Progressive lengthening of 3’ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development

Journal clubs in bioinformatics

by Tõnu Margus

nature
International weekly journal of science

Journal club

Robert Blelloch[1]

A computational biologist looks at how mRNA length changes during development.

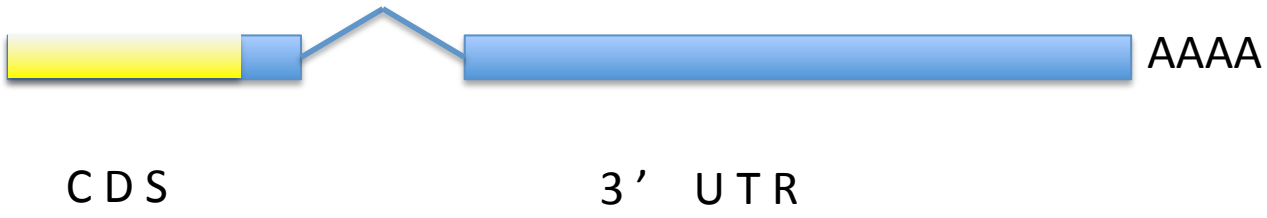
I am always amazed by how we start as a fertilized egg and develop into a complex, multicellular organism. This feat occurs despite the fact that the DNA in every cell — even the most specialized ones — remains, for the most part, unchanged.

PNAS

Progressive lengthening of 3′ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development

Zhe Ji[1], Ju Youn Lee[1], Zhenhua Pan[1], Bingjun Jiang, and Bin Tian[2]

introduction



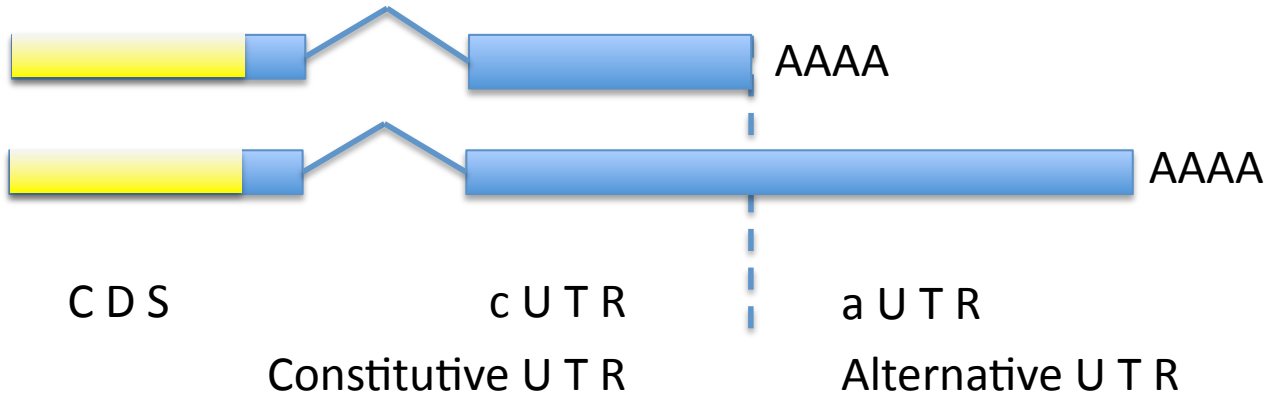
- The 3’ UTRs of mRNA harbor various *cis*-acting elements involved in posttransc-riptional regulation

