

RNA Genomics

BME 110: CompBio Tools

Todd Lowe

May 17, 2012

In Genomes, Two Types of Genes

Protein-coding:

[Start codon] [codon 1] [codon 2] [...] [Stop codon]

+ DNA codons translated to amino acids to form a protein

Non-coding RNAs (NcRNAs)

No consistent patterns common to all RNA genes

+ Not translated to proteins, functional as RNA molecules

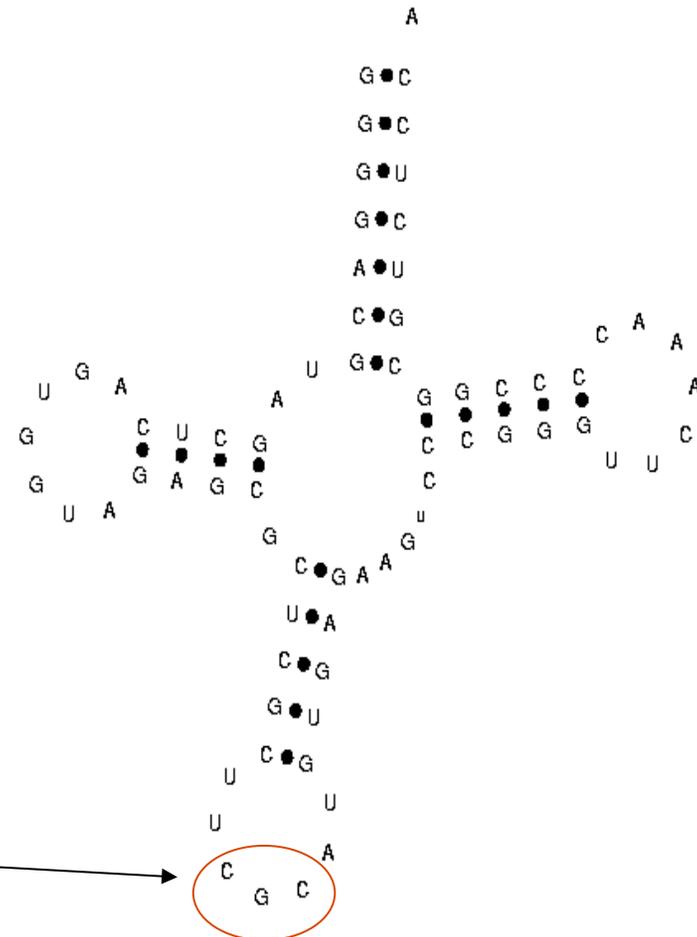
Rules of Engagement

- Alphabet: A, C, G, U (T~U)
- Canonical pairings: A-U, G-C
- Non-canonical: G-U (very common), A-C, A-G, etc.
- ncRNAs are riddled with nucleotide modifications which may affect H-bond base pairing
- Depending on the modification, can make non-canonical pairings more *or* less favorable
Example: wobble base modif. in tRNAs

