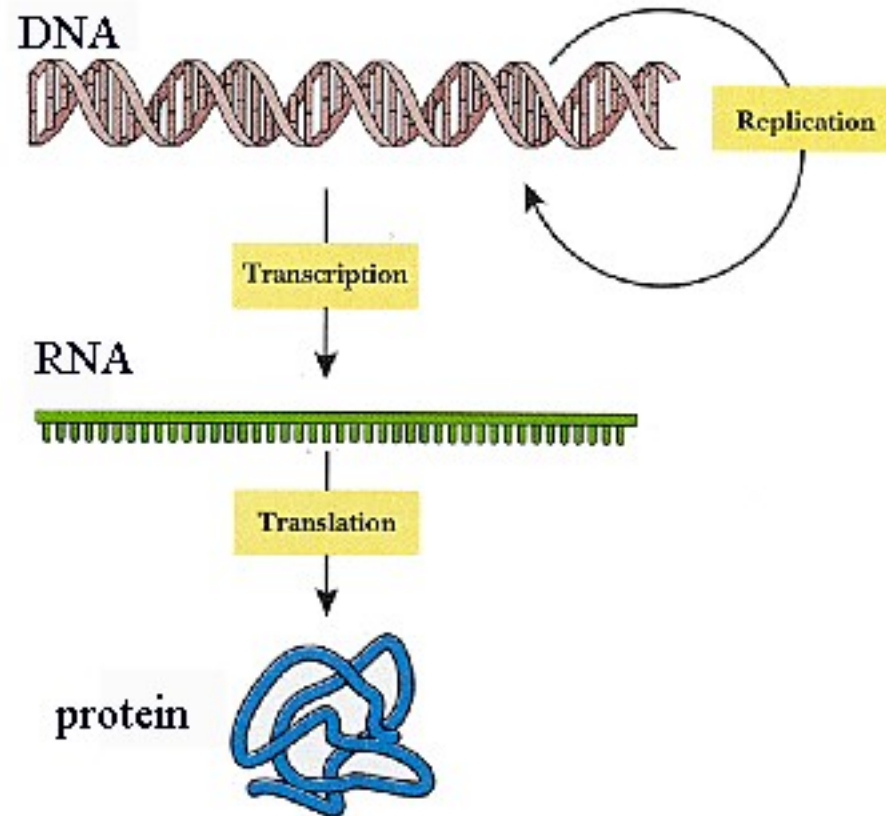


# A glimpse of Non Coding RNA



# Introduction

- Central dogma :



But which came first ??

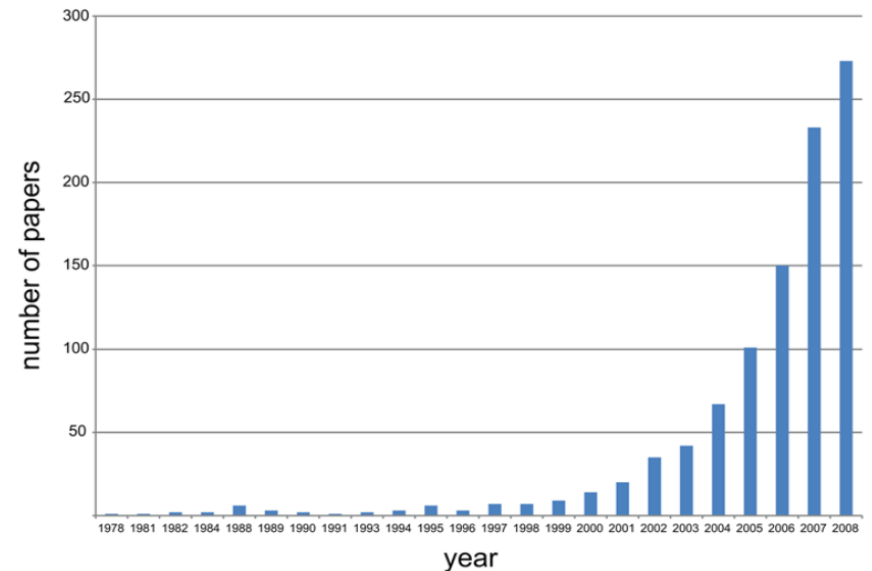
# The RNA World

Sidney Altman and Thomas Cech (1980's) : Ribozymes (RNA able to cleave itself)

RNA role :

- Information, genetic material (DNA role),
- Function, enzymes (Protein role).

