

lncRNAs

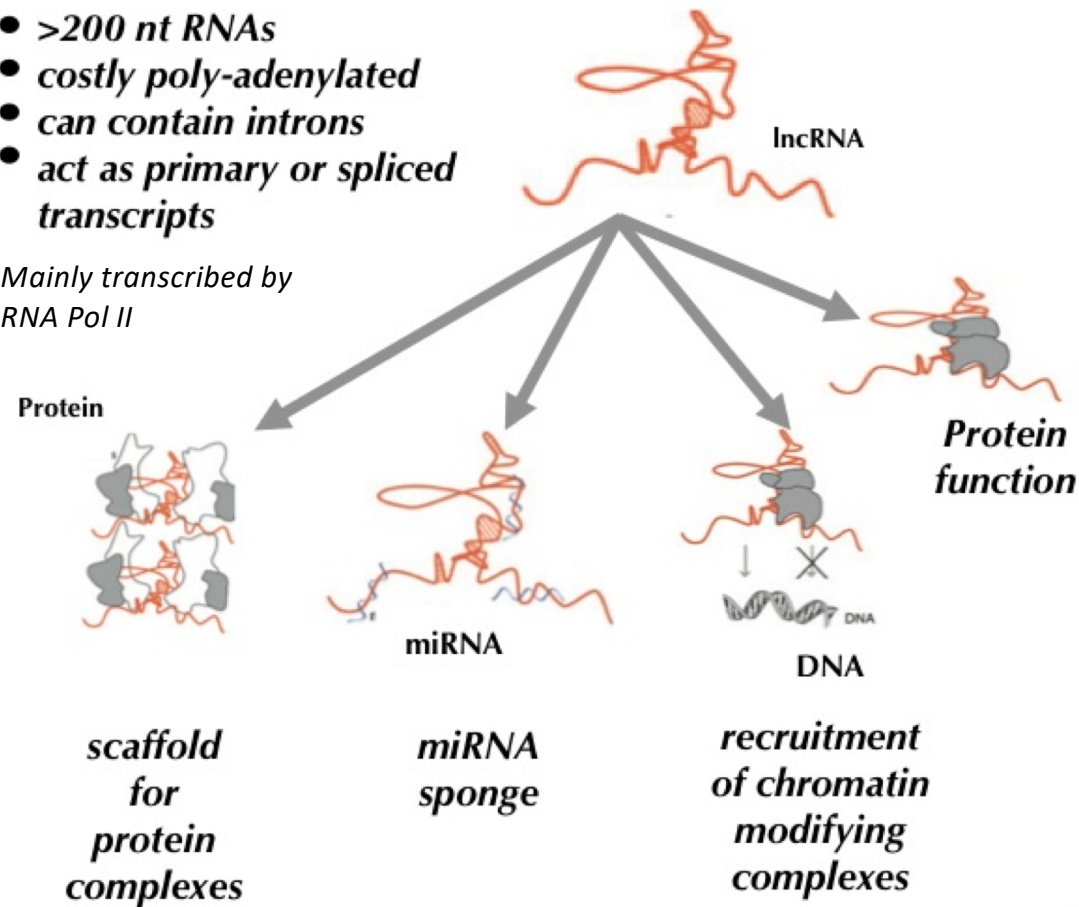
long, non-coding RNAs

Characteristics of lncRNAs

Long, non-coding RNAs (lncRNAs)