Regulatory elements in 3'UTR

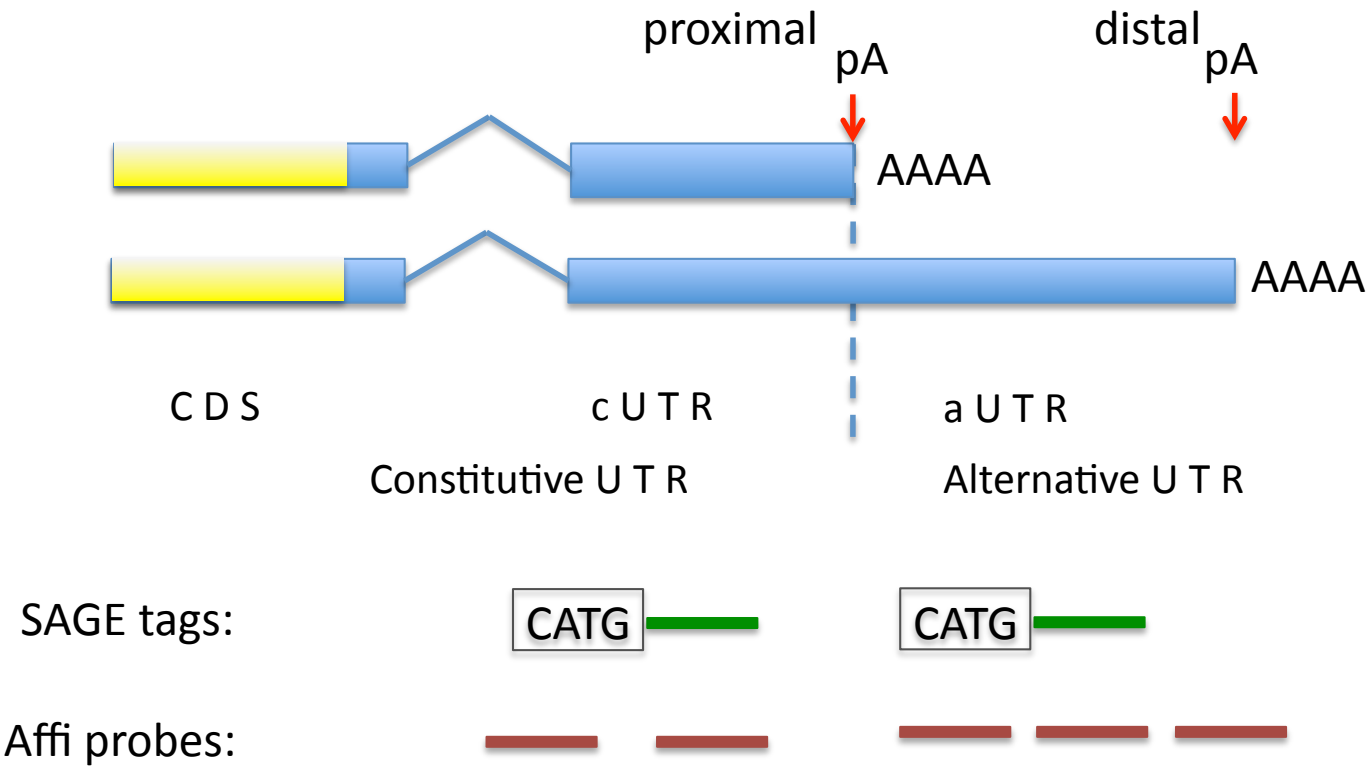
- RNA–protein complexes control mRNA localization, translation and stability
- Some interacts with miRNA that cause inhibition of translation and/or stability
- In humans, the AU-rich elements (AREs) involved in control of mRNA stability are found in 5-8% of all genes, and over 30% of genes are predicted to have miRNA target sites

Alternative poly(A) pattern

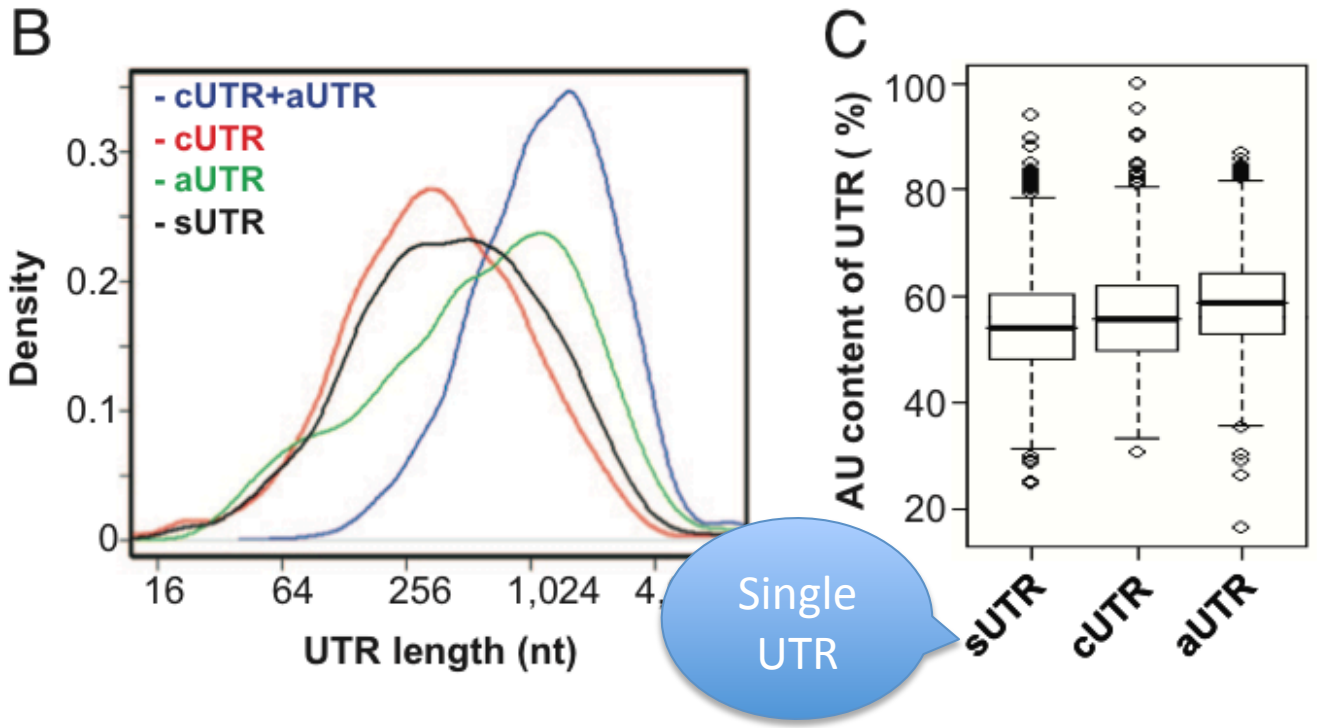
- Over half of mammalian genes contain multiple polyadenylation sites
- Alternative polyadenylation pattern (APA) of genes varies across human and mouse tissues, and can be affected by genomic imprinting
- Regulation of polyadenylation factors and regulatory proteins have been found to modulate APA