Walter Gibson : RNA World Hypothesis (1986)



# Types and Functions of RNAs

- CODING = messenger RNA (mRNA)
  - translated into a protein
- NON-CODING (ncRNA) = not translated into a protein
  - lot of different functions : splicing, translation, gene regulation ...
    - rRNA Ribosomal RNA
    - tRNA Transfer RNA
    - snoRNA nucleosome, rRNA-editing
    - miRNA Control of gene expression on the translation level
    - snRNA nucleus, spliceosomal RNA
    - ... and more (Ribozymes, transcription regulation)

# Different model of evolution compared to coding RNA

- coding RNA : function linked to the sequence of Amino acids
- non coding RNA : function linked to the sequence and the structure of the molecule
  - > no selection on Amino acids sequence

Can not assume independent evolution of sites (paired regions)

# Differences of rates in RNA evolution

High heterogeneity of rates of evolution :

- Among different types ncRNAs
- Same ncRNA among lineages
- Among structure categories within ncRNA
- Within structure categories

# Different rates of evolution among RNAs

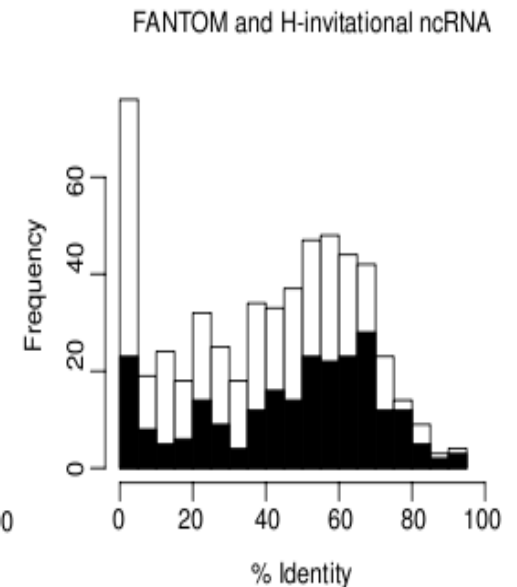
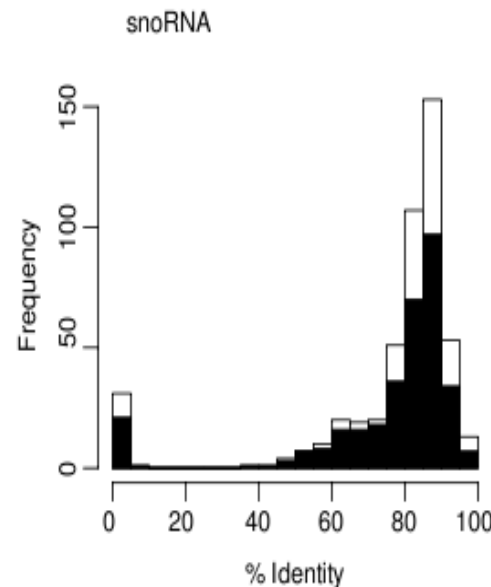
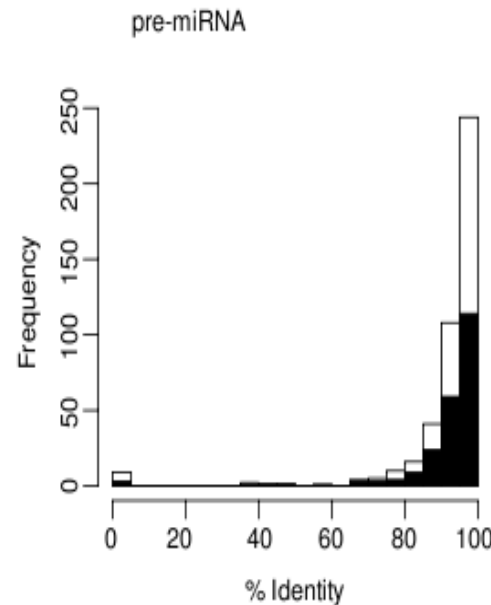
Example : Mammalian

