ASSESSING TRANSLATIONAL EFFICIACY THROUGH POLY(A)-TAIL PROFILING AND IN VIVO RNA SECONDARY STRUCTURE DETERMINATION

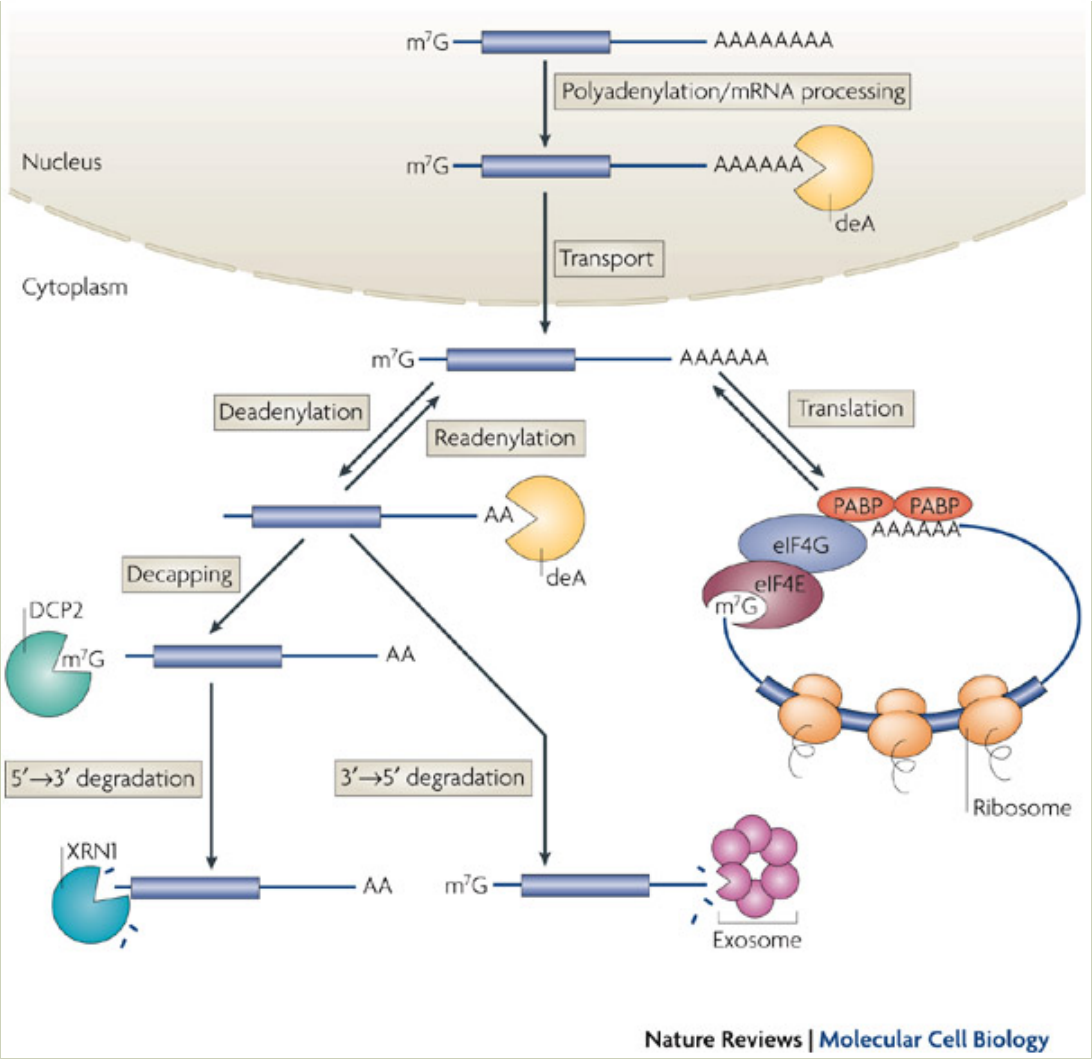
Journal Club,
April 15th 2014

Karl Frontzek,
Institute of
Neuropathology

POLY(A)-TAIL PROFILING

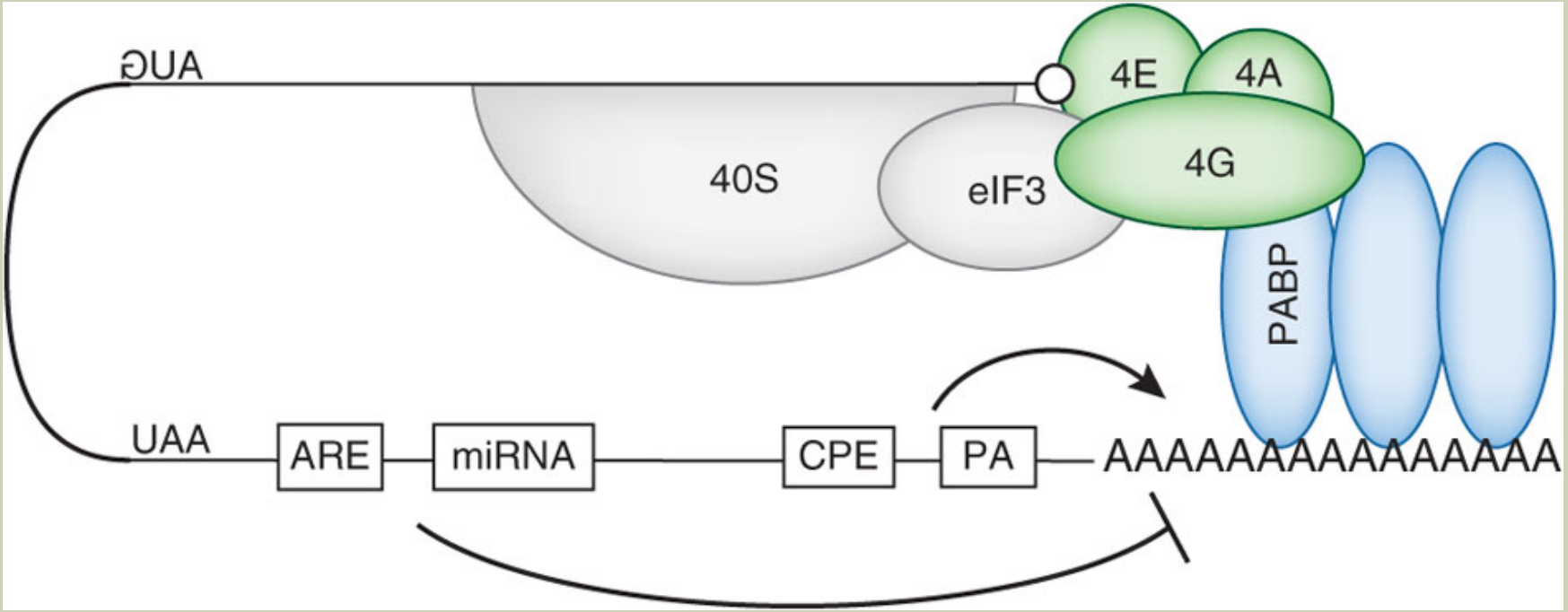
(IMAGES FROM SUBTELNY ET AL., NATURE 2014 UNLESS STATED OTHERWISE)

