

THE FUNCTIONAL REPERTOIRE OF RNA AND ITS RELATIONSHIP TO THE EARLY PROTEOME.

A.D. Goldman^{1,*}, Y.A. Takagi¹, T.M. Bernhard², T.W. Tullius¹, L.F. Landweber². ¹Oberlin College, Oberlin, OH, ¹Princeton University, Princeton, NJ, *agoldman@oberlin.edu

A minimal expression of the RNA world hypothesis is that the current genetic system in which informational nucleic acids encode functional proteins was preceded by one based on RNAs or RNA-like polymers that performed both the informational and functional roles. A number of corollary hypotheses follow from this model. Some nucleotide-derived cofactors used in modern enzymes are proposed to be vestiges of the active sites of ribozyme involved in an RNA-based metabolism¹. Some amino acid-derived cofactors are thought to have been catalysts in an RNA world metabolism and to have provided the initial impetus for the evolution of a genetic code prior to the establishment of protein synthesis by translation². The early evolution of protein-mediated metabolism may have involved pathways partially composed of ribozymes with functions that were ultimately displaced by protein enzymes³. Here we use a comprehensive survey of RNA enzymatic functions to test these hypotheses about the early evolution of an RNA world. We also explore the known regulatory functions of RNAs in extant life to demonstrate the potential for trans-regulation of RNA world ribozymes.

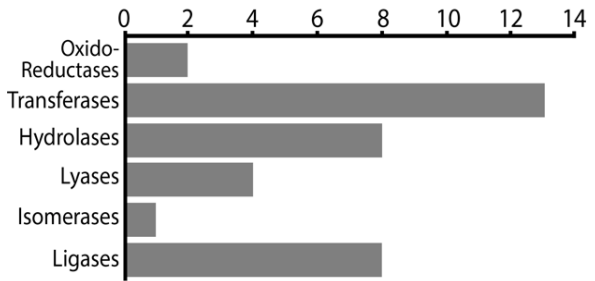


Figure 1: A histogram of known RNA functions found in extant cells or designed in laboratories⁴. There are 33 functions in total that span all major functional categories defined by the Enzyme Commission.

Dataset Type	Dataset	Ribozyme Overlap
Functions of proteins from LUCA gene family predictions	Delaye ⁵	p < 0.02 [§]
	Harris ⁶	p < 0.04 [§]
	Mirkin ⁷	p < 0.03 [§]
	Wang ⁸	p < 0.00005*
	Yang ⁹	p < 0.0007*
	Srinivasan ¹⁰	p < 0.09
Functions of proteins using proposed ancient enzyme cofactors	Amino acid ¹	p < 0.017 [§]
	Nucleotide ²	p < 0.0042*
	Iron-Sulfur ¹¹	p = 0.72
	Zinc ¹²	p = 0.71

Table 1 (bottom left): Comparisons between the ribozyme functional set to datasets from the LUCAPedia database⁴ representing protein functions in predicted Last Universal Common Ancestor (LUCA) gene families and functions of enzymes using proposed ancient cofactors. Significance is determined by comparison of RNA functions to simulated random datasets. The ribozyme repertoire shares a statistically significant overlap with LUCA models as well as with those cofactors predicted to have played a role in the transition from RNA-mediated to protein-mediated metabolism.
* below Bonferroni-corrected threshold of p < 0.005.
§ below uncorrected threshold of p < 0.05.

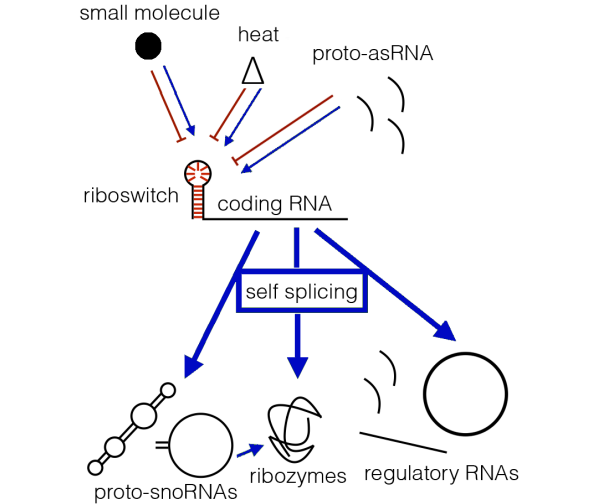


Figure 2: An example of the regulatory potential of RNA. RNA may be converted to its functional form by self-splicing. This conversion can be regulated in *cis* by riboswitches than can, themselves, be activated or stabilized in *trans* by antisense RNAs, small molecules, and environmental conditions like heat.

References: [1] Kyrpides N.C. and Ouzounis C.A. (1995) *J Mol Evol*, 40, 564–569. [2] Szathmáry E. (1999) *Trends Genet* 15, 223-229. [3] Lazcano A. and Miller S.L. (1999) *J Mol Evol*, 49,424-31. [4] Goldman A.D., et al. (2013) *Nucleic Acids Res*, 41, D1079-82. [5] Delaye L., et al. (2005) *Orig Life Evol Biosph*, 35, 537–554. [6] Harris J.K., et al. (2003) *Genome Res*, 13, 407-12. [7] Mirkin B.G., et al. (2003) *BMC Evol Biol*, 3, 2. [8] Wang M., et al. (2007) *Genome Res*, 17, 1572–1585. [9] Yang S., et al. (2005) *Proc Natl Acad Sci USA*, 102, 373–378. [10] Srinivasan V. and Morowitz H.J. (2009) *Biol Bull*, 216, 126–130. [11] Wächtershäuser G. (1990) *Proc Natl Acad Sci USA*, 87, 200–204. [12] Mulkidjanian A.Y. and Galperin M.Y. (2009) *Biol Direct*, 4, 27.