Most mouse genes with APA contain 2 poly(A) sites



Distribution of length of different UTR groups



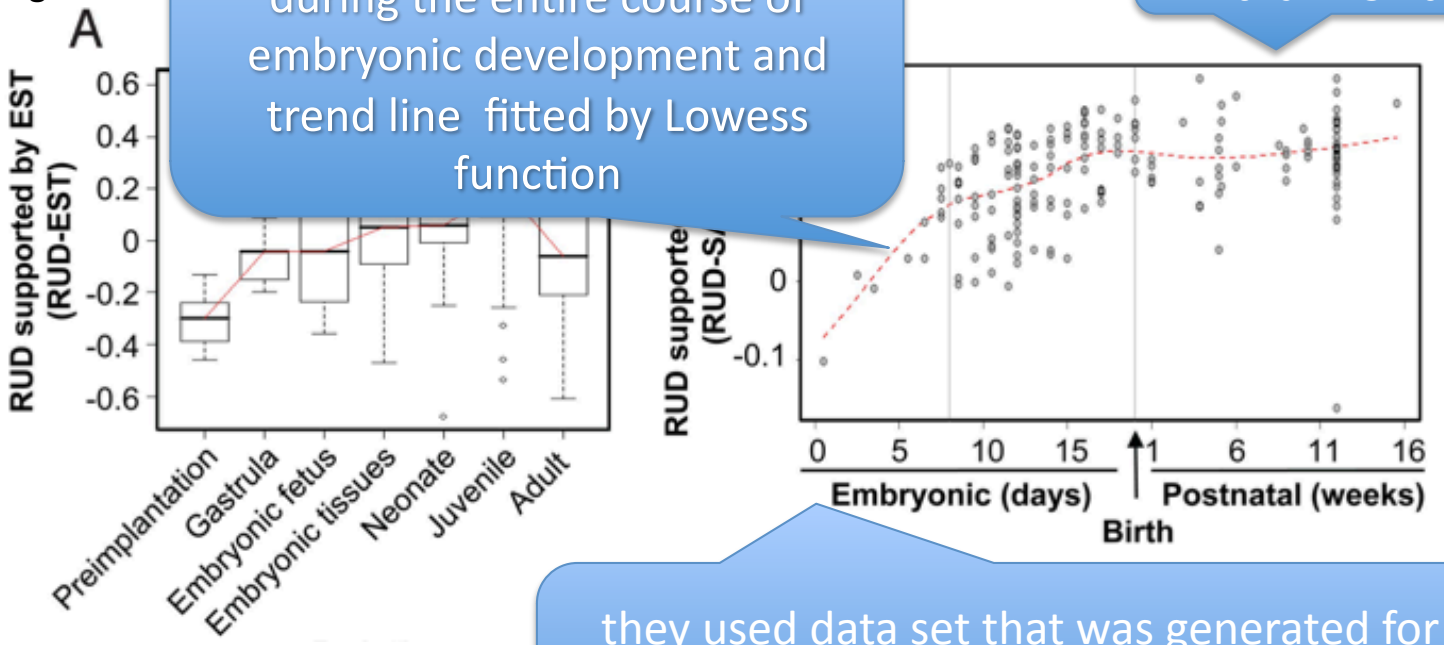
(B) Distribution of length for different UTR groups. The full 3' UTRs for genes with APA are shown by cUTR + aUTR [4,139 genes, median = 1,288 nucleotide (nt)]. Constitutive and alternative regions are represented by cUTR (median = 358 nt) and aUTR (median = 685 nt), respectively. sUTRs are 3' UTRs not affected by APA (7,242 genes, median = 439 nt). (C) AU content for different UTR sequences.