MULTIFUNCTIONAL DEADENYLASE COMPLEXES DIVERSIFY MRNA CONTROL



Goldstrohm & Wickens – Nat Rev Mol Cell Biol 2008

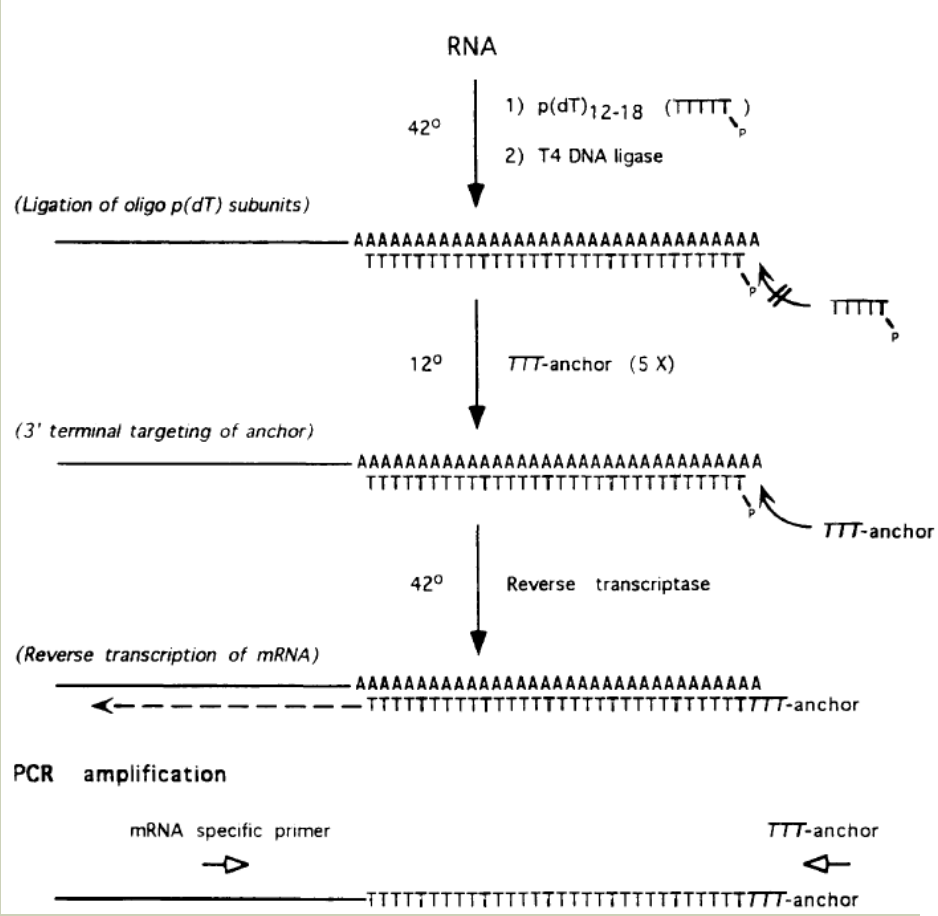
REGULATION OF TRANSLATION INITIATION AND PSEUDO-CIRCULARIZATION OF MRNA



- **Weill et al., Nat Struc Mol Biol 2012**

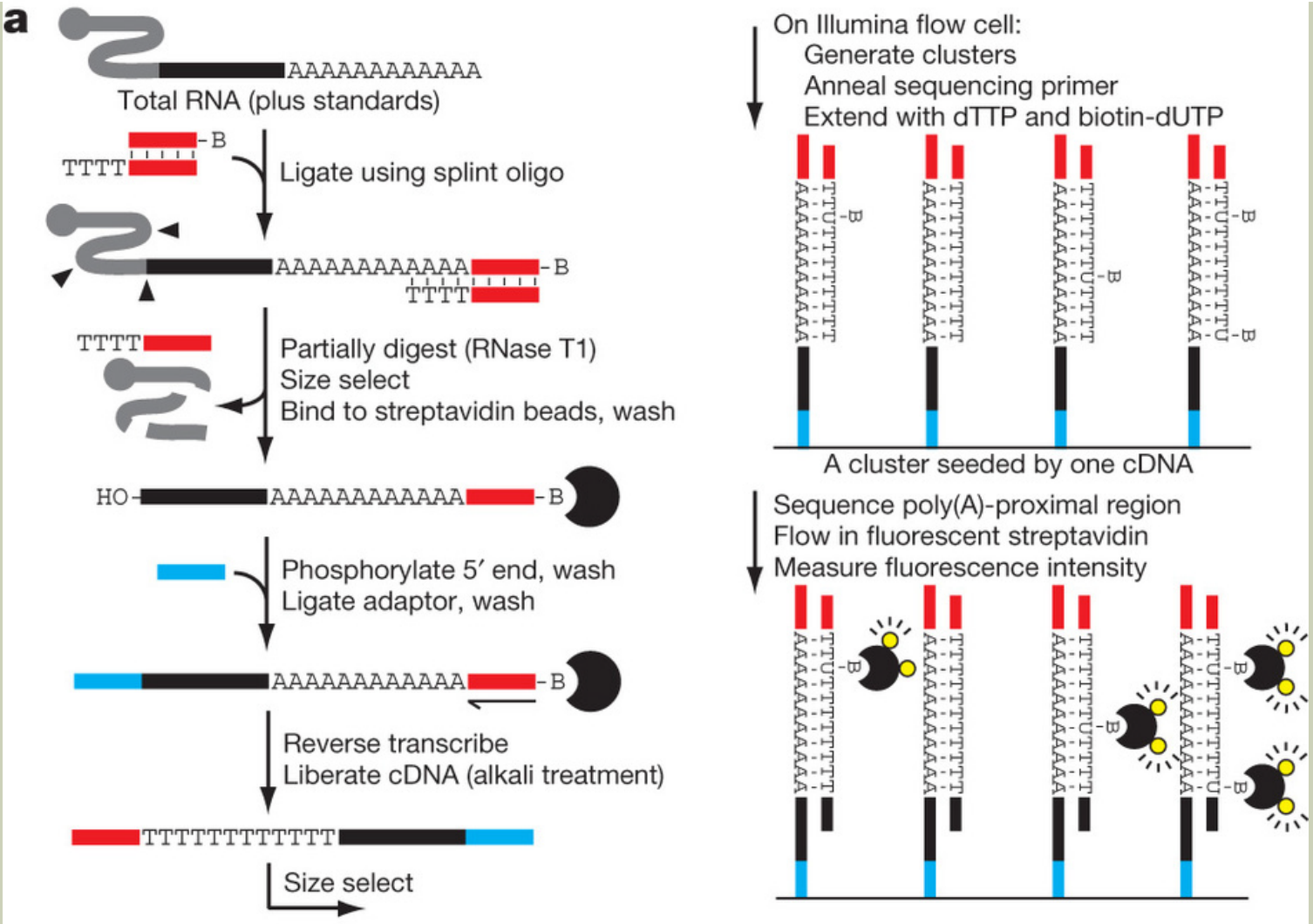
POLY(A)-TAIL LENGTH IN YEAST

- Mature mRNAs in S. cerevisiae have poly(A)-lengths of 70-90 adenosines (vs 150-250 in mammals)
- Generally prevailed view: longer tails > higher translational efficiacy
- Other Poly(A)-tail profiling methods were not feasible in mammalian cells primarily due to length (as yet mostly microarray based) and e.g. relied on
  - discrimination of alternative polyadenylation sites combined with splicing sites
  - stepwise thermal elution from poly(U)-Sepharose (*polyadenylation state microarray, PASTA*)
  - …