- miRNA and snoRNA slow rate of sequence evolution

-> function linked

to hybridisation with other nucleic acids

- Other types of ncRNA (longer) evolve faster

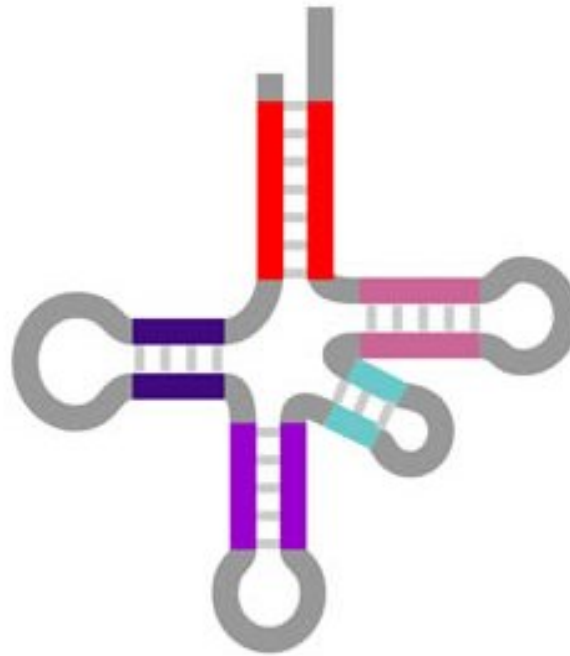


# Non coding RNA structure

Primary

... AAUGCCG ...