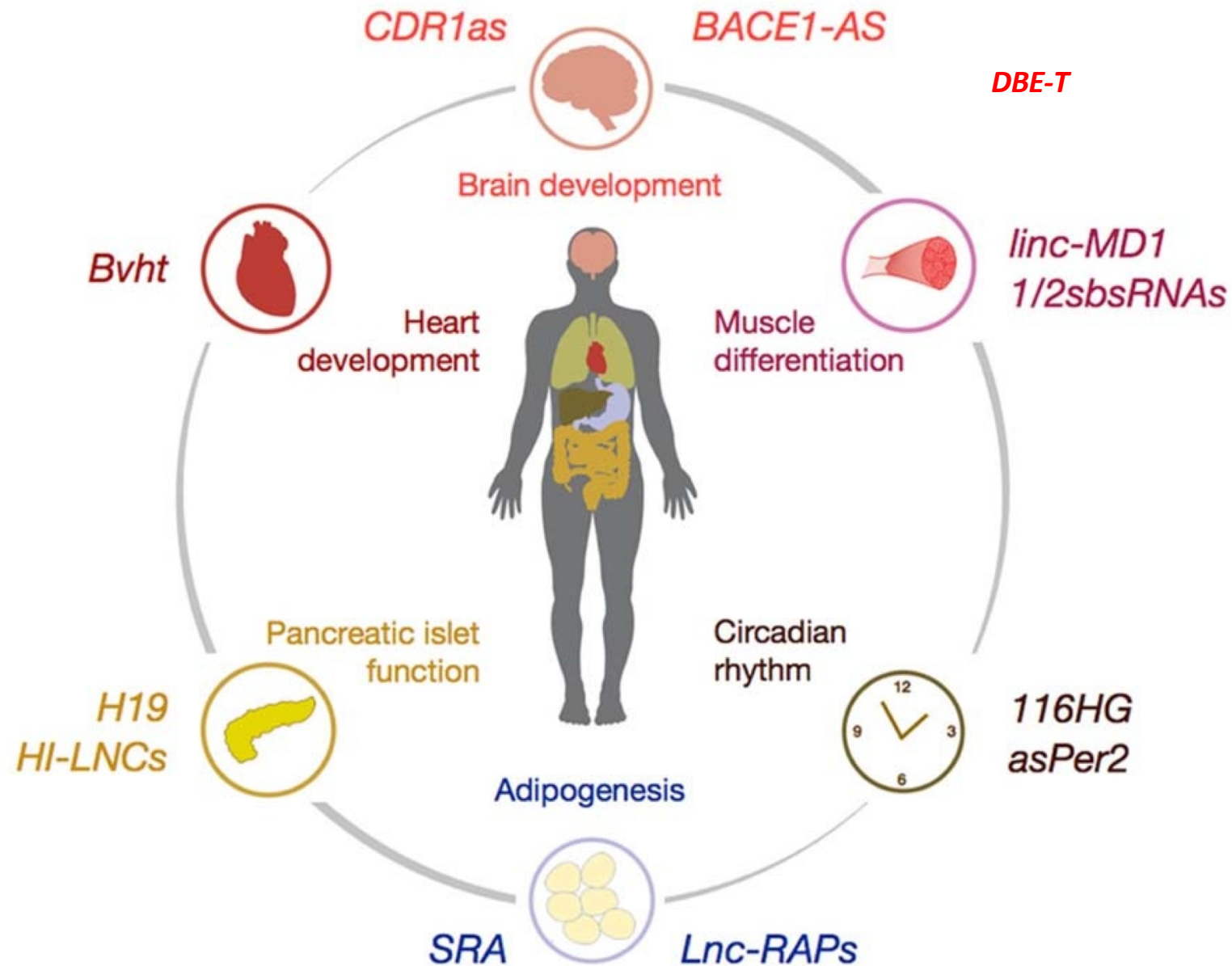
- *>200 nt RNAs*
- *costly poly-adenylated*
- *can contain introns*
- *act as primary or spliced transcripts*

*Mainly transcribed by
RNA Pol II*



Nuclear and/or
cytoplasmatic
localization

lncRNAs in physiology



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 Xiaolan 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}

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

COMPARING mRNAs - lncRNAs

lncRNAs can be

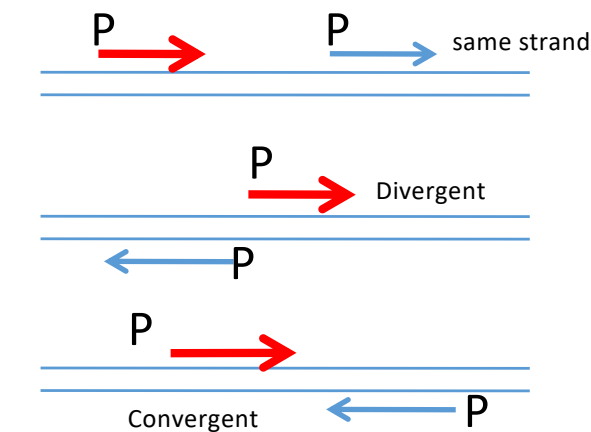
1. Intergenic lncRNAs (lincRNA):