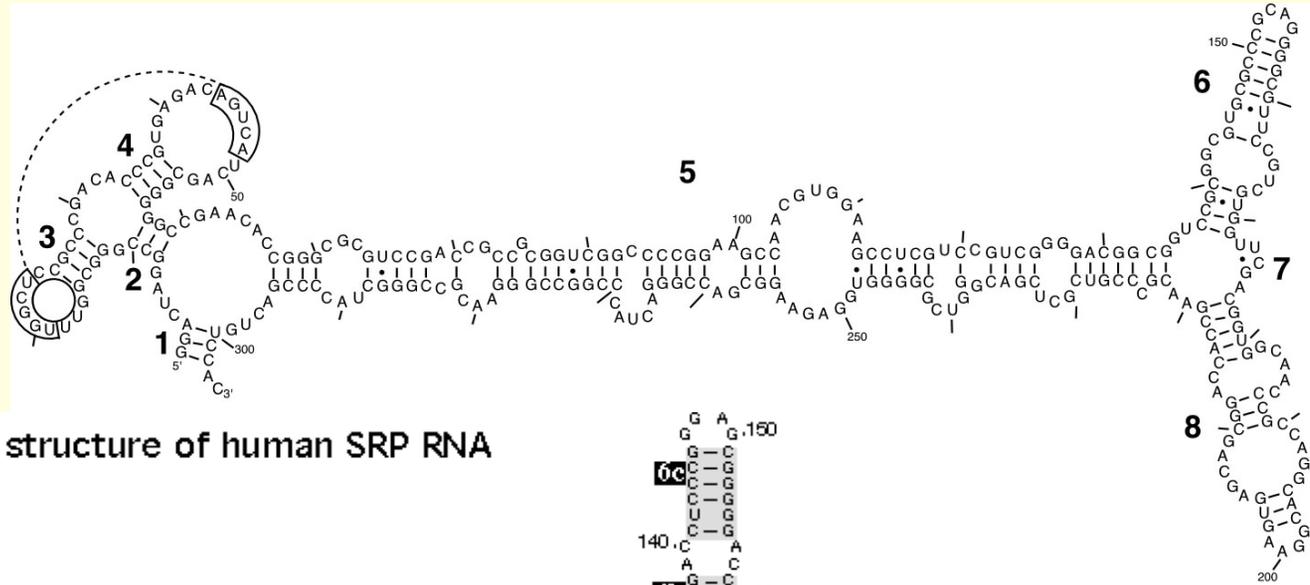
Features of ncRNAs

- Secondary structure is conserved, but often not primary sequence
- Additional sources of functional information easily derived from sequence and/or structure (i.e. anticodon in tRNA)

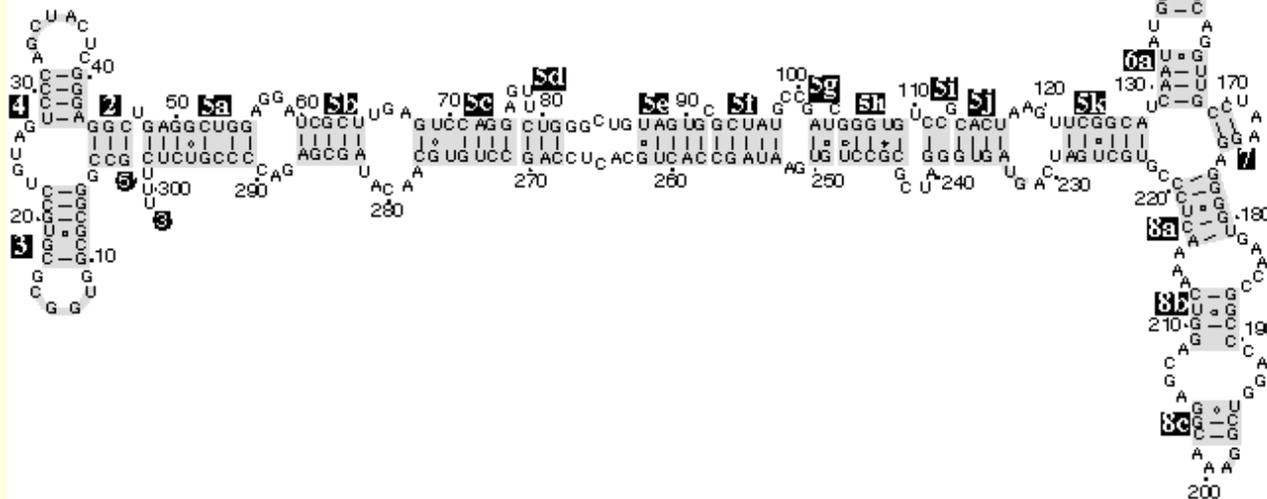
Drosophila-tRNA Ala (CGC) 74.79 bits



SRP RNA Orthologs: Same Function, Similar Structure, < 70% sequence identity



Secondary structure of human SRP RNA



Halobacterium halobium SRP RNA
(SRPDB, March 10, 2000)

Common Attributes of ncRNAs

- Structural – many fold into complex 3-D structures
- Antisense – many form specific base pairings with external “target” RNAs, no structure required for function
- Catalytic – some catalyze biochemical reactions “ribozymes” (RNaseP, group I & II introns, hammerhead RNA, & others)
- Regulatory – some interact with DNA or RNA targets for gene regulation (microRNAs, many bacterial small RNAs)

RNA + Proteins = Complexes

- RNAs usually do not act alone, but are complexed with proteins
- Sites of interaction with RNAs or proteins exert selective pressure to conserve sequence or structure
- Sites of interaction make up conserved motifs one looks for to help identify RNA and/or determine function

How To Keep Track of All Types of RNA?