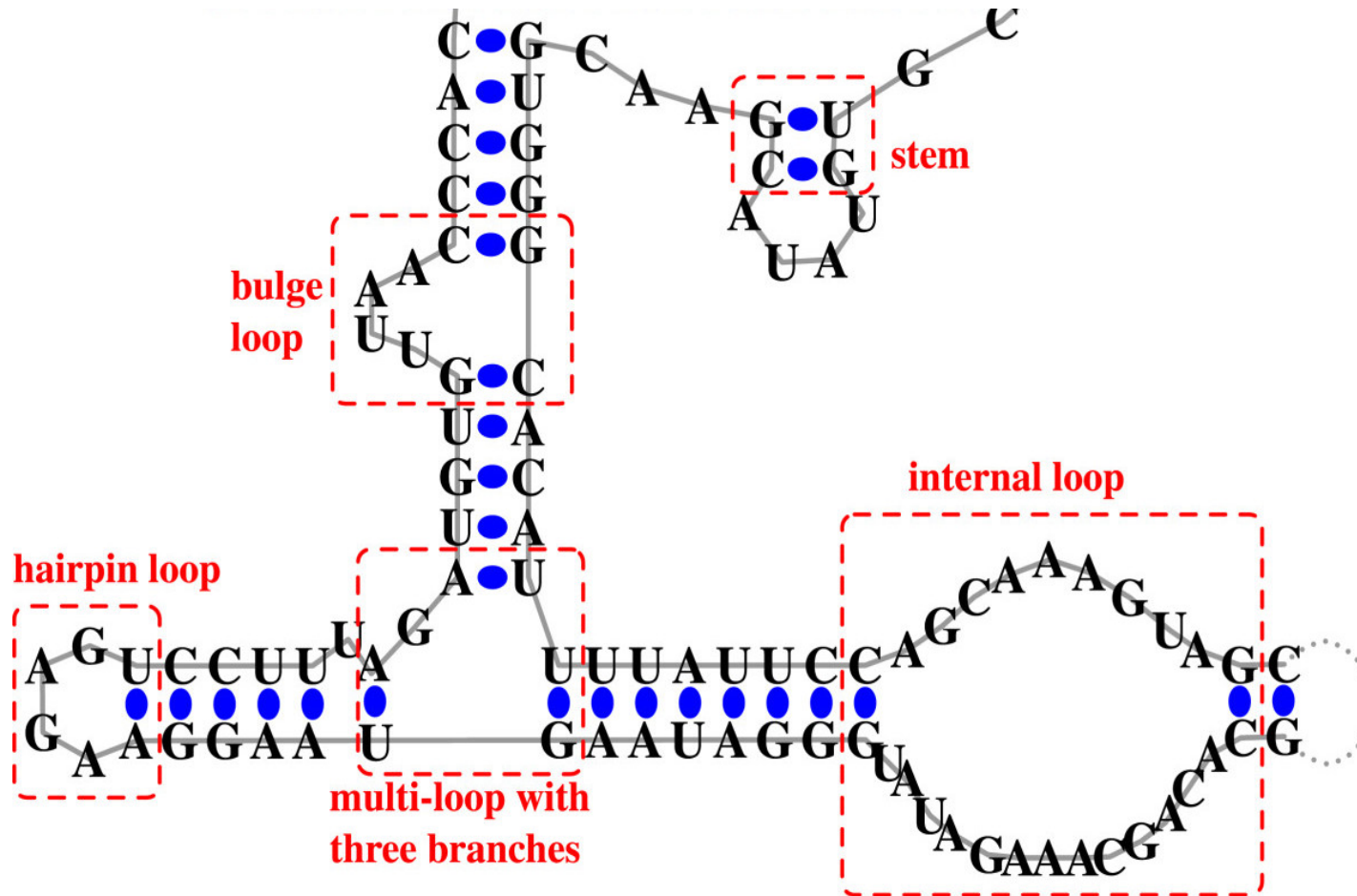
Secondary



Tertiary

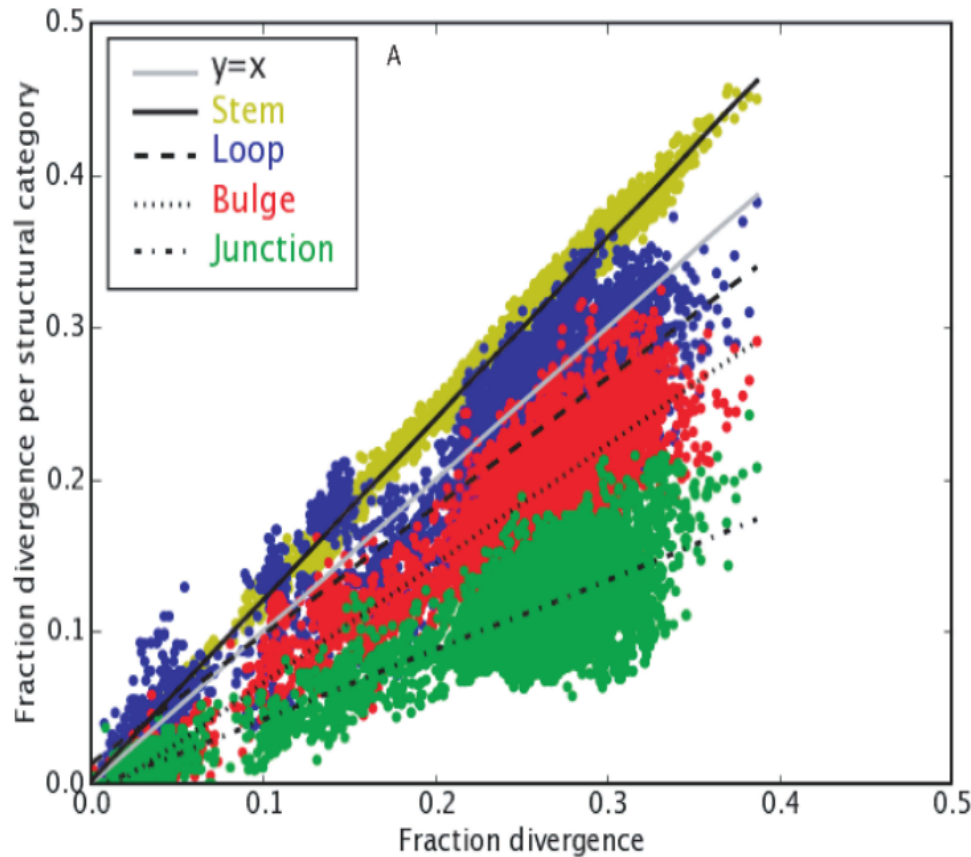


# Secondary structure

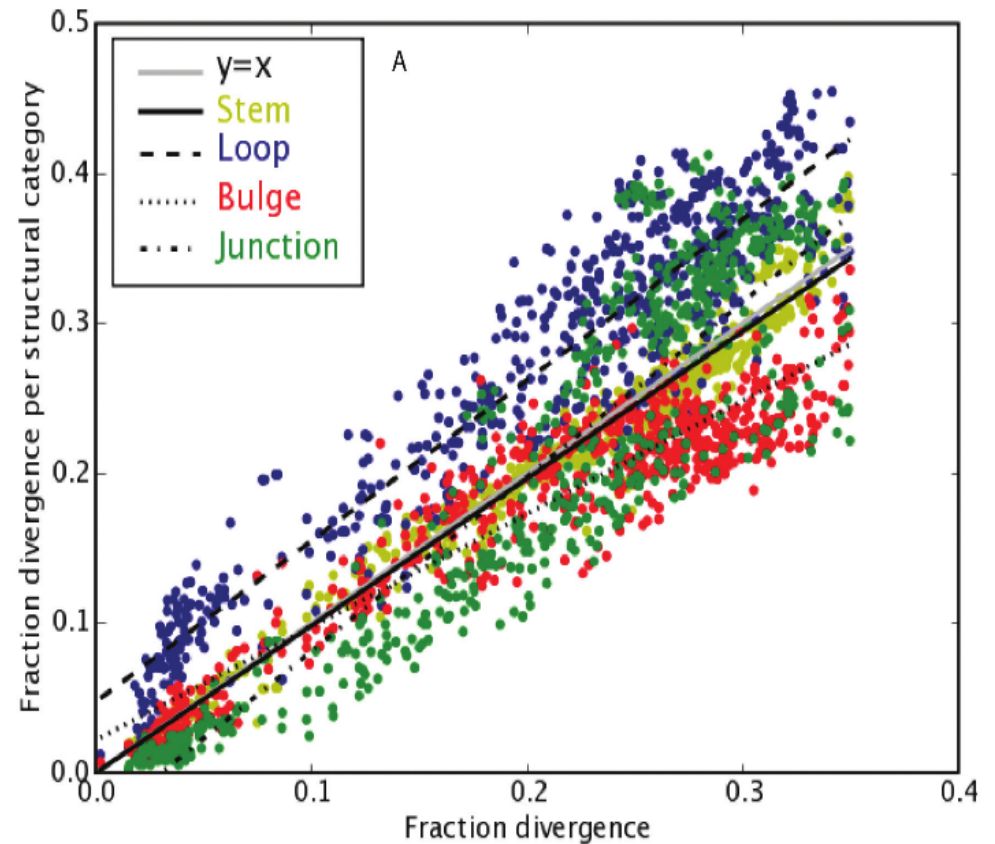


# Different rates among structure categories

## Small Sub Unit RNA



Bacteria

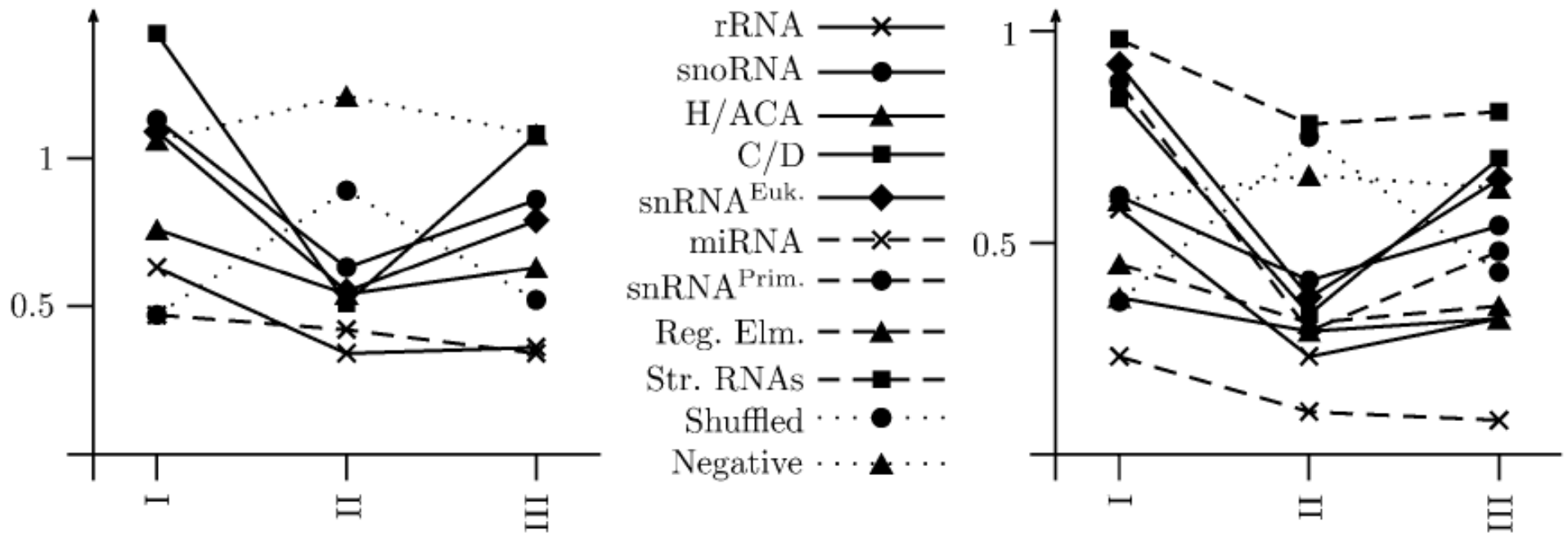


Eukaryotes

Smit et al. 2007

# Different rates within structure categories

Within stem



Valid for highly structured RNAs

