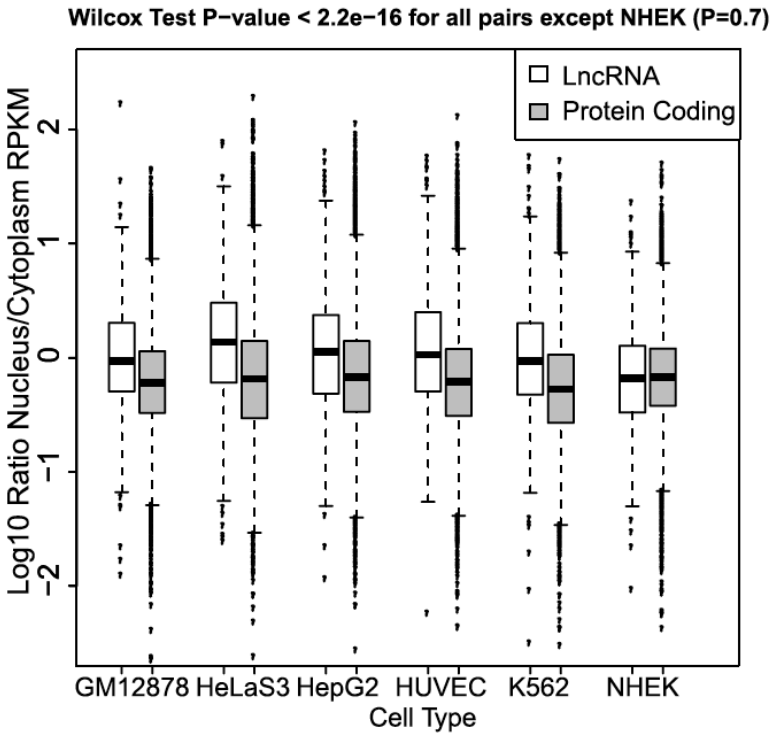
Determination of transcript abundance
Chromatin/nucleus



lncRNAs locate more frequently to chromatin than mRNAs

Data: RNA seq of RNAs located in the nucleoplasma and sticking to chromatin

Determination of transcript abundance
Nucleus/cytoplasma



lncRNAs locate more frequently to the nucleus than mRNAs

Data: RNAseq of RNAs located in the nucleoplasma and cytoplasma

COMPARING mRNAs - lncRNAs

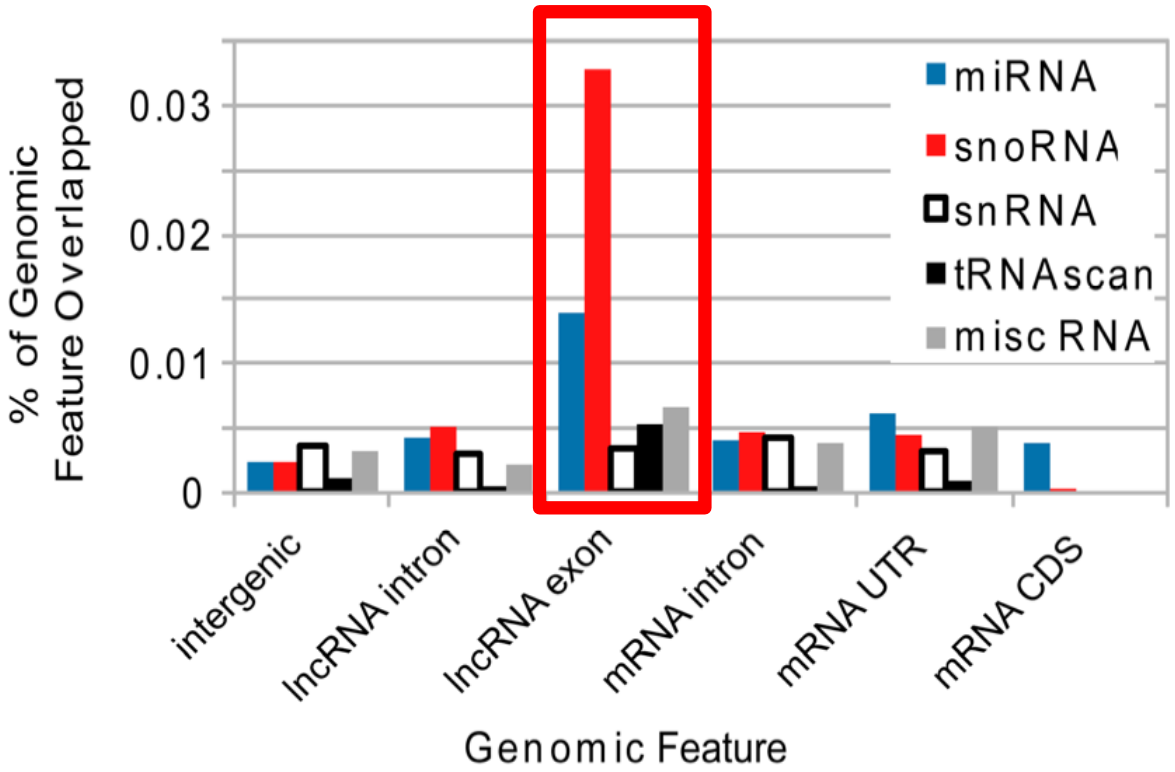


Figure S7: LncRNAs as a source of small RNAs. Shown is the proportion of nucleotides in exons/introns/(UTRS) from protein-coding genes and lncRNAs that overlap different classes of annotated small RNAs. Note that all values refer to cases where the small RNA is on the same strand as the indicated lncRNA/mRNA.

lncRNA exons have a remarkable potential to encode functional **small RNAs** (sRNAs)

HIGHER THAN EXONS OF mRNAS!!!!

miRNA: small RNAs that target mRNAs to reduce protein expression

snoRNA: Small nucleolar RNAs (snoRNAs) are a class of small RNA molecules that primarily guide chemical modifications of other RNAs, mainly ribosomal RNAs, transfer RNAs and small nuclear RNAs.

snRNA: Small nuclear ribonucleic acid (snRNA), also commonly referred to as U-RNA, is a class of small RNA molecules that are found within the splicing speckles and Cajal bodies of the cell nucleus in eukaryotic cells.

tRNA: translation

miscRNAs: MiscRNA is short for miscellaneous RNA, a general term for a series of miscellaneous small RNA. It serves a variety of functions, including some enzyme-like catalysis and processing RNA after it is formed. Besides, some of these small RNAs may serve as switches.

COMPARING mRNAs - lncRNAs

lncRNAs represent a big class of functional elements that
- show controlled gene expression
- are processed
- lack protein coding potential
- defined localization
- frequently encode small RNAs
- low overall conservation
- 35% of lncRNAs are primate specific
- expression is rather low – but controlled!
- mostly transcribed by RNA Pol II

FUNCTION: FOR THE VAST MAJORITY OF lncRNAs THE BIOLOGICAL FUNCTION IS UNKNOWN!!!!