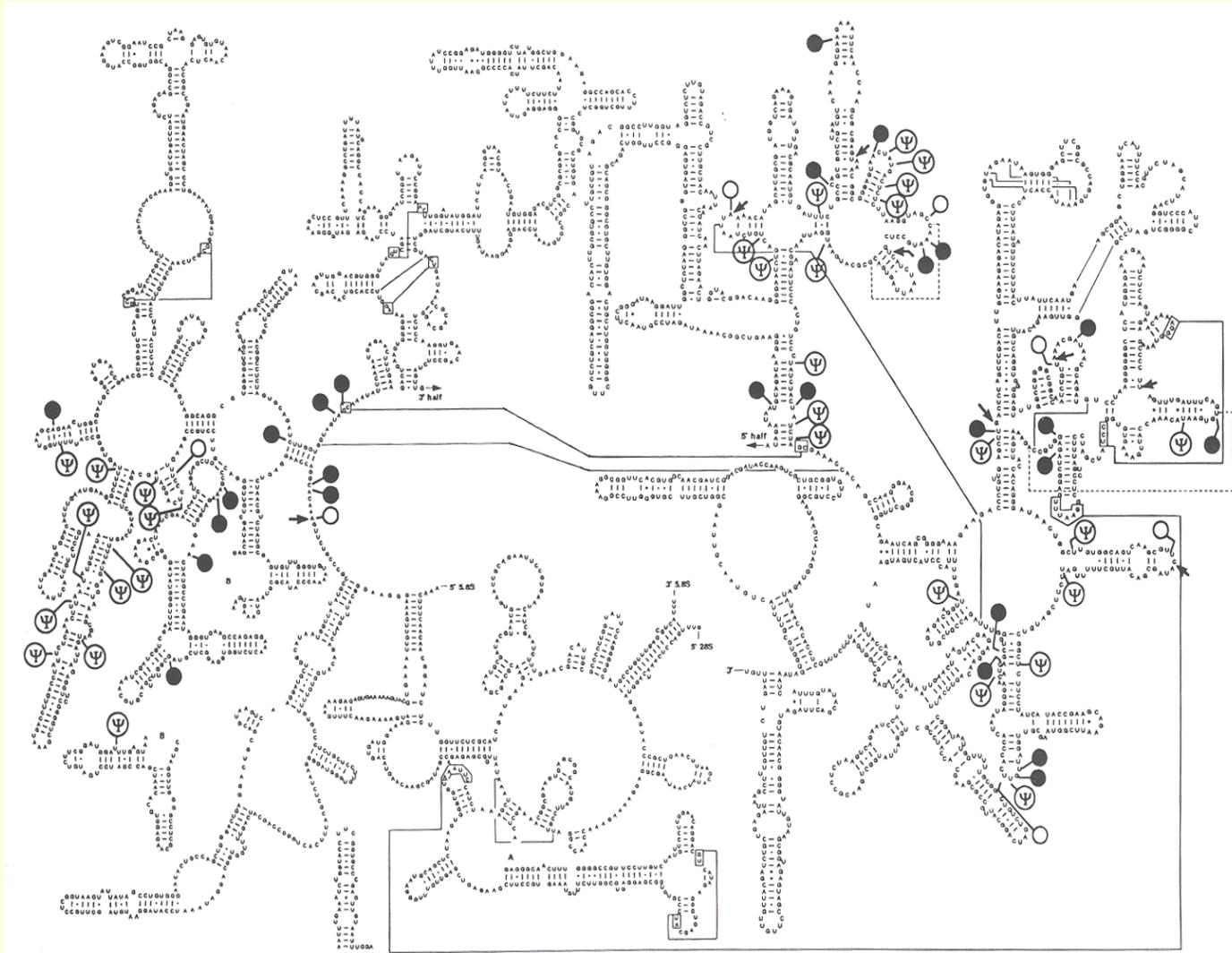
- General Database: RFAM
<http://rfam.sanger.ac.uk/>
- Attempt to catalogue, align, and create search models for all known RNAs
- Just like “PFAM” (protein sister site) can paste your sequence in and look for matches to RNA gene models
- Nice alignments, excellent resource
- Sometimes not as good as customized search programs, when they exist