story rolls

Interestingly, we noticed that many mouse EST sequences supporting proximal poly(A) sites came from cDNA libraries that were annotated with early development stages, particularly those before implantation.

RUD – Relative usage of Distal Poly(A)

Zhang et al 20



3'UTR progressively lengthen during the entire course of embryonic development and trend line fitted by Lowess function

More quantitative than EST data

they used data set that was generated for almost all stages of mouse development, from fertilized egg to postnatal development (Table S2)

The **RUD-EST** value for a cDNA library is the averaged usage of distal poly(A) sites minus the averaged usage of proximal sites across all genes.

The **RUD-SAGE** value for a given SAGE library is the averaged percentage of tags mapped to aUTRs minus the averaged percentage of tags mapped to up- stream regions across all genes.

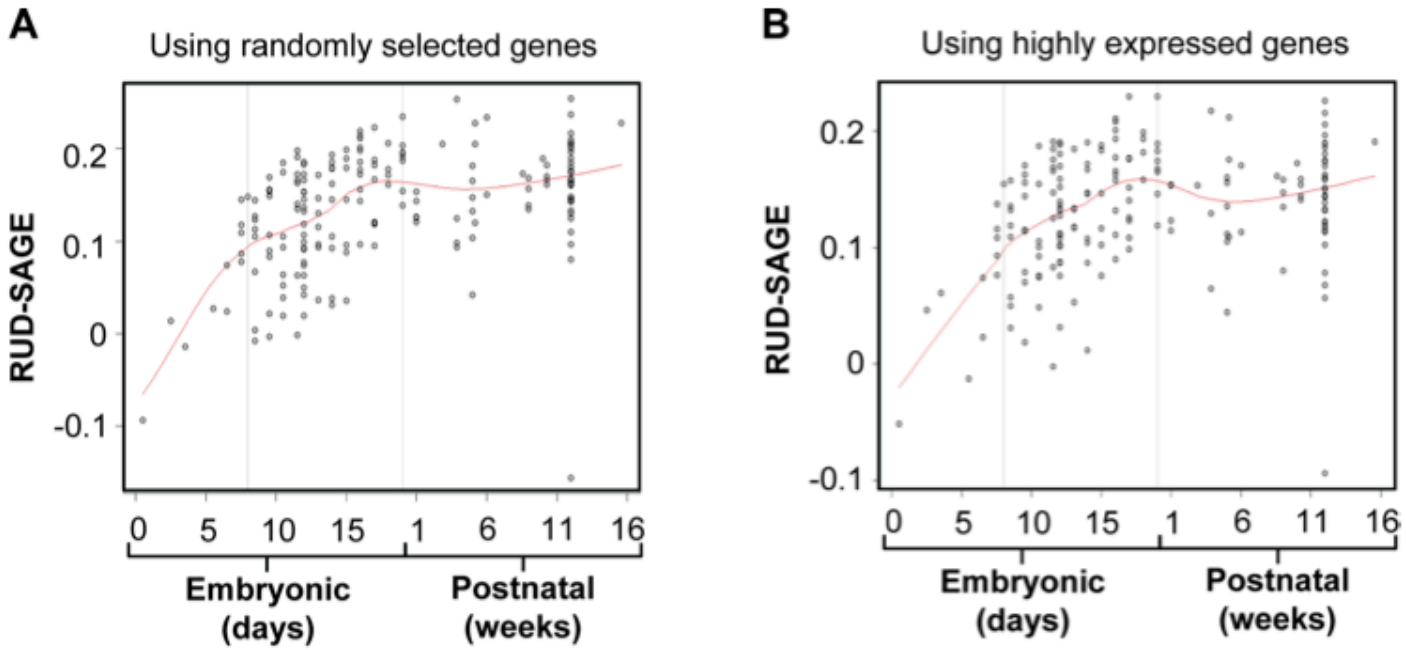
APA variants - stability versus synthesis

Because additional *cis* elements influencing mRNA stability can exist in aUTRs compared with cUTRs, **the difference** in the steady-state level of mRNA **between APA variants** may result from their **difference** in **stability rather** than **synthesis.**

This issue can be addressed for global analysis of trend:

We reasoned that **randomly selected genes** are **not likely** to share the **same regulatory mechanism for mRNA stability** across different conditions, and **highly expressed genes** or genes expressed at their high levels are **not likely** to be under the **condition** under which they are **subject to degradation**.

