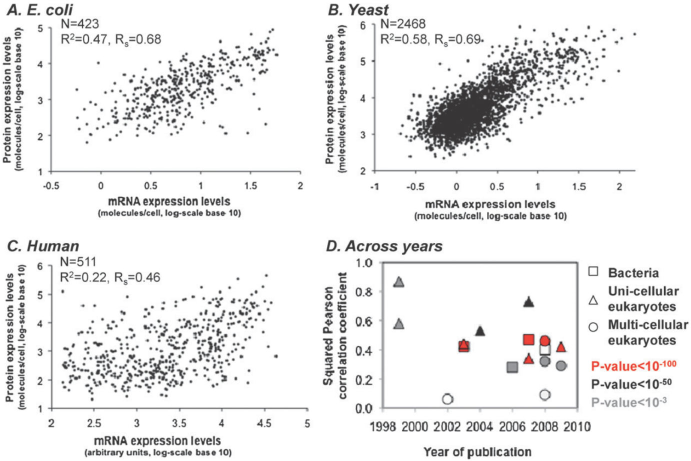
Stochastic interactions between ncRNAs and mRNAs inhibit translation and are selected against

Paul Gardner

July 28, 2015

The hard work of Sinan Uğur Umu

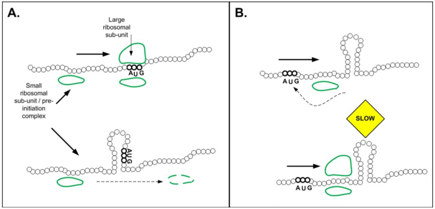
mRNA levels are poorly correlated with protein levels across domains of life



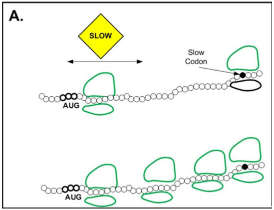
de Sousa Abreu, Penalva, Marcotte & Vogel (2009) Global signatures of protein and mRNA expression levels. *Molecular BioSystems.*

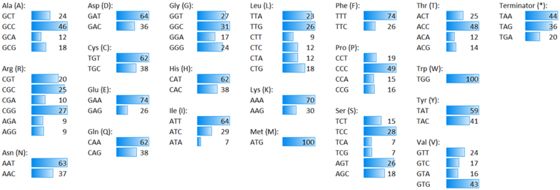
Currently only two general models of translation inhibition exist

- RNA structure



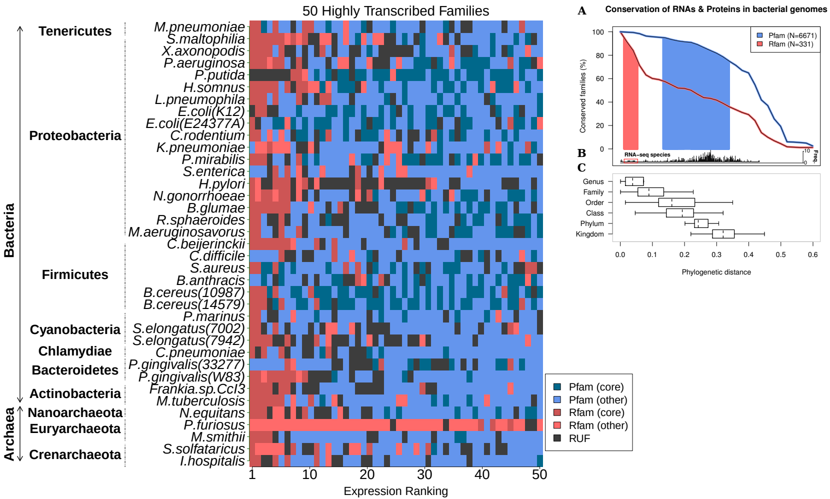
- Codon usage





Tuller & Zur (2015) Multiple roles of the coding sequence 5′ end in gene expression regulation. *Nucl. Acids Res.*

Non-coding RNAs are ubiquitous and abundant



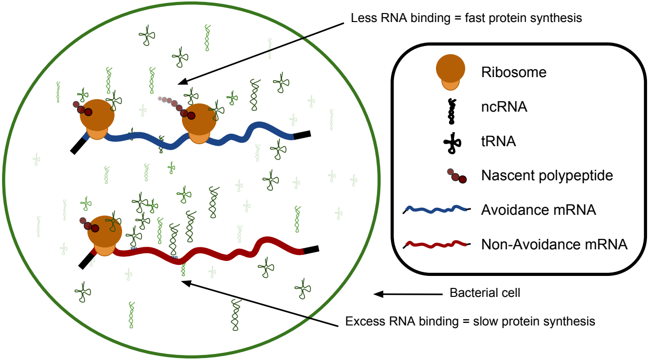
Lindgreen *et al.* (2014) Robust Identification of Noncoding RNA from Transcriptomes Requires Phylogenetically-Informed Sampling. *PLOS Computational Biology.*

Functionally characterising RUFs

- RUFs are some of the most abundant transcripts in the cell
- General approaches for functional characterisation are hard and has been largely unsuccessful, trying:
  - Comparative RNA-seq
  - Motif annotation with RMfam
  - Synteny information
  - Checking potential RNA:RNA interactions

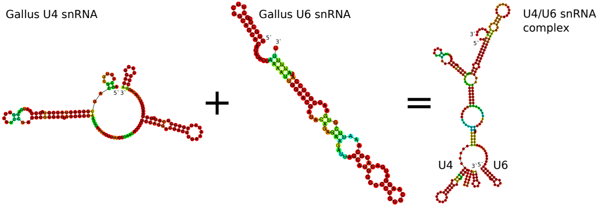
Do ubiquitous and abundant RNAs influence translation?