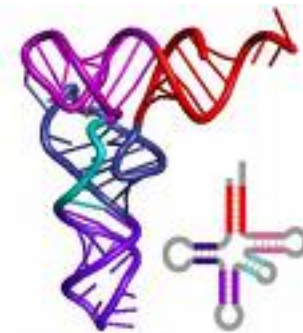
# Why understanding pattern of evolution ?

- Structure prediction
- Accuracy of phylogenetic inferences
- ncRNA gene screening

Creation of databases on RNA structures (RNA STRAND)

# Where are you, ncRNA genes?

- A major goal of bioinformatics
- De novo prediction still a largely unsolved problem
- Evolutionary conservation of secondary structures
- Comparative studies!



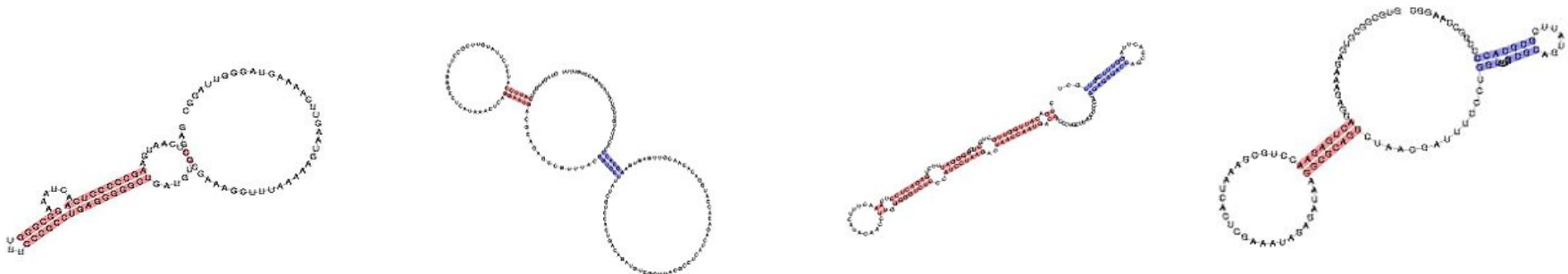
# Approaches to finding ncRNA genes

- EvoFold
- RNAz
- QRNA
- DDBRNA
- MSARI

# EvoFold

- Comparative method for identifying fRNA structures in multiple-sequence alignments
- Phylo-SCFG: Phylogenetic Stochastic Context-Free Grammar
- Input: a multiple alignment and a phylogenetic tree
- Output: a specific RNA secondary-structure prediction and an folding potential score (fps)