APA variants - stability versus synthesis



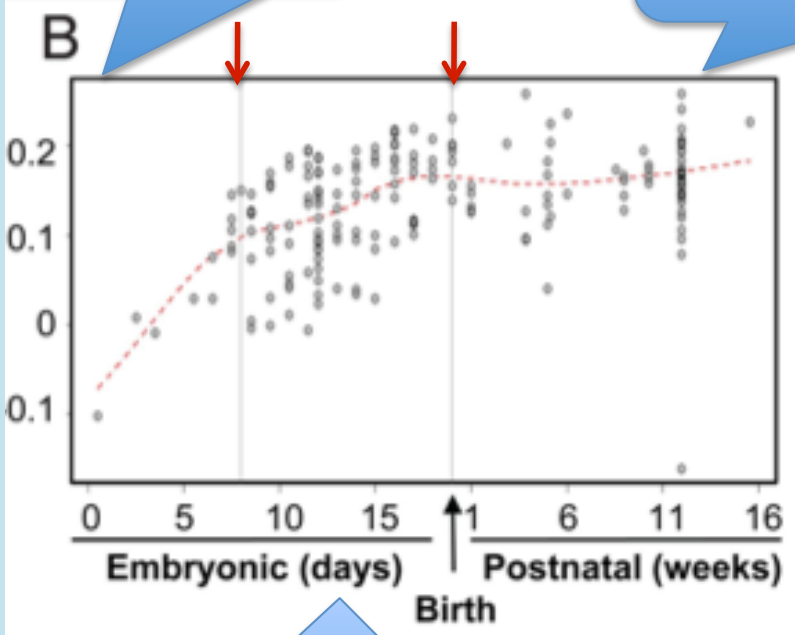
They found that all gene selection methods resulted in **similar trends** (Fig. S1), indicating that the **observed lengthening of 3'** UTR most likely **results from regulation of APA**.

Three phases of 3’ UTR lengthening

First phase:
before the 8th embryonic day E 8.0

Third phase:
Postnatal development

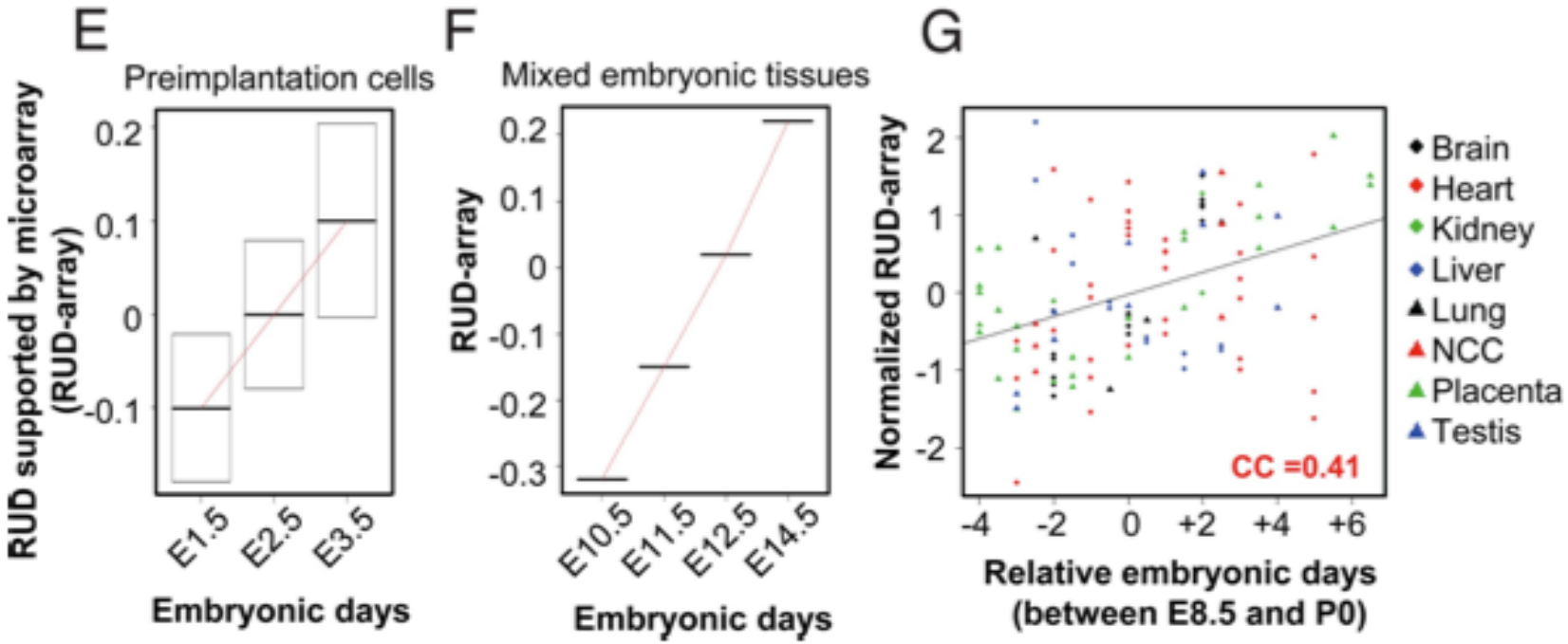
Because during mouse embryonic development, 3 germ layers, i.e., **ectoderm, endoderm**, and **mesoderm**, begin to **develop into organs** around E8.0, our **result suggests** that **3’ UTR lengthening** in the first phase **coordinates with initiation of organogenesis**