Do not intersect with protein coding gene

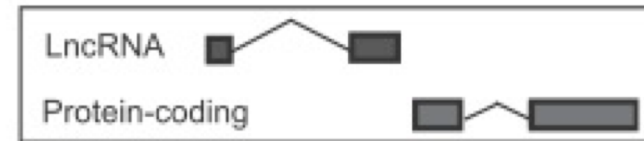
2. Genic lncRNAs:

Intersect a protein coding gene

- Exonic
- Intronic
- overlapping

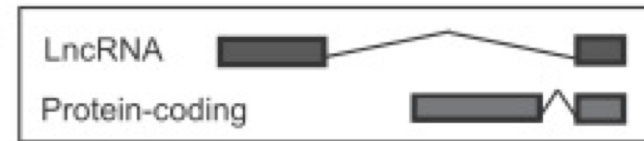


Intergenic lncRNA

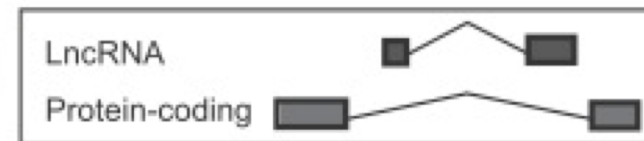


Genic lncRNA

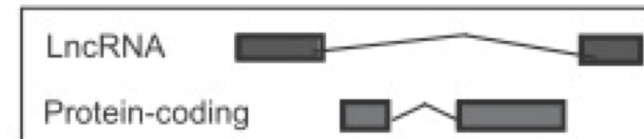
Exonic



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	

P: promoter

Why?

Why?

Why?