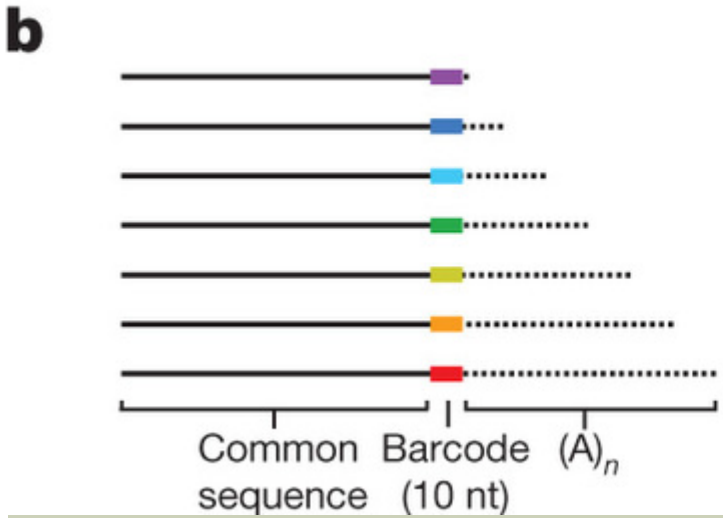


Sallés & Strickland – Genome Res 1995

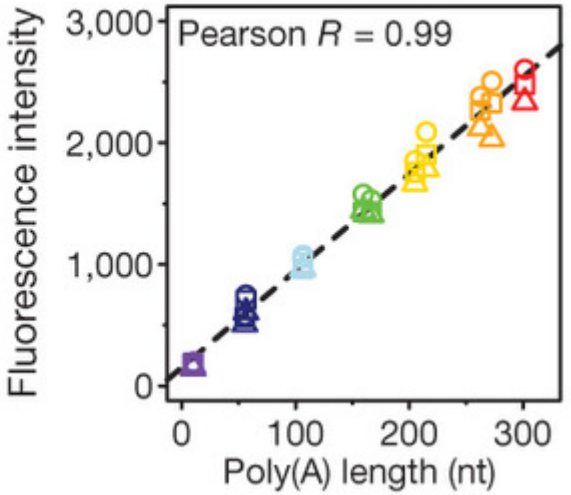
POLY(A)-TAIL LENGTH PROFILING BY SEQUENCING (**PAL-SEQ**)



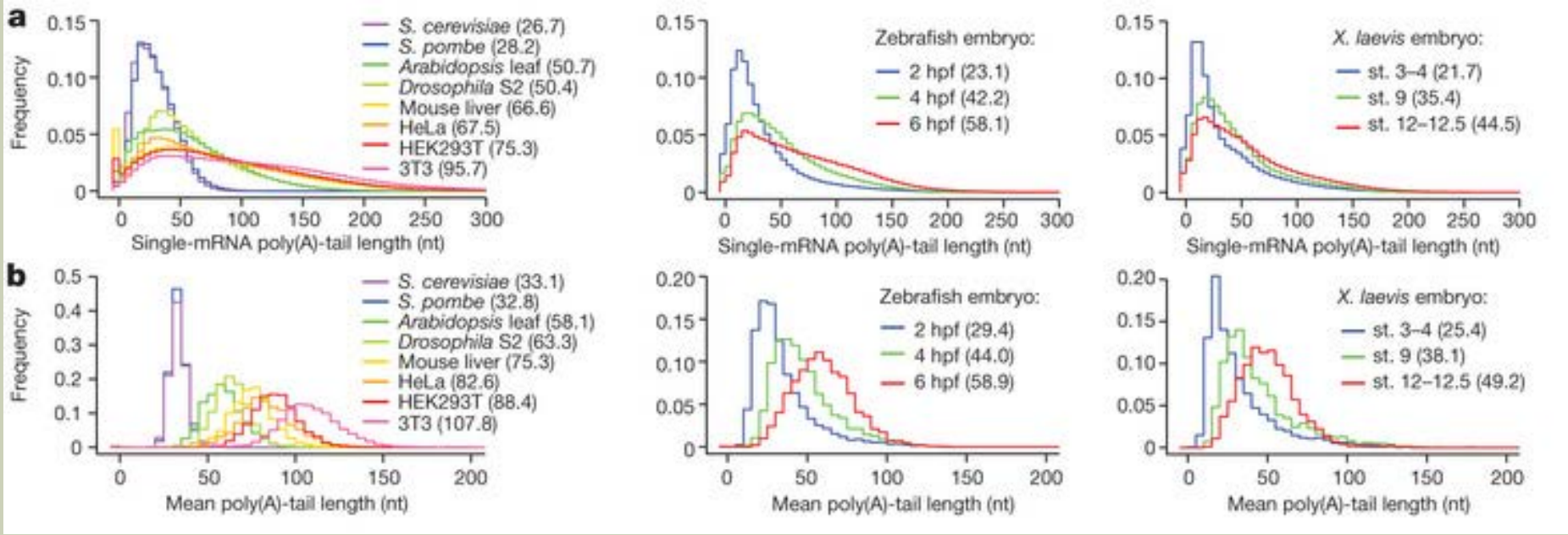
PAL-SEQ WORKFLOW #2



Set 1	Set 2
10.0	10.0
55.5	56.5
106.5	106.5
166.5	159.5
205.0	215.0
262.5	272.5
301.5	



POLY(A)-TAIL PROFILING ACROSS SPECIES

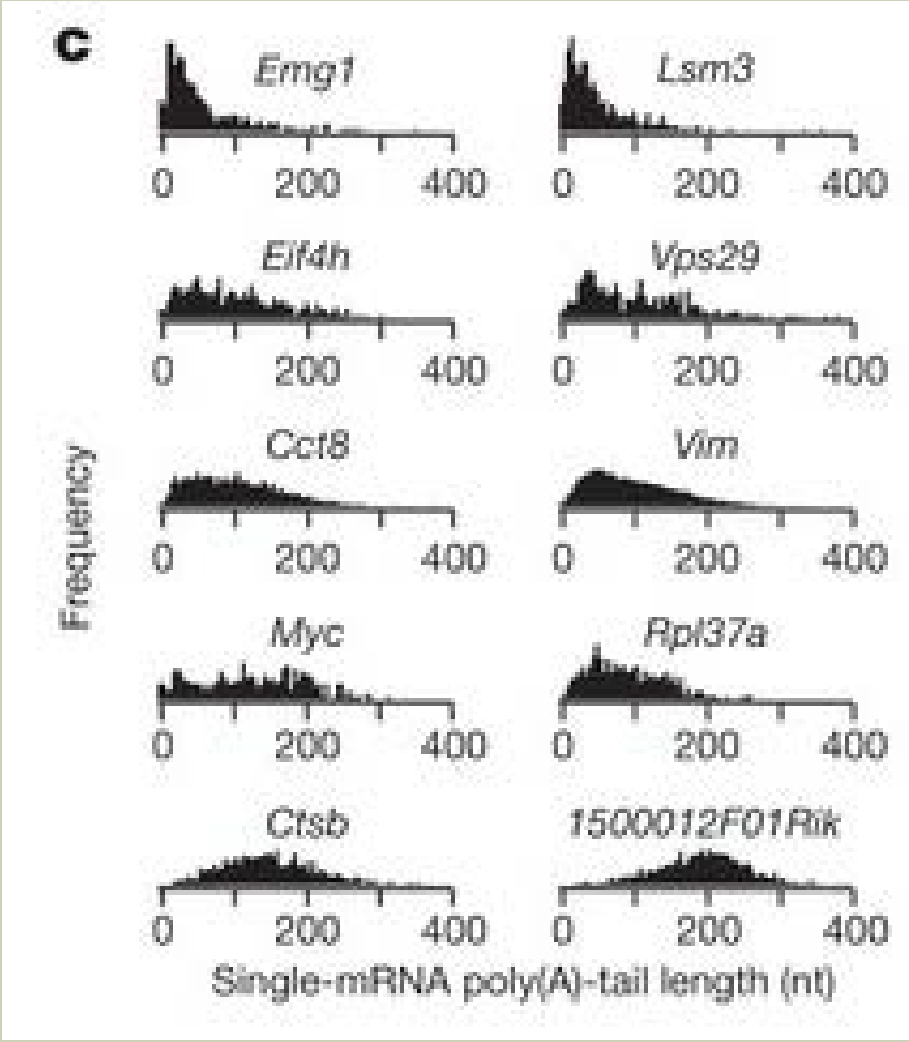


a) poly(A)-tail length pooled and compared across species and developmental stage (for zebrafish and X. laevis embryos)