Second phase:
From E 8.0 to birth

Microarray data related to mouse embryonic development

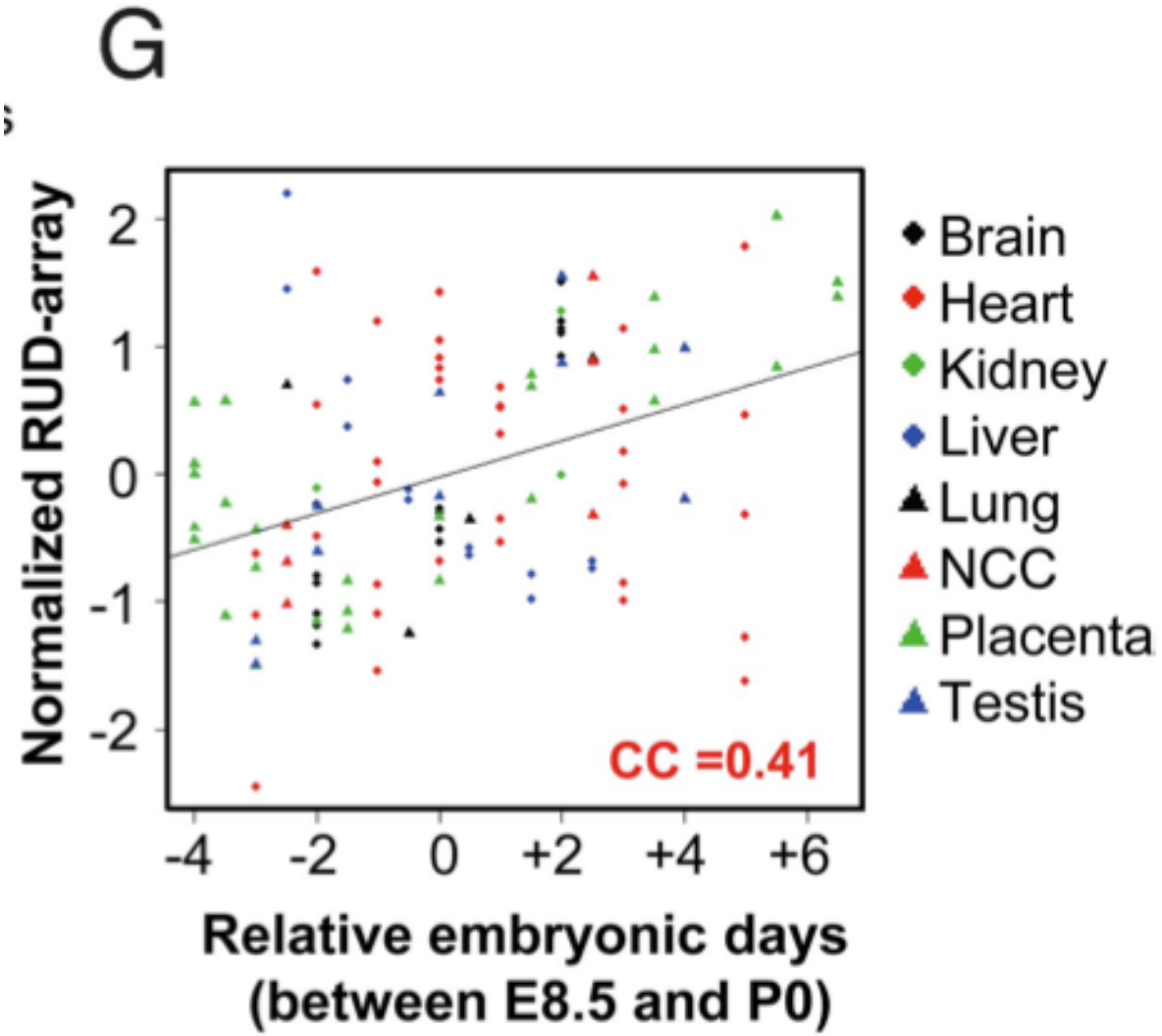
Many types of Affymetrix GeneChip arrays contain probes that target aUTR regions