COMPARING mRNAs - lncRNAs

A. lncRNAs do not have coding potential:

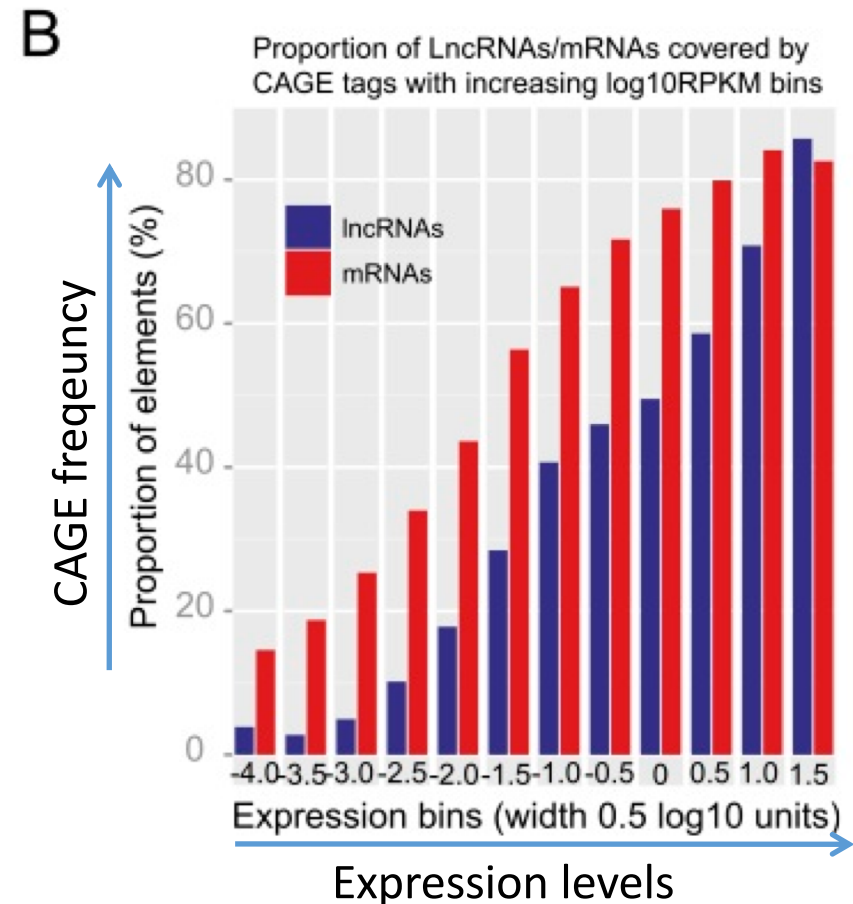
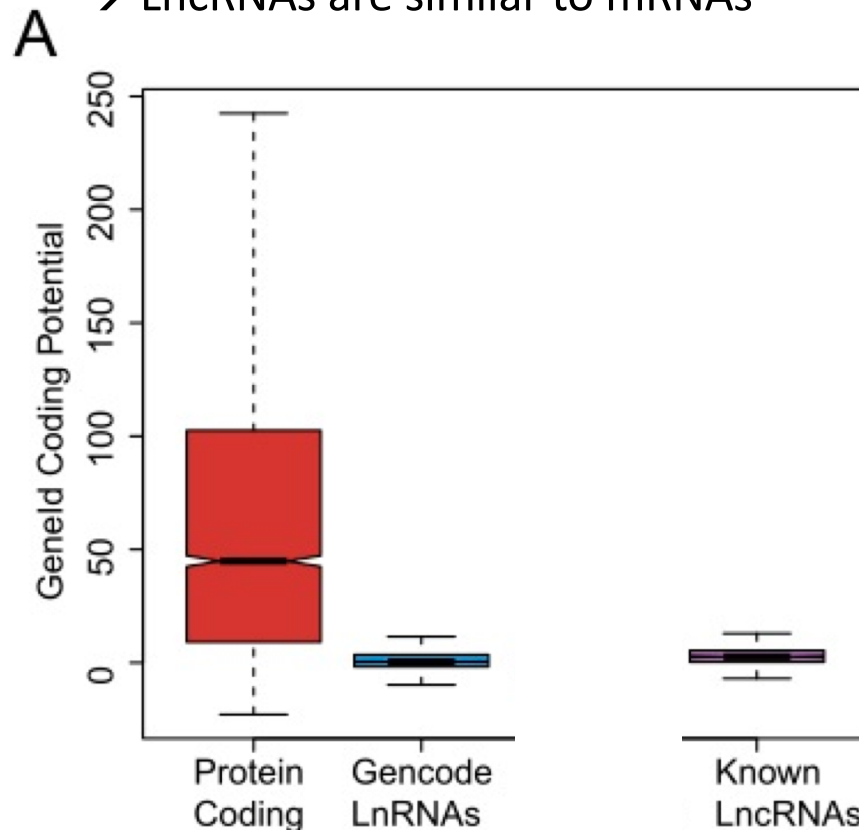
→ Longest possible ORF was searched in mRNA/lncRNAs