Pedersen, et al. 2006



# EvoFold

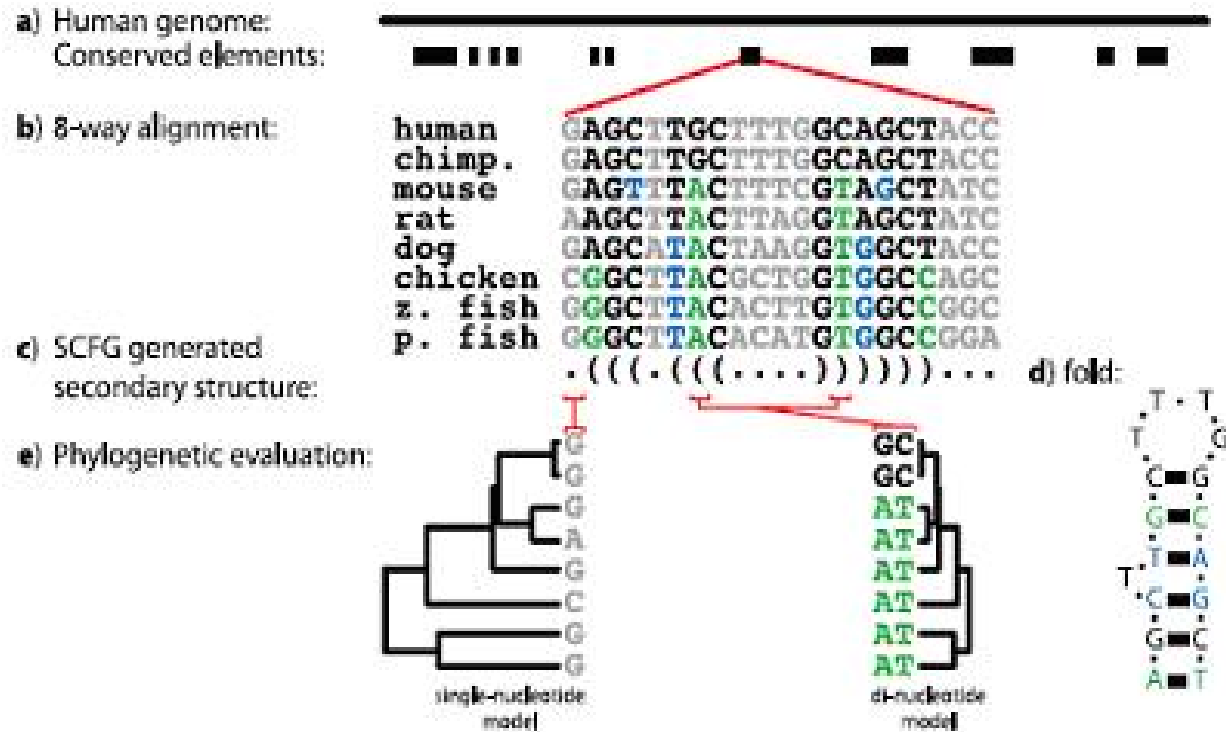


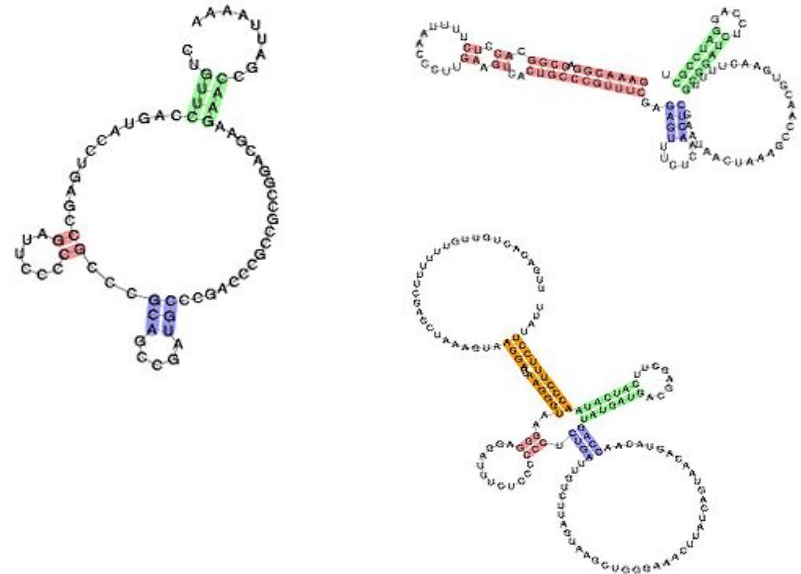
Figure 1. Outline of EvoFold Prediction Method