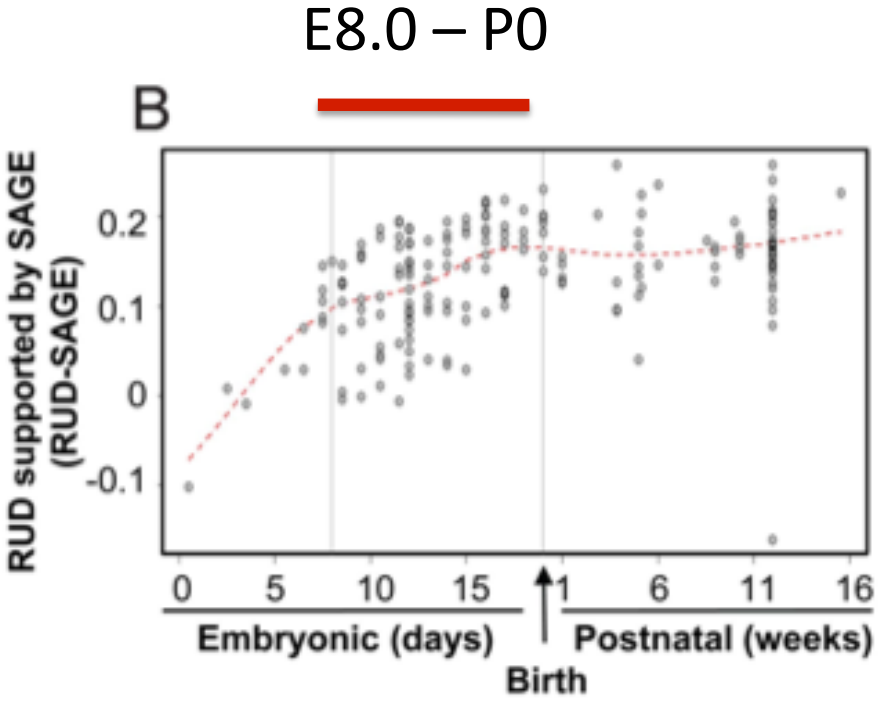
Lengthening of 3' UTR by APA can be detected for preimplantation cells between E1.5 and E3.5, and mixed embryonic tissues between E10.5 and E14.5.