B. CAGE tags can be found in lncRNAs → defined transcriptional start site

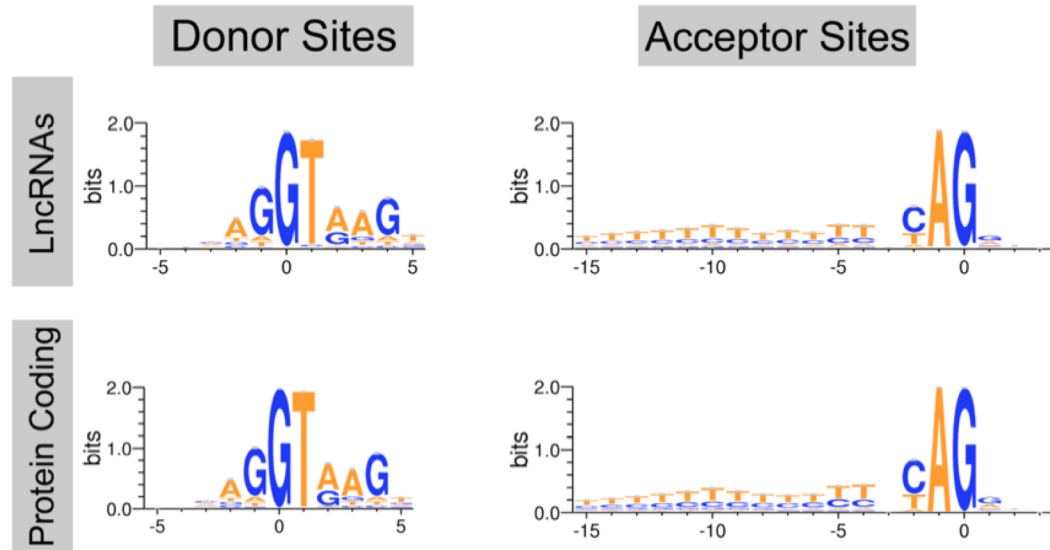
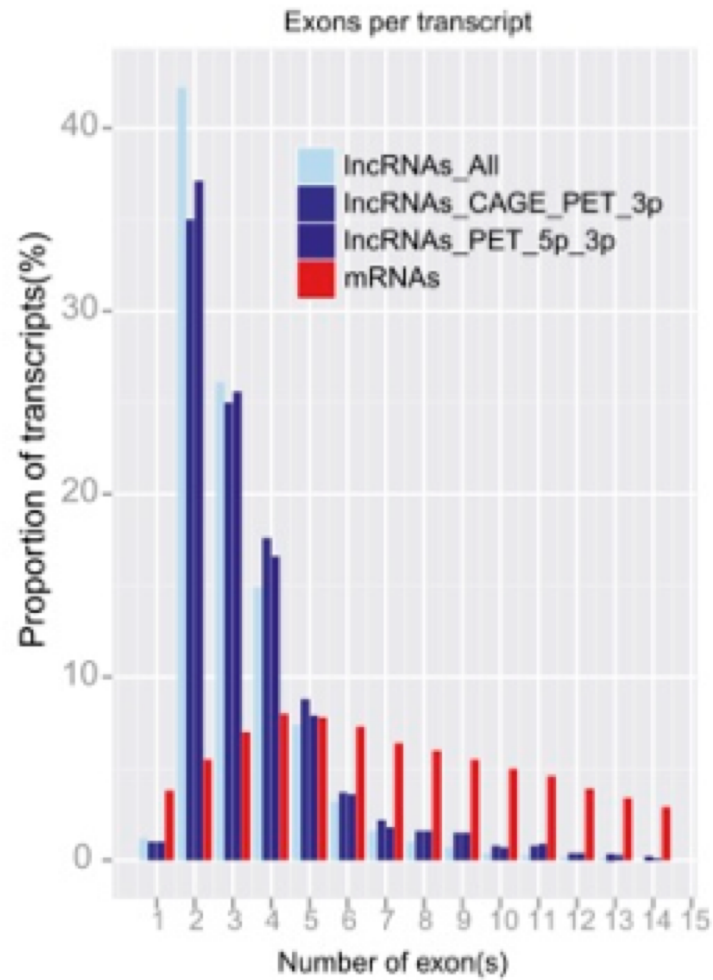
CAGE tag frequency increases with increased lncRNA expression levels.