- Given that ncRNAs like rRNA & tRNA, RNase P RNA, SRP RNA, tmRNA and 6S RNA are among the most abundant RNAs in the cell ($[ncRNA] >> [mRNA]$)
- AND that RNAs are sticky (like to hybridise)
- THEN maybe stochastic interactions with mRNAs inhibit translation



How can this hypothesis be tested?

- An accurate hybridisation prediction method:
    - RNAup: nearest neighbour model, computes unfolding as well as hybridisation energy
- Focus on 6 RNA families & 114 proteins that are highly conserved. Test the 1st 21 nucleotides of CDS (critical region & saves LOTS of compute)
- We predict that: (**1**) there will is selection against ncRNA:mRNA interactions
- and (**2**) that ncRNA:mRNA interactions will influence [protein]:[mRNA] ratios

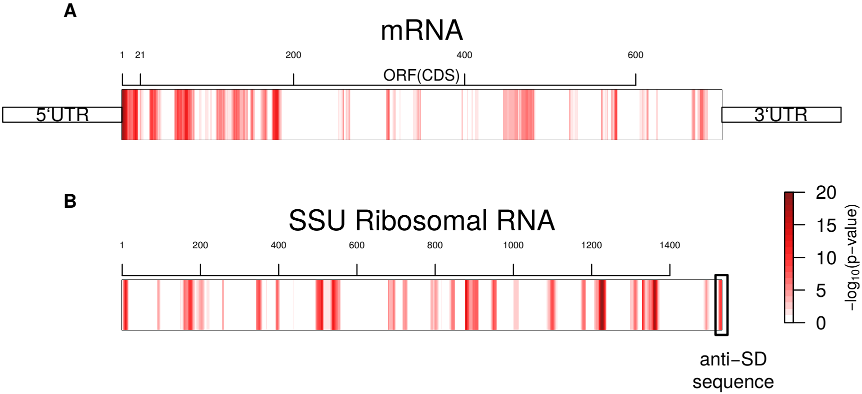


ΔG = -39.70 kcal/mol

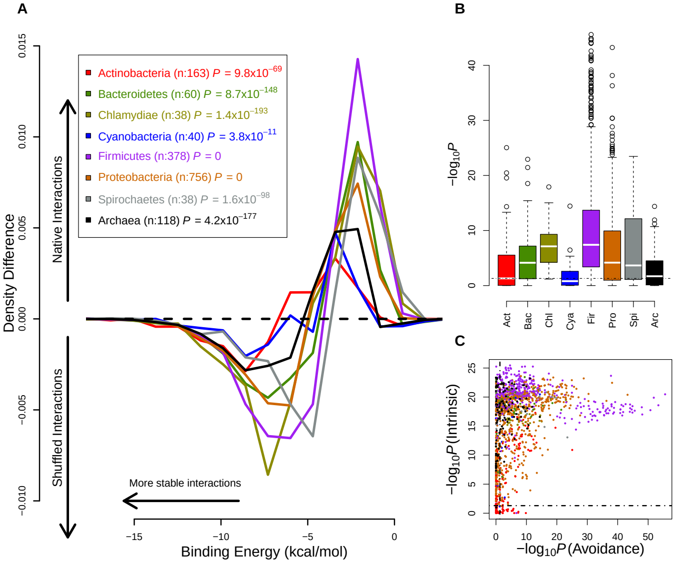
ΔG = -32.60 kcal/mol

ΔG = -73.80 + 13.10 + 19.10
= -41.6 kcal/mol

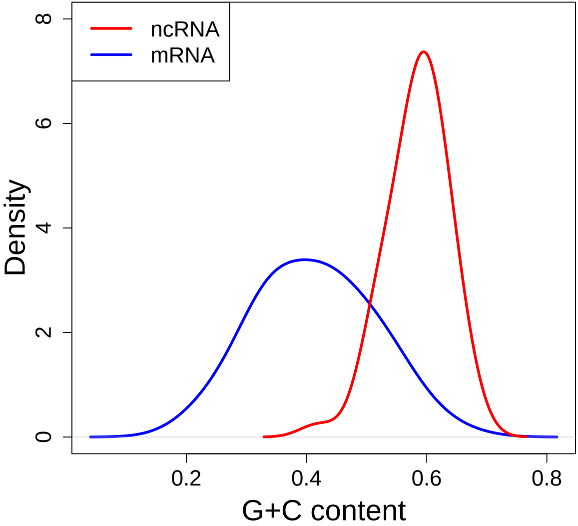
Justifying 1:21



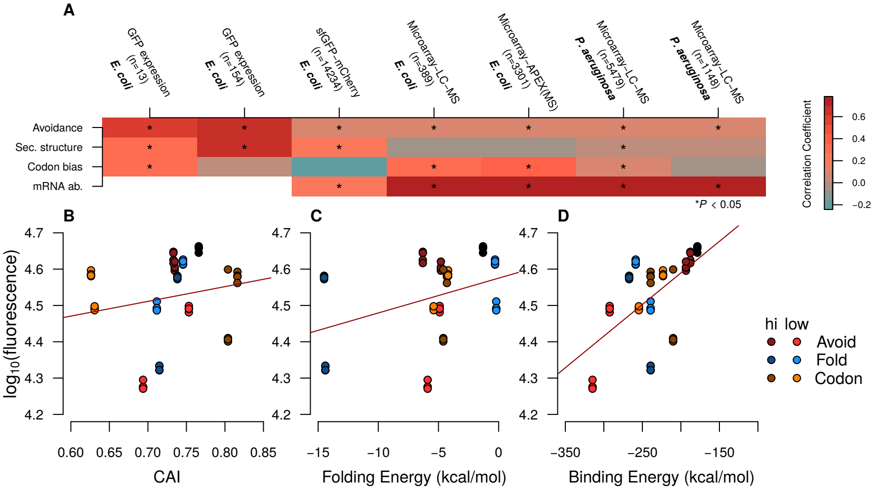
Are ncRNA:mRNA interactions selected against?



"Intrinsic" avoidance: *E. coli*



Do ncRNA:mRNA interactions influence protein expression?



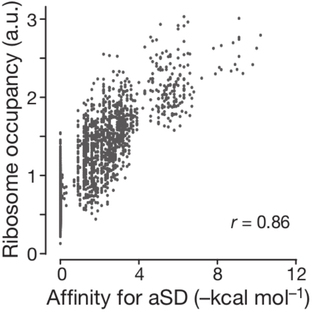
CAI: $R_s = 0.29$, $P = 0.016$; Sec. Struct.: $R_s = 0.34$, $P = 0.006$; Avoidance: $R_s = 0.56$, $P = 6.9x10^{-6}$