Different tissues between E8.5 – P0

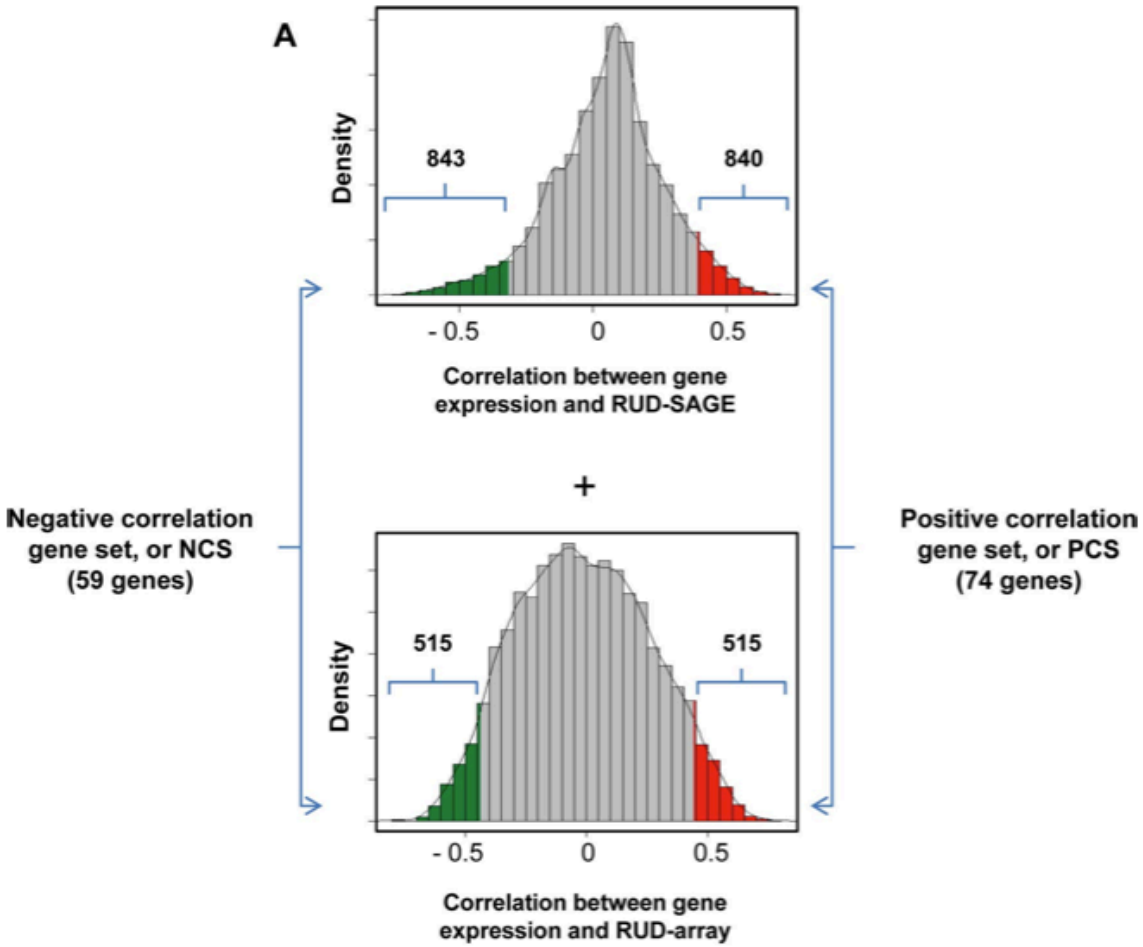
Variations of 3' UTR lengthening can be discerned among different tissues between E8.0 and P0, in both SAGE (Fig. 2B) and microarray (Fig. 2G**) results, suggesting 3' UTR lengthening is regulated differently in different tissue types in this phase of development**



3’ UTR and gene expression correlation



3’ UTR and gene expression correlation



(A) Genes of which expression levels are positively and negatively correlated with RUD-SAGE (Top) and RUD-array (Bottom) values between E8.0 and P0 were selected.

Table 1. GO entries associated with genes that have significant correlation with RUD-SAGE and RUD-array

Type	$-\log_{10}(P)$	GO ID, GO term
Associated with genes that positively correlate with RUD	3.51	GO:0006817, phosphate transport
	2.86	GO:0043062, extracellular structure organization
	2.60	GO:0007167, enzyme linked receptor protein signaling pathway
	2.19	GO:0007259, JAK-STAT cascade
	2.15	GO:0045595, regulation of cell differentiation
	2.02	GO:0030029, actin filament-based process
	1.97	GO:0030324, lung development
	1.79	GO:0040008, regulation of growth
	1.60	GO:0006182, cGMP biosynthetic process
	1.55	GO:0050982, detection of mechanical stimulus
	1.49	GO:0000902, cell morphogenesis
Associated with genes that negatively correlate with RUD	15.28	GO:0006260, DNA replication
	14.15	GO:0022403, cell cycle phase
	13.99	GO:0006259, DNA metabolic process
	12.73	GO:0051301, cell division
	11.95	GO:0006396, RNA processing
	10.72	GO:0006412, translation
	3.62	GO:0051716, cellular response to stimulus
	3.47	GO:0065002, intracellular protein transmembrane transport
	2.90	GO:0022607, cellular component assembly
	2.83	GO:0051726, regulation of cell cycle
	2.82	GO:0009262, deoxyribonucleotide metabolic process

We considered the GO terms that are identified by both SAGE and microarray analyses to be the most significant and reliable.

They found that 3’ UTR lengthening coincides with **upregulation** of genes involved in **morphogenesis and differentiation**, such as
‘‘extracellular structure organiza- tion,’’
‘‘regulation of cell differentiation,’’
‘‘actin filament-based process,’’ and
‘‘cell morphogenesis,’’ and with
downregulation of genes involved **in proliferation**, such as
‘‘DNA replication,’’
‘‘cell cycle phase,’’
‘‘DNA metabolic process,’’ and
‘‘cell division,’’ RNA processing, and translation.

Thanks !