mRNAs are characterized by more CAGE tags

→ lncRNAs are similar to mRNAs



COMPARING mRNAs - lncRNAs



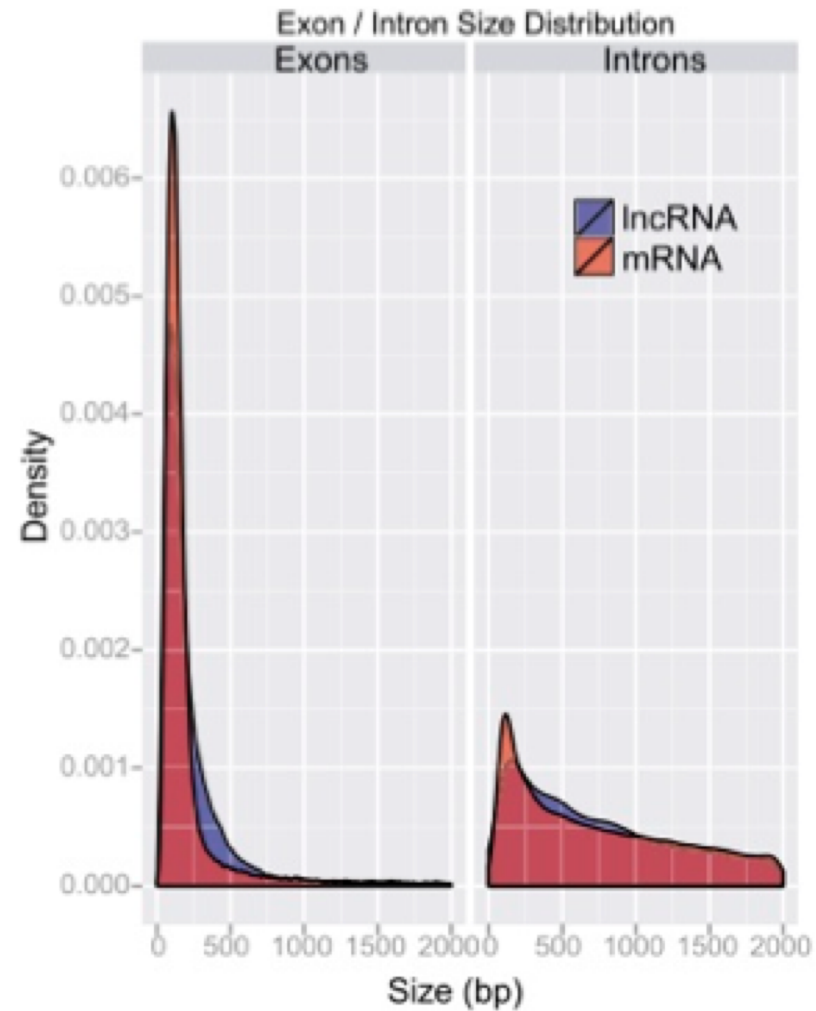
SPLICING:

- **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%)

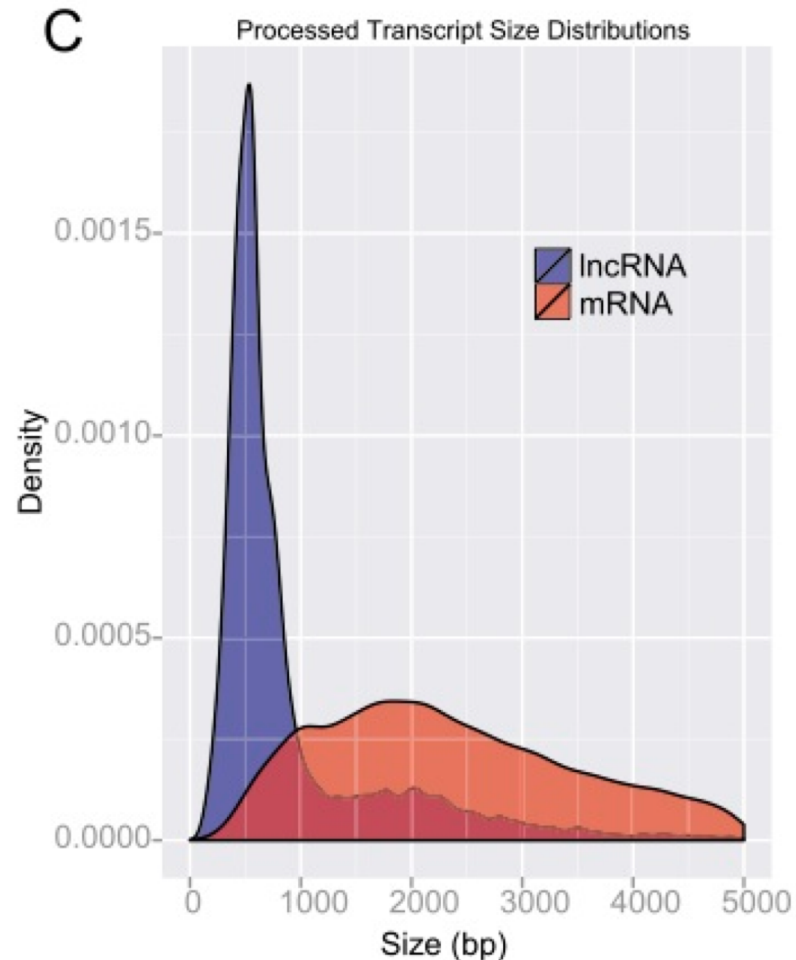
BUT: lncRNAs contain fewer INTRONS!! most lncRNA have only 1 intron!!!; mRNAs 3-6