Examples of ncRNA Genes

Ribosomal RNAs (rRNAs) – protein translation

- Examples : large subunit rRNA (aka LSU/28S/25S/23S), small subunit rRNA (aka SSU/18S/16S), 5.8S, and 5S rRNA
- Highly structured scaffold for dozens of ribosomal proteins; catalytic role in forming peptide bonds

Baker's Yeast (*S. cerevisiae*) Large Subunit rRNA 2-D structure

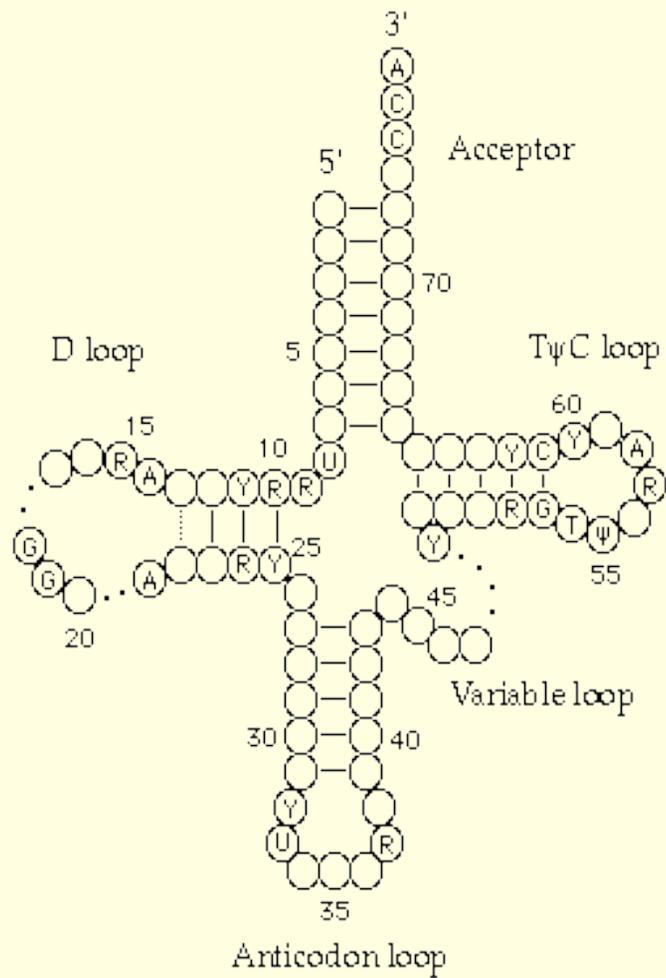


tRNAs: A well-studied ncRNA Gene family

Transfer RNAs (tRNAs) – “decode” mRNA codons into cognate amino acids in protein translation

- Examples: tRNA-AGC(Ala), tRNA-ACC(Gly)...
(62 kinds possible)
- Structured RNA and “antisense” interaction between tRNA anticodon and mRNA codon
- Can find them easily using the tRNAscan-SE web server:
<http://lowelab.ucsc.edu/tRNAscan-SE/>
- Can get example sequences from the Genomic tRNA Database: <http://gtrnadb.ucsc.edu/>