Pedersen, et al. 2006

- Result: 48.479 candidate RNA structures
- Estimate: candidate set contains ~18.500 substructures of ~10.000 RNA transcripts

# RNAz

- Method for predicting structurally conserved and thermodynamically stable RNA secondary structures in multiple sequence alignments
- Two basic components
  - a measure for RNA secondary structure conservation
  - a measure for thermodynamic stability
- Time complexity:  $O(N \times n^3)$ 
  - $N$ : number of sequences
  - $n$ : length of the alignment



# RNAz

- Comparative Regulatory Genomics database
  - screened all alignments of length  $> 50$



Washietl, Hofacker and Stadler. 2005

# RNAz

- Comparative screen of vertebrate genomes
- Predicted >30.000 structured RNA elements in the human genome
  - ~1.000 conserved across all vertebrates
  - ~1/3 found in introns of known genes
  - ~1/6 potential regulatory elements in UTR of protein-coding mRNAs
  - ~1/2 far away from any known gene
- Only a small fraction previously described

# MicroRNA

- Pretty new discovery, still kind of hot
- Gene regulation on the translation level
- Short sequences (~22bp)
- Hairpin structure (processed)
- [How is miRNA processed](#) (Ghosh et al. 2007)
- Binding to mRNA
  - Binding sites on 5' and 3' UTRs
  - hindering translation (mRNA ring formation)
  - cleavage of mRNA

# More of microRNA

- Finding miRNA genes and mRNA targets in the genome
  - Ghosh et al. (2007) give a list of available programs in mini review.
  - Kuhn et al. (2008) recommend [miRanda](#), [TargetScan](#) and [PicTar](#) for identifying mRNA targets
  - Ragan et al (2009) have introduced a new program called FASTAH
  - The programs use basepairing, free energy calculations and conservation.

# Reference

Zhumur Ghosh, Jayprokas Chakrabarti and Bibekanand Mallick. miRNomics—The bioinformatics of microRNA genes. *Biochemical and Biophysical Research Communications*, 363 (1): 6-11, 2007

Donald E. Kuhn, Mickey M. Martin, David S. Feldman, Alvin V. Terry Jr., Gerard J. Nuovo and Terry S. Elton. Experimental validation of miRNA targets. *Methods*, 44(1):47-54, 2008

Ragan C, Cloonan N, Grimmond SM, Zuker M, Ragan MA. Transcriptome-Wide Prediction of miRNA Targets in Human and Mouse Using FASTH. *PLoS ONE* 4(5): e5745, 2009 doi: 10.1371/journal.pone.0005745

# References

Jakob Skou Pedersen, Gill Bejerano, Adam Siepel, Kate Rosenbloom, Kerstin Lindblad-Toh, Eric S. Lander, Jim Kent, Webb Miller, and David Haussler. Identification and classification of conserved RNA secondary structures in the human genome. *PLoS Computational Biology* , 2(4):e33, 2006.

Stefan Washietl, Ivo L. Hofacker, and Peter F. Stadler. Fast and reliable prediction of noncoding RNA. *Proceedings of the National Academy of Sciences of the United States of America* , 102(7):2454–59, 2005.

Stefan Washietl, Ivo L. Hofacker, Melanie Lukasser, Alexander Huttenhofer and Peter F. Stadler. Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nature Biotechnology* , 23(11):1383–90, 2005.