COMPARING mRNAs - lncRNAs

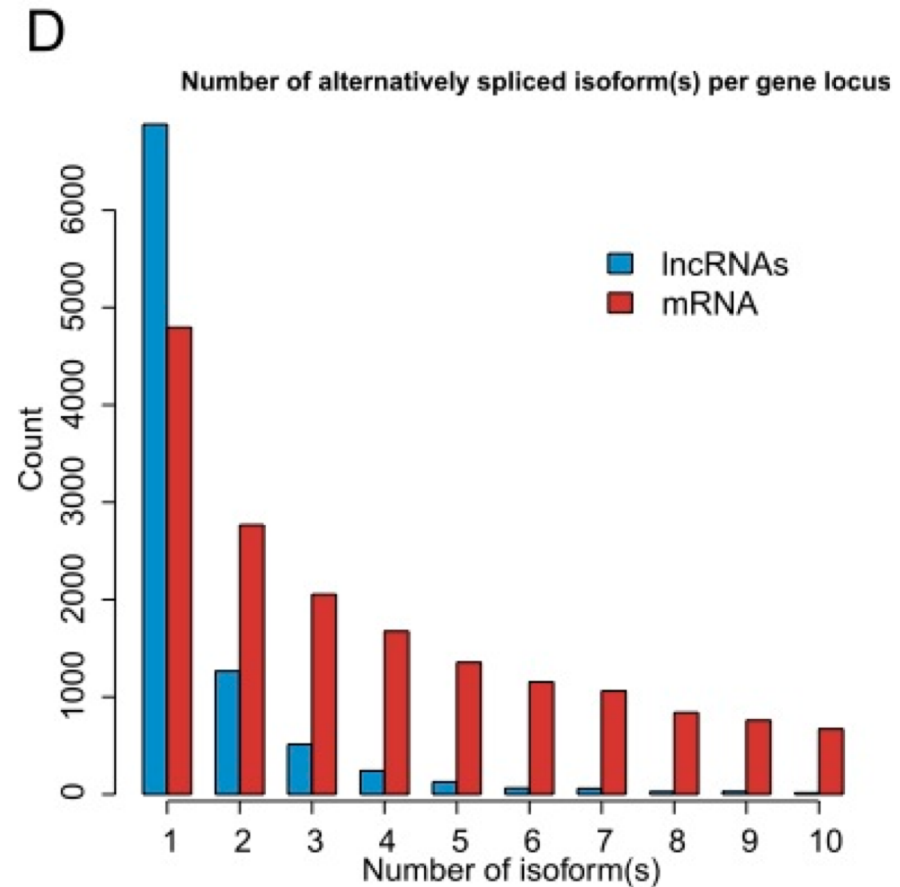
Introns/Exons from lncRNAs are slightly longer



COMPARING mRNAs - lncRNAs



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



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

COMPARING mRNAs - lncRNAs

EXONS:

mRNA: high conservation

lncRNA: reduced conservation

But: conservation is higher than mRNA intron conservation

INTRONS:

mRNA: higher conservation than lncRNAs