b) Intergenic tail-length distributions across species and developmental stage

INTRAGENIC TAIL-LENGTH DISTRIBUTION

- Tail-length spectrum of 3T3 genes



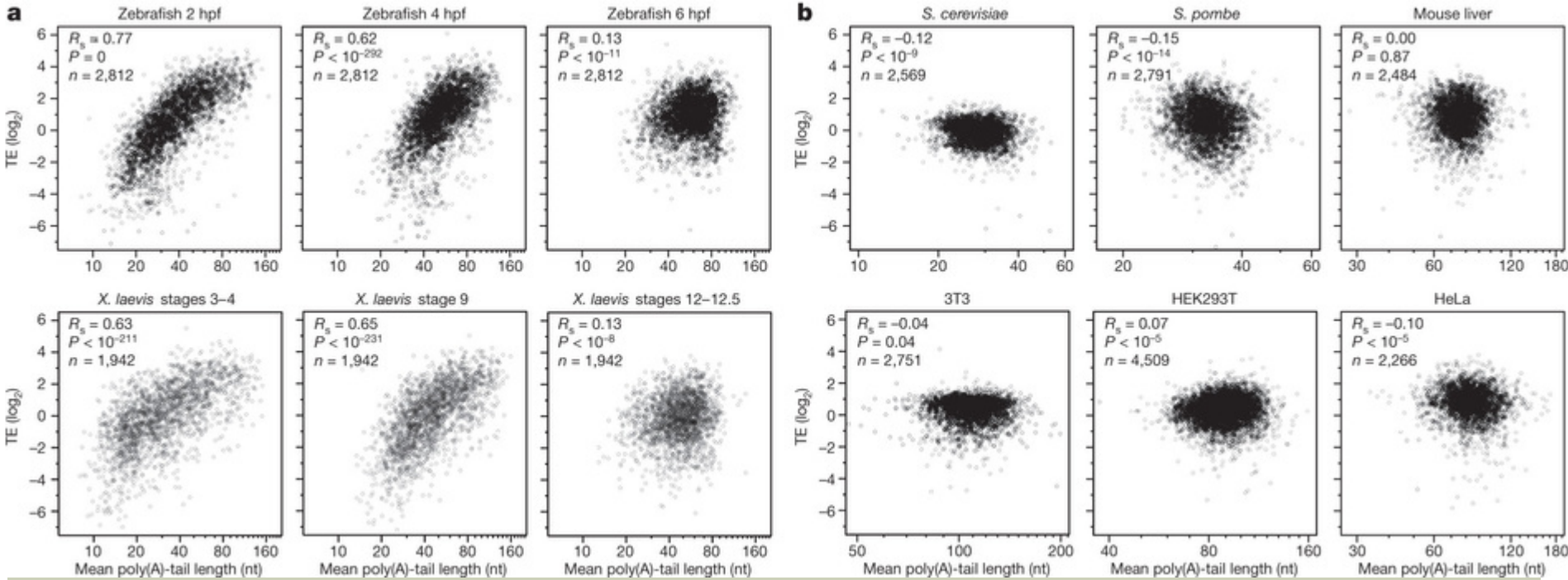
(MODERATE) TAIL-LENGTH CONSERVATION ACROSS SPECIES

- Overall Spearman's R = 0.46
- High enrichment of short-tail genes in for ribosomal (also across species) and other 'housekeeping' genes (i.e. ATP synthesis, not across species); *gene ontology*

Tail length conservation

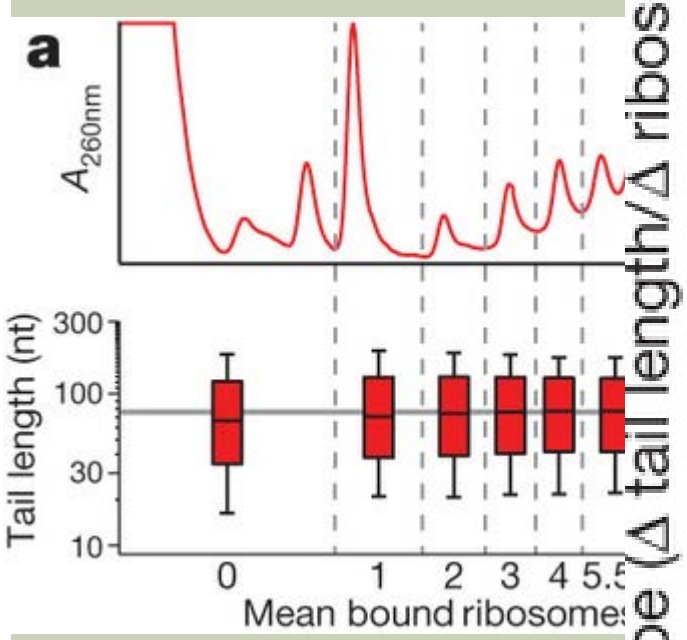
Samples	*n*	R_s	*P* value
HeLa to HEK293T	1620	0.74	$< 10^{-276}$
HeLa to 3T3	1259	0.46	$< 10^{-67}$
HeLa to mouse liver	1095	0.21	$< 10^{-11}$
HeLa to S2	1087	0.16	$< 10^{-6}$
HeLa to *S. cerevisiae*	671	0.056	0.14
HEK293T to 3T3	1907	0.40	$< 10^{-74}$
HEK293T to mouse liver	1815	0.22	$< 10^{-20}$
HEK293T to S2	1877	0.16	$< 10^{-11}$
HEK293T to *S. cerevisiae*	1221	0.078	0.0063
3T3 to mouse liver	1548	0.37	$< 10^{-51}$
3T3 to S2	1194	0.20	$< 10^{-11}$
3T3 to *S. cerevisiae*	737	0.11	0.0034
mouse liver to S2	1238	–0.068	0.016
mouse liver to *S. cerevisiae*	784	0.0028	0.94
S2 to *S. cerevisiae*	959	0.094	0.0038
S. pombe to *S. cerevisiae*	1379	0.22	$< 10^{-15}$

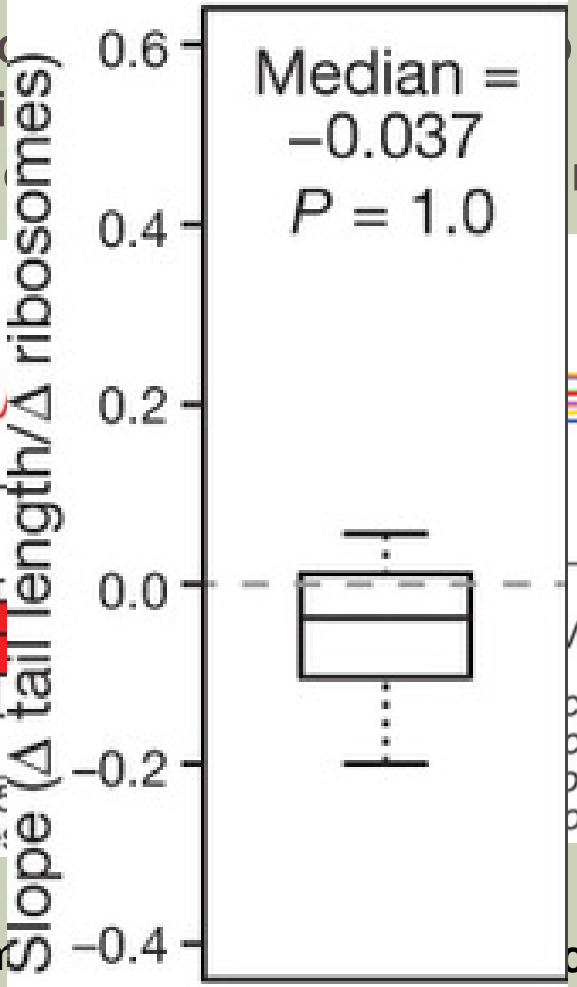
STRONG BUT TRANSIENT CORRELATION OF POLY(A)-TAIL LENGTH WITH TRANSLATIONAL EFFICIACY AS MEASURED THROUGH RIBOSOME FOOTPRINTING

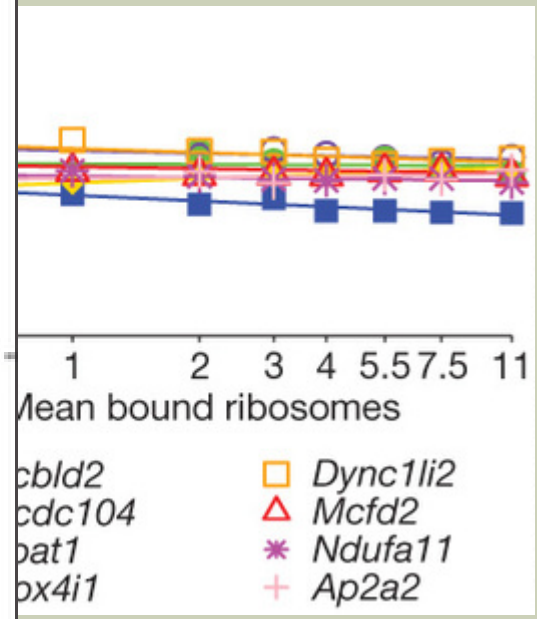


TAIL-LENGTH AND TRANSLATION #2

- Weak to negative c…oly(A)-tail length and translational effici…rough data pooling with lack of intrag…re: 3T3 cell lysates)







Stepwise sucrose gradient … between tail length and fractination with subsequent m… bound per mRNA for poly(A)-tail profiling … mRNAs from the same gene

CONCLUSIONS

- Poly(A)-tail profiling through PAL-SEQ refutes as yet reported findings of broadly assumed proportional relationsship between tail length and translational efficiacy
- Two regulatory regimes were observed: (I) in yeast, mammalian and gastrulation-stage cells (II) in metazoan embryos

(I) TRANSCRIPTIONALLY ACTIVE CELLS
(YEAST, MAMMALIAN & GASTRULATION-STAGE CELLS)

- Have lots of opportunities for nuclear control of gene expression
- At high transcriptional rates, instable mRNAs can be replaced, putting mRNA stability into regulatory focus
- Old mRNAs mostly tend to have shorter poly(A) tails due to the absence of cytoplasmatic polyadenylation, making poly(A) length as a regulatory feature dispensable
- In this regime, older mRNAs would have less value if they are translated to a lesser amount because of their shorter tails

(II) METAZOAN EMBRYOS

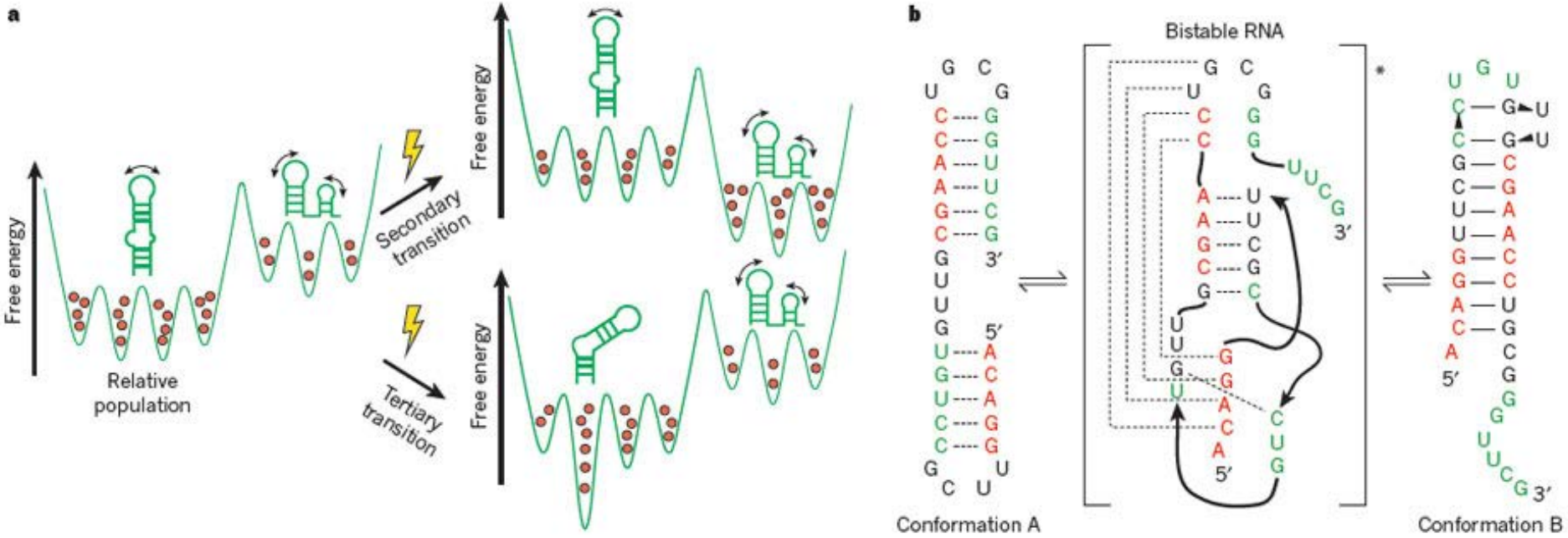
- Embryos were transcriptionally inactive
  - Precludes use of transcription and other nuclear programs to alter gene expression programs
  - Also precludes mRNA stability as regulatory factor because old mRNA can only be replaced when zygotic transcription begins
- Apparently these early embryonic cells use differential tail length to control gene expression (many mRNA with short tails observed at this stage – see above)
- Consistently, early embryonic cells were shown to have robust cytoplasmatic polyadenylation
- Cleavage-stage embryos also had more uniform intragenic tail lengths and more variable intergenic tail lengths than cells subjected to the other regime (i.e., (I)
- **Poly(A)-profiling may prove a useful tool for regulatory function in other systems that show transcriptional repression with active cytoplasmatic polyadenylation (i.e. early embryos of other metazoan species, maturing oocytes and neuronal synapses)**

IN VIVO RNA 2ND STRUCTURE DETERMINATION

RNA DYNAMICS

EQUILIBRIUM FLUCTUATIONS VS CONFORMATIONAL TRANSITIONS

- Although the 2 states are intricately related, EF are due to cellular thermal variations while cellular cues – as changes in metabolite concentrations (> bolts) - can alter the free energy landscape thus leading to CT



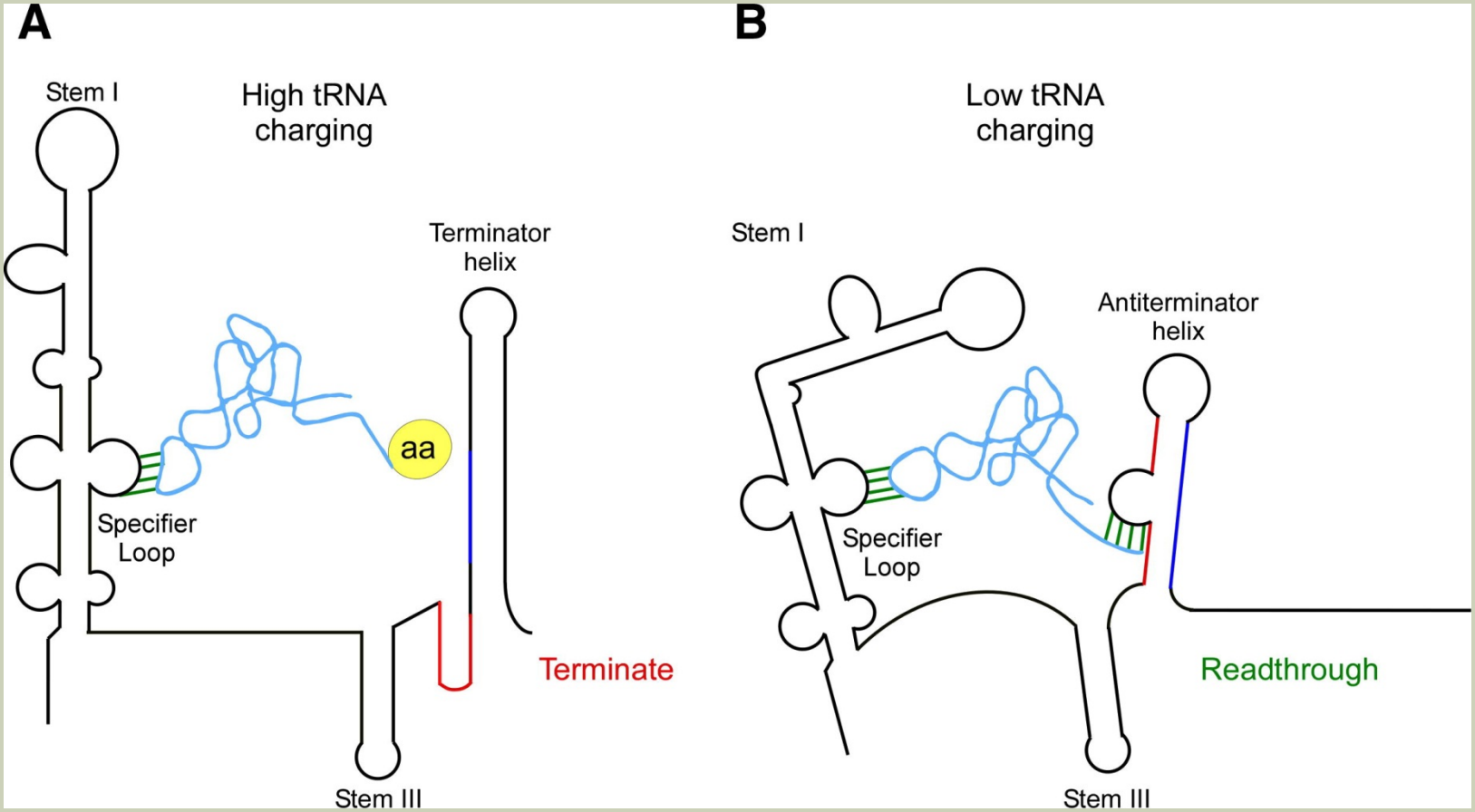
Dethoff EA et al., Nature 2012

RNA CONFORMATIONAL TRANSITIONS

- Triggering cues: proteins binding/liberating RNA resulting in (de-)stabilization of conformation, stabilizing catalytic activity
- RNA 'chaperones' and helicases
- Metabolites - AAs, coenzymes, nucleotides
- Physiochemical conditions – $[Mg^{2+}]_i$, pH (>> riboswitches)
- Thermosensors …

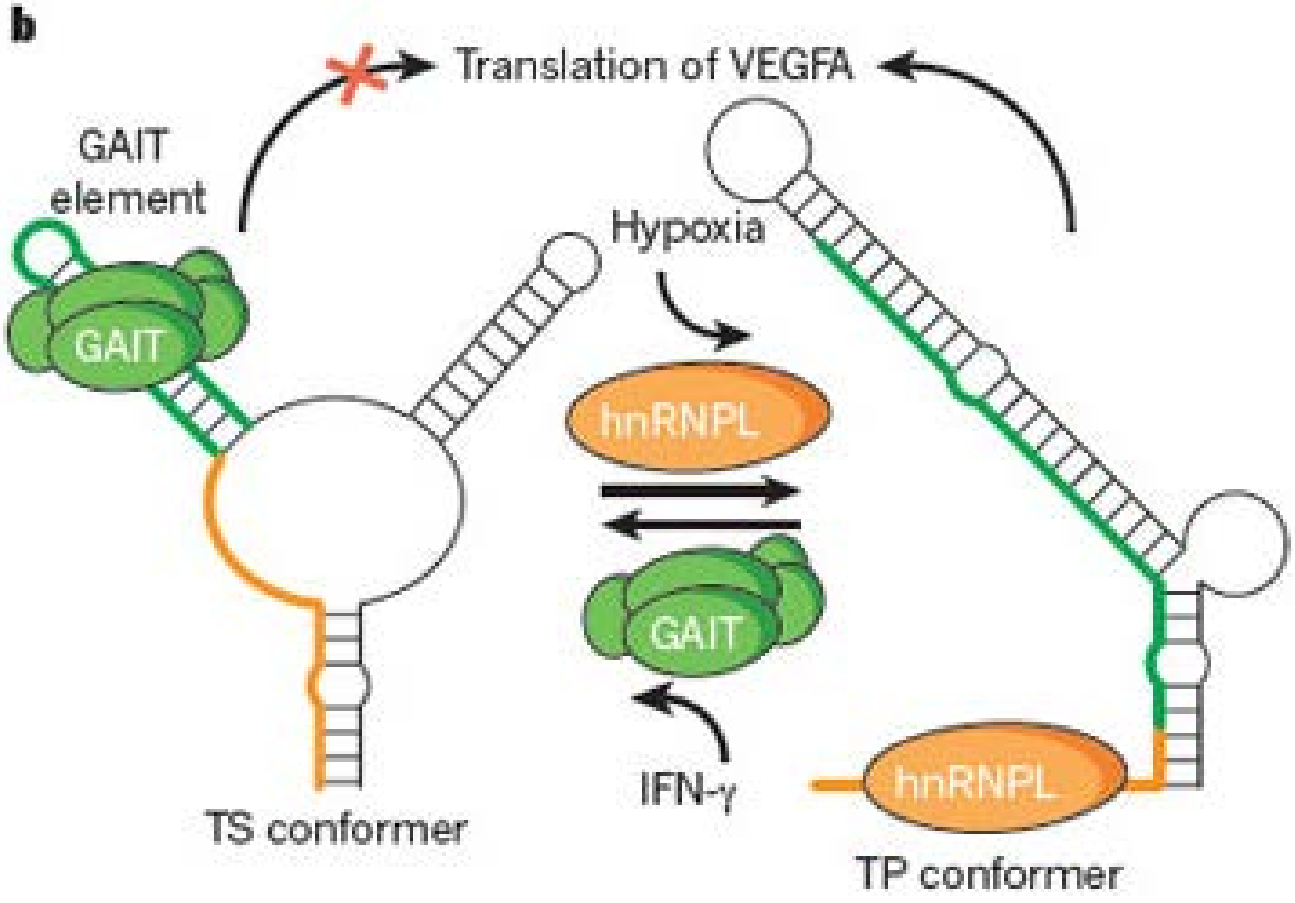
FUNCTIONS OF SECONDARY STRUCTURAL TRANSITIONS – THE T-BOX MECHANISM

- Transcription-terminating helices



Green NJ et al., FEBS Let 2010

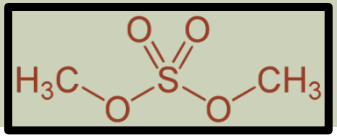
FUNCTIONS OF SECONDARY STRUCTURAL TRANSITIONS – TRANSLATION OF VEGF-A

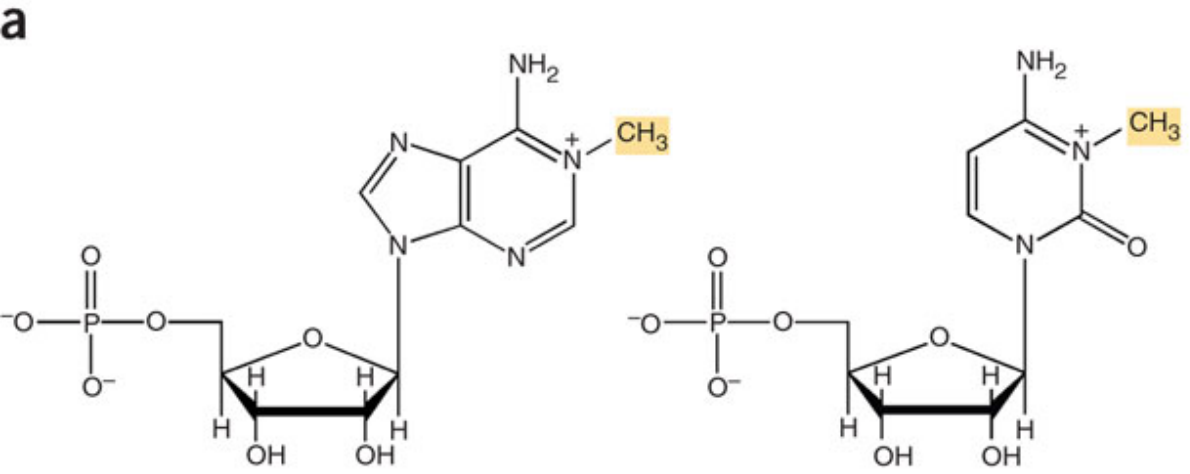


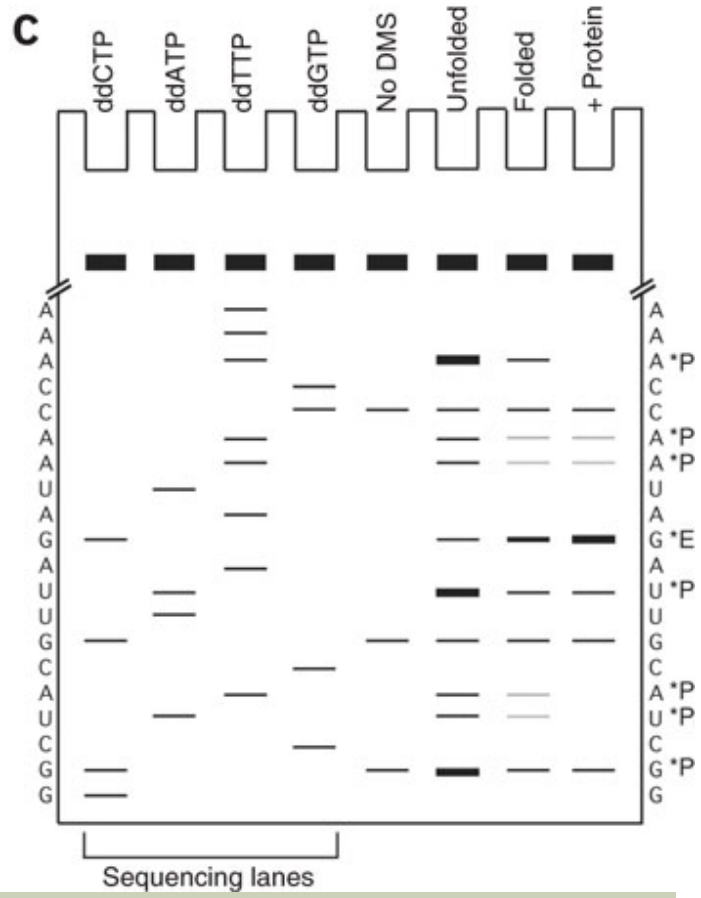
- From Dethoff EA et al., Nature 2012 adapted from Ray PS et al., Nature 2009

CHEMICAL METHODS FOR NUCLEIC ACID STRUCTURE DETERMINATION - DMS

- Dimethylsulfate footprinting

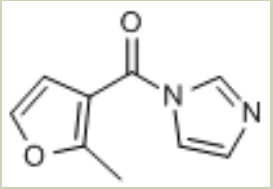




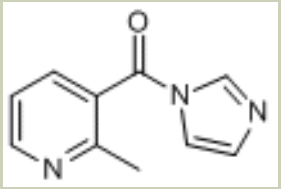


CHEMICAL METHODS FOR NUCLEIC ACID STRUCTURE DETERMINATION - SHAPE

- **2' OH** group of ribose in RNA is universal feature amongst all 4 RN acids
  - High 2' hydroxyl reactivity indicates single-stranded or flexible RNA regions
  - Low 2'OH reactivity usually indicates base pairing or other interactions
- **SHAPE - S**elective 2′-**h**ydroxyl **a**cylation analyzed by **p**rimer **e**xtension (detects 2' OH groups of all ribonucleic acids)
- **FAI & NAI** – acylation electrophiles with selective reaction towards hydroxyl groups, soluble at high concentrations and useful in living cells (Spitale RC et al., Nat Chem Biol 2012)

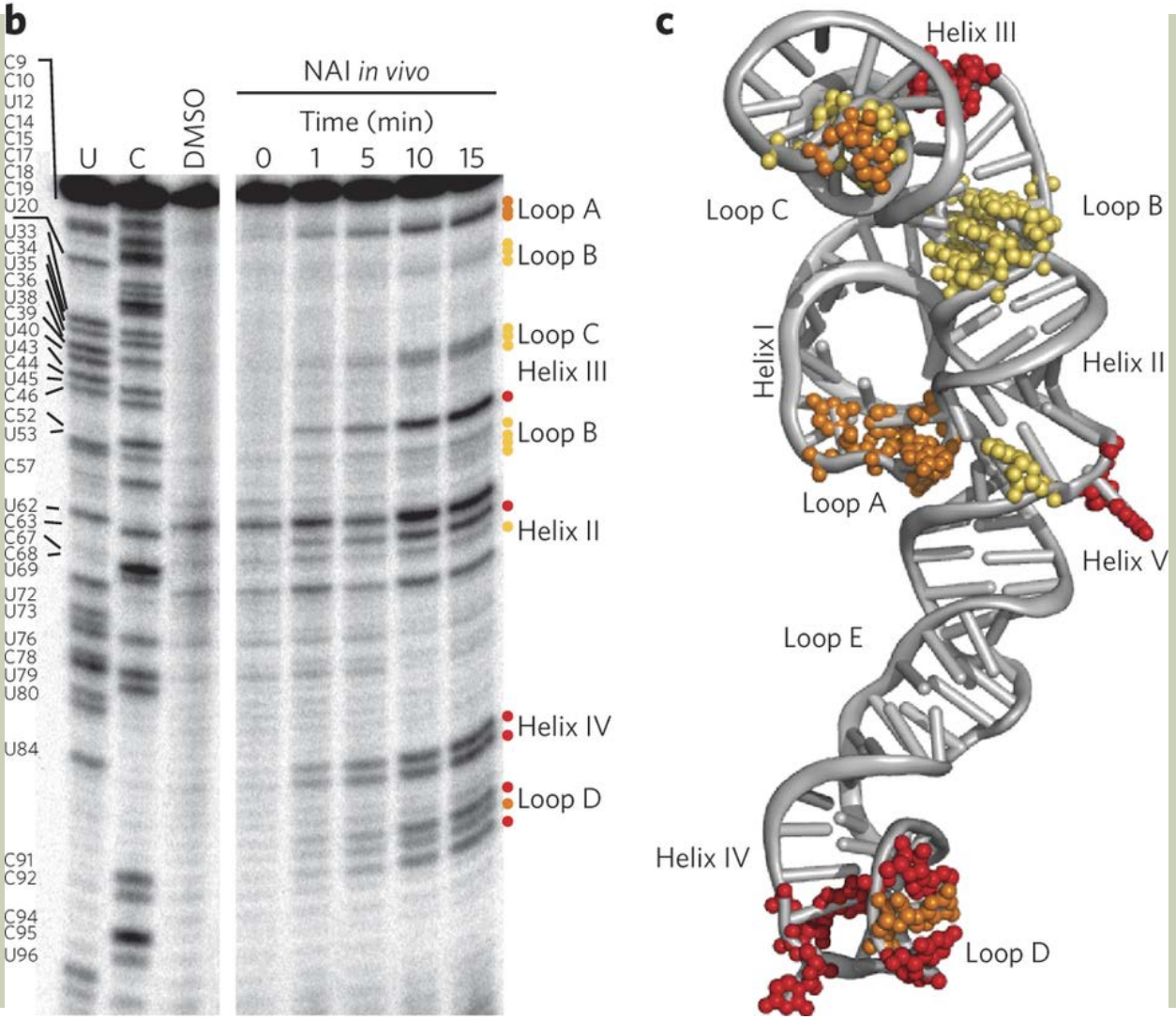


2-methyl-3-furoic acid imidazolide, FAI



2-Methylnicotinic acid imidazolide, NAI

SHAPE PROBING OF 5S_RRNA IN MESC

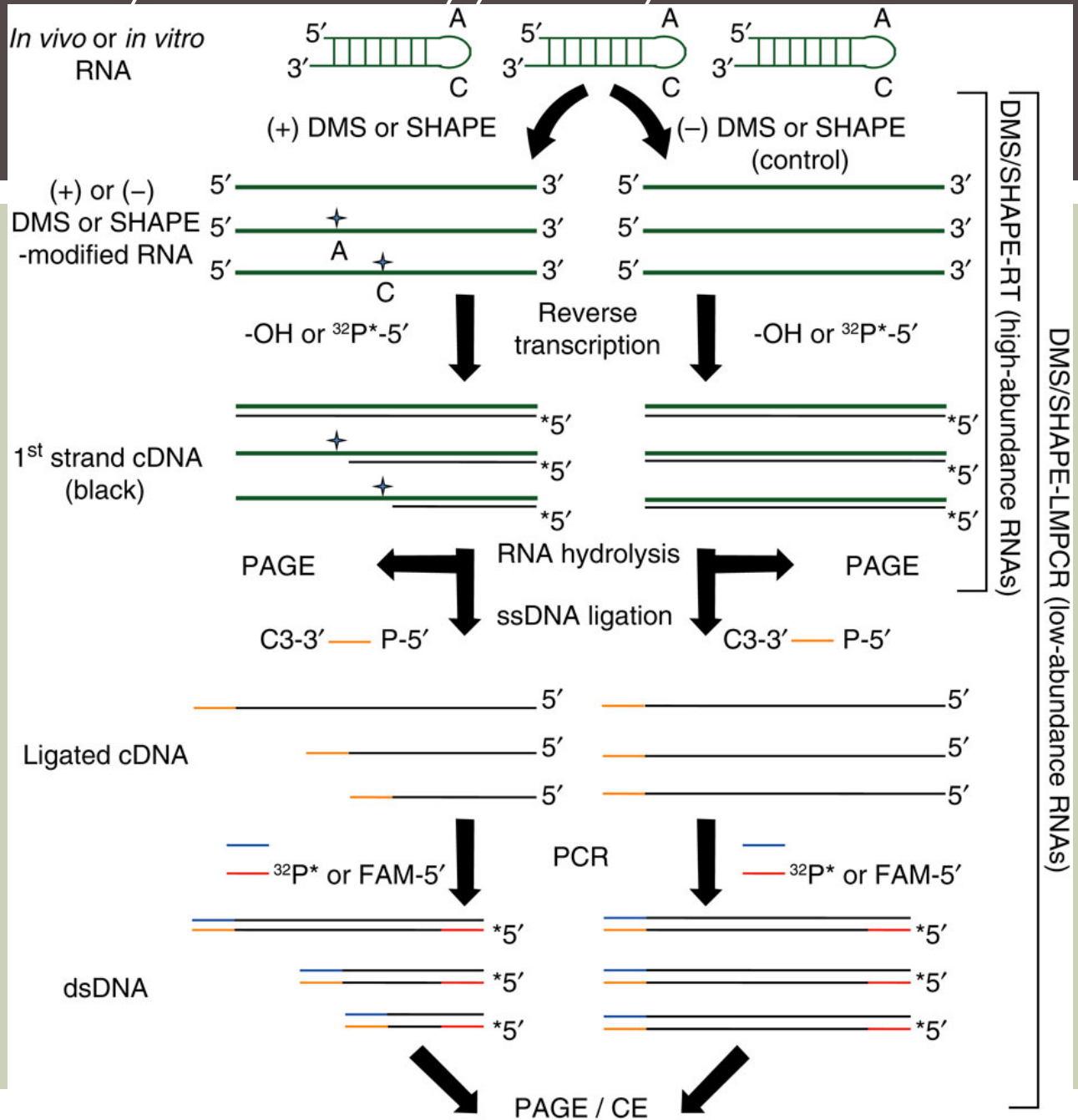


DMS/SHAPE-RT // DMS/SHAPE-LMPCR
KWOK CK ET AL., NAT COMM 2013