tRNA “cloverleaf” 2-D structure



3-D structure

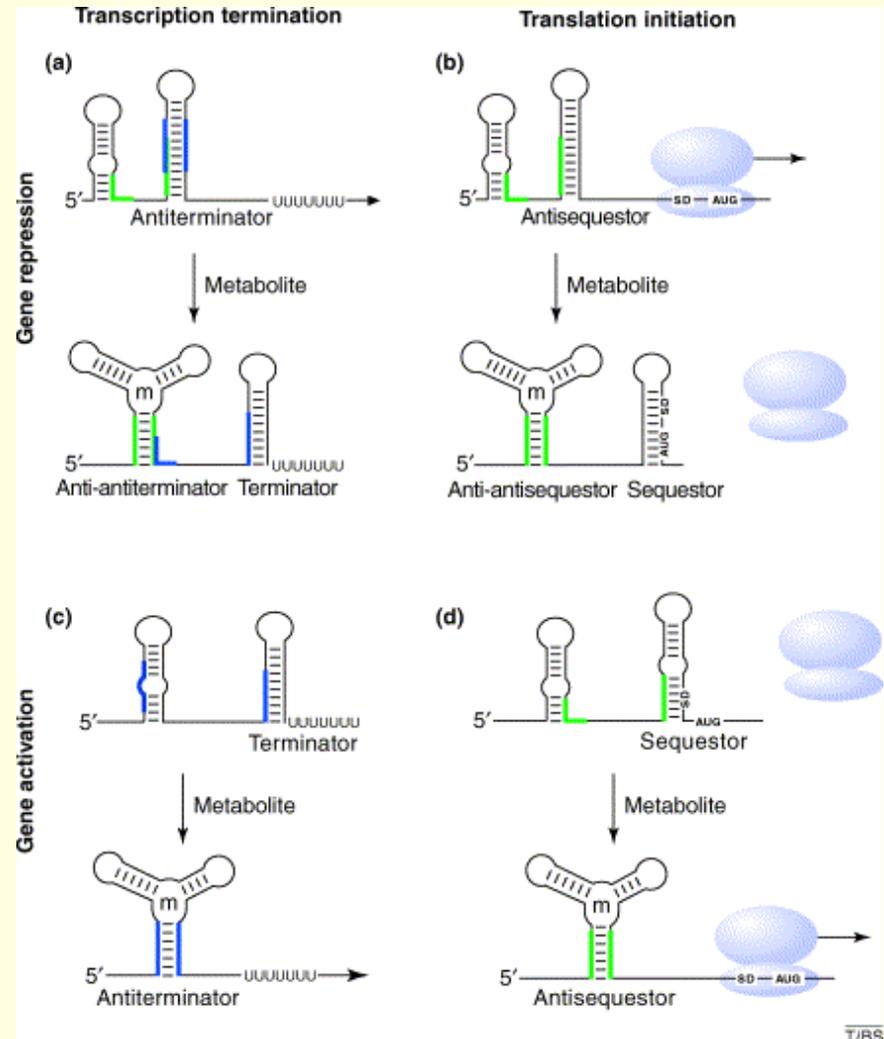


Some Functional RNA Structures are found in mRNAs

- Specific secondary structure can act with regulatory roles
- Examples
 - Riboswitches (example: TPP riboswitch)
 - Selenocysteine insertion sequences (SECIS)
 - Iron Response Element (IRE)

Many Different Types of Riboswitches

- Functional RNA is part of mRNA molecule, *usually* at 5' end
- Riboswitch functions by binding a specific small molecule (i.e. free lysine, or free purine nucleotides)
- Binding state of riboswitch modulates on/off control of transcription or translation of attached gene



From Nudler & Mironov, *TIBS*
29:11-17, 2004.

Riboswitch Web Analyses

Ribex:

<http://www.ibt.unam.mx/biocomputo/ribex.html>

Riboswitch Finder:

<http://riboswitch.bioapps.biozentrum.uni-wuerzburg.de/>

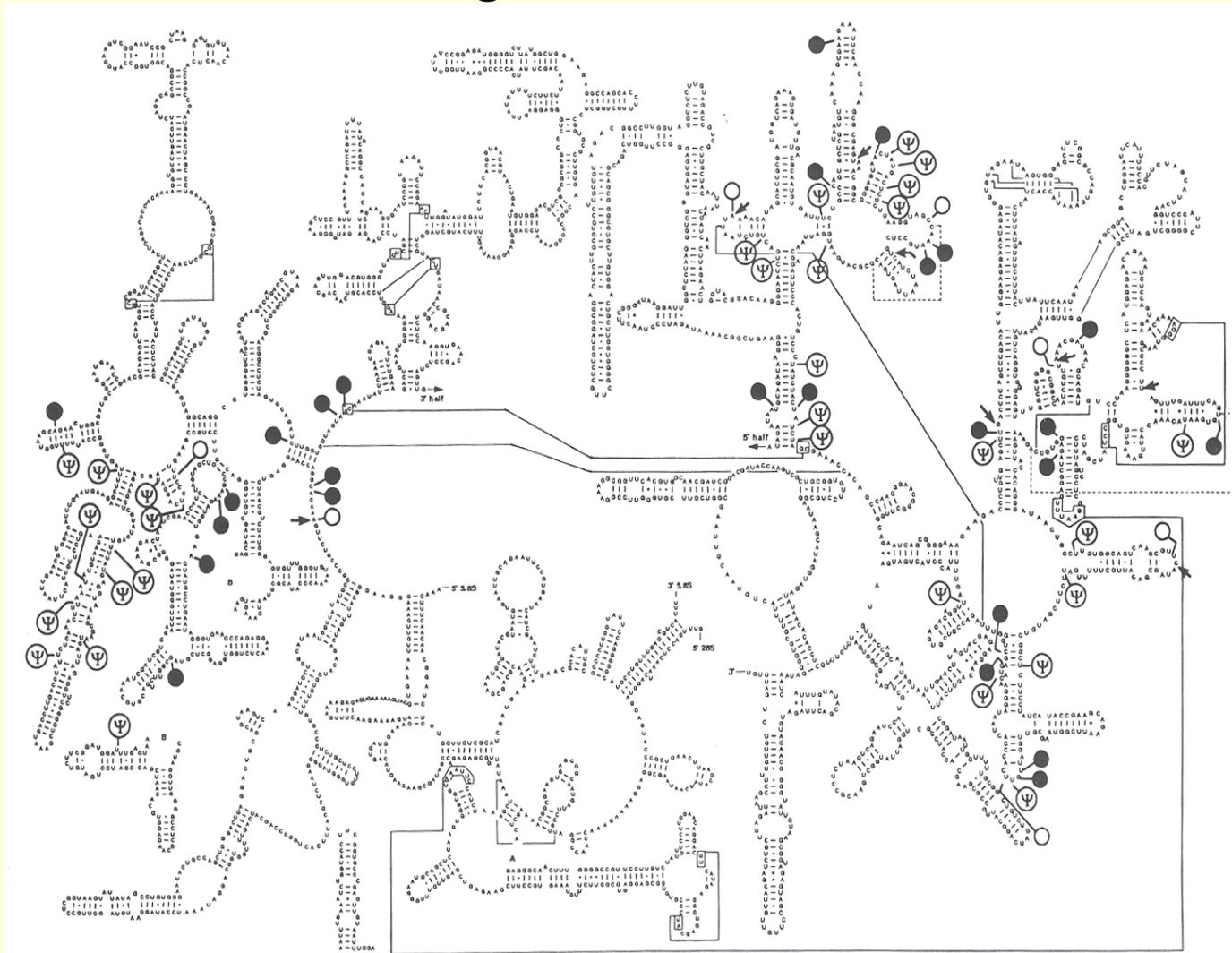
Other kinds of ncRNAs

- Small nuclear RNAs (snRNAs)

Examples:

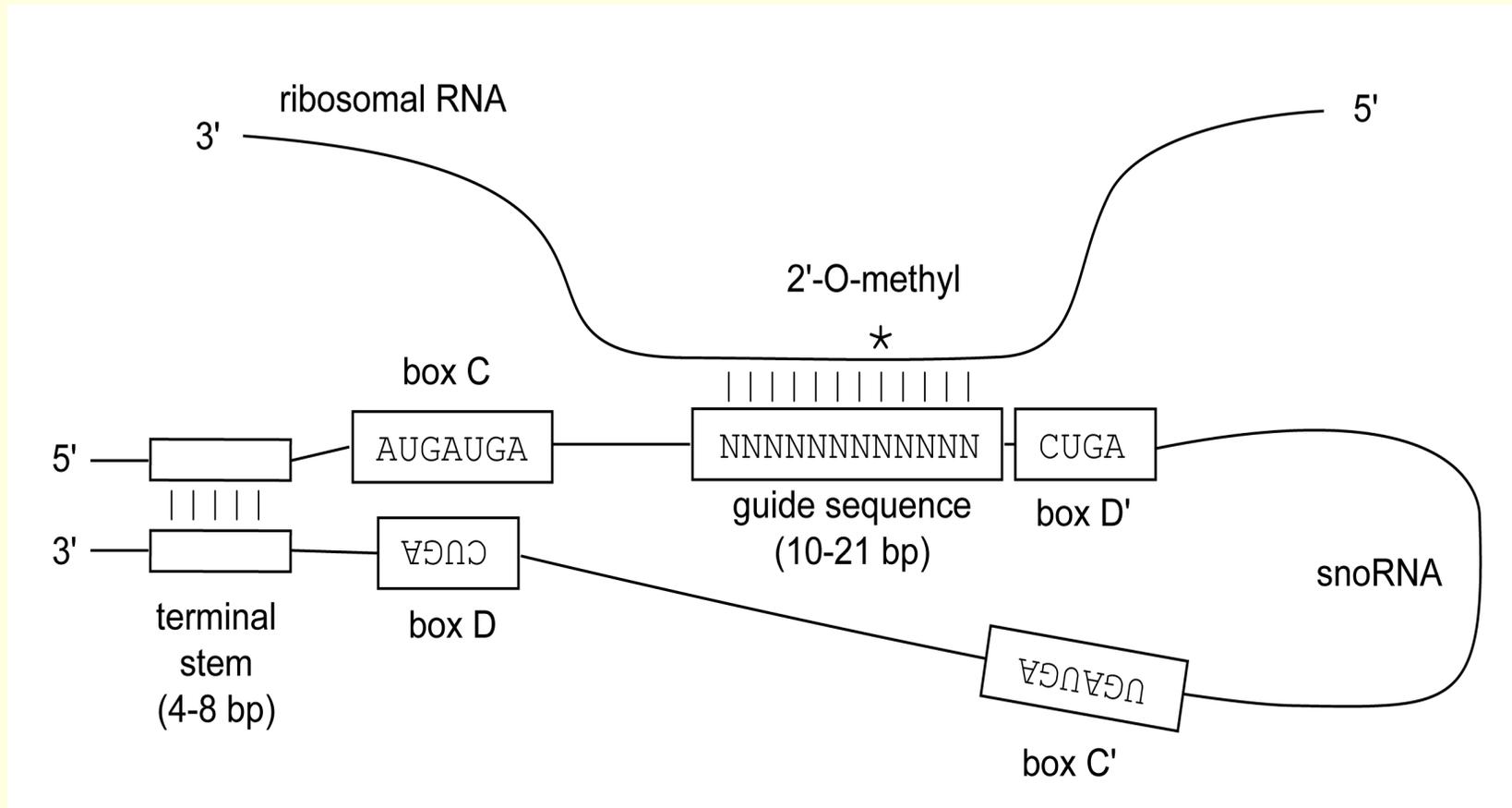
- Spliceosomal RNAs: U1, U2, U4, U5, U6, & others found in spliceosome, involved in removing introns from mRNAs in eukaryotes
- Can usually detect by BLASTN since highly conserved, like ribosomal RNA
- Small nucleolar RNAs (snoRNAs): located in the nucleolus, these process ribosomal RNA (chopping up & adding nucleotide modifications)
- BLASTN only works for some family members from very closely related species

S. cerevisiae Large Subunit Ribosomal RNA



- Highly conserved primary sequence and secondary structure
- Many interacting molecules constrain changes