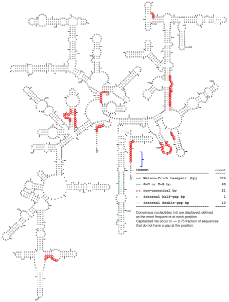
A specific case of ncRNA:mRNA interactions influencing protein expression

doi:10.1038/nature10965

The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria

Gene-Wei Li[1], Eugene Oh[1] & Jonathan S. Weissman[1]





Li, Oh & Weissman (2012) The anti-ShineDalgarno sequence drives translational pausing and codon choice in bacteria. *Nature.*

Conclusion & Further Work

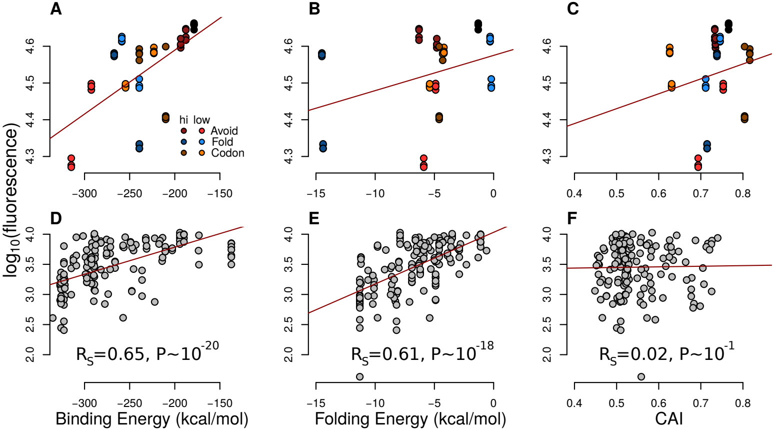
- Can we detect negative selection against ncRNA:mRNA interactions over evolutionary time-scales?
    - Yes!
- Can we determine if ncRNA:mRNA interactions influence protein expression?
    - Yes! $\sim$ 42% of variation in GFP expression, 0.8-1.2% in native data
- Further work:
    - Do ncRNA:mRNA interactions influence viral infection, hybridisation, HGT & transformation expts?
    - Do ncRNA:mRNA interactions influence Eukaryotic gene expression?
    - Are ncRNA:mRNA interactions a driver of compartmentalisation?

Thanks & Plugs

- Sinan Uğur Umu
- Anthony Poole & Renwick Dobson

- Rfam/RNAcentral are hiring:
  - Project Leader
  - Software Developer
  - Database Biocurator

Effect sizes & scatter plots



Kudla, Murray, Tollervey, Plotkin (2009) Coding-sequence determinants of gene expression in Escherichia coli. *Science.*