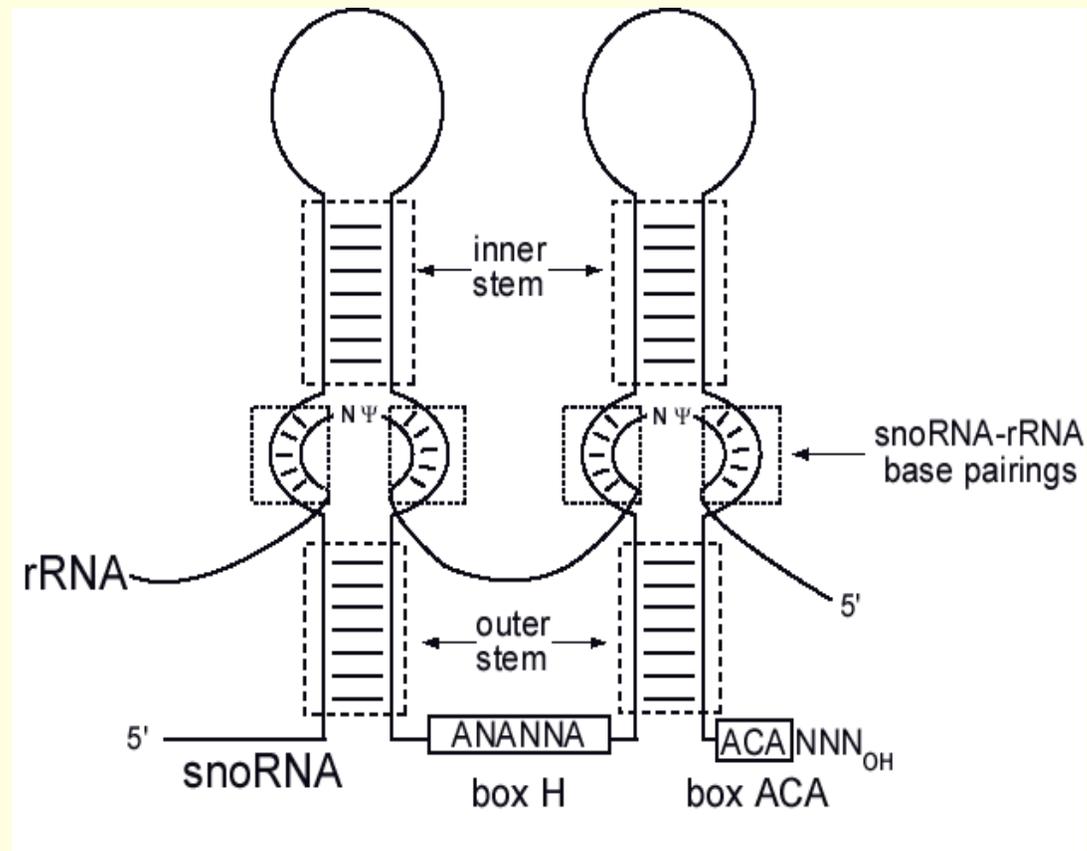
Methylation guide C/D box snoRNA



Very little secondary structure, but short “box” motifs and antisense “guide sequence” conserved

Pseudouridine guide

H/ACA snoRNA



Weak primary sequence motifs, but conserved secondary structure

snoRNA Web Searches

- C/D Box methylation guide snoRNAs
<http://lowelab.ucsc.edu/snoscan/>
- H/ACA pseudouridine guide snoRNAs
<http://lowelab.ucsc.edu/snoGPS/>

More NcRNAs...

microRNAs / siRNAs - short ~22nt RNAs that pair with mRNA to regulate expression

MiRscan: <http://genes.mit.edu/mirscan/>

Bacterial small RNAs: OxyS RNA (antisense temp. sensor), tmRNA in bacteria (incomplete protein translation termination), CsrB (carbon storage regulation)

guide RNAs in trypanosomes – post-transcriptional addition/deletion of nucleotides from mRNAs

RNA Tracks in Genome Browsers

- In UCSC genome browsers, look for:

“RNA Genes”

“sno/miRNA genes”

“transfer RNAs (or tRNAs)”

“Genbank RNAs”

“RFAM RNAs”

Or anything listed at “ncRNA”...

Some On-line Databases

Ribosomal RNA databases:

Ribosomal Database Project: <http://rdp.cme.msu.edu/html/>

rRNA WWW Server: <http://rrna.uia.ac.be/>

tRNA Databases

Genomic tRNA Database <http://lowelab.ucsc.edu/GtRNADB/>

Sprinzi tRNA Database <http://trnadb.bioinf.uni-leipzig.de/>

snoRNA Databases

Yeast snoRNA database

<http://people.biochem.umass.edu/fournierlab/snornadb/>

More Databases

- **SRP RNA Database:**
 - <http://bio.lundberg.gu.se/dbs/SRPDB/SRPDB.html>
- **RNase P Database:**
<http://jwbrown.mbio.ncsu.edu/RNaseP/>
- **tmRNA Database:**
 - <http://psyche.uthct.edu/dbs/tmRDB/tmRDB.html>
- Soon: RNA Central – pooling access to all RNA databases

Other RNA Lists of Links

- RNA World @ IMB Jena (software & databases)
<http://www.imb-jena.de/RNA.html>
- NAR Databases Index (annual update)
<http://www.oxfordjournals.org/nar/database/cat/2>
- NAR Web Server List
http://nar.oxfordjournals.org/content/39/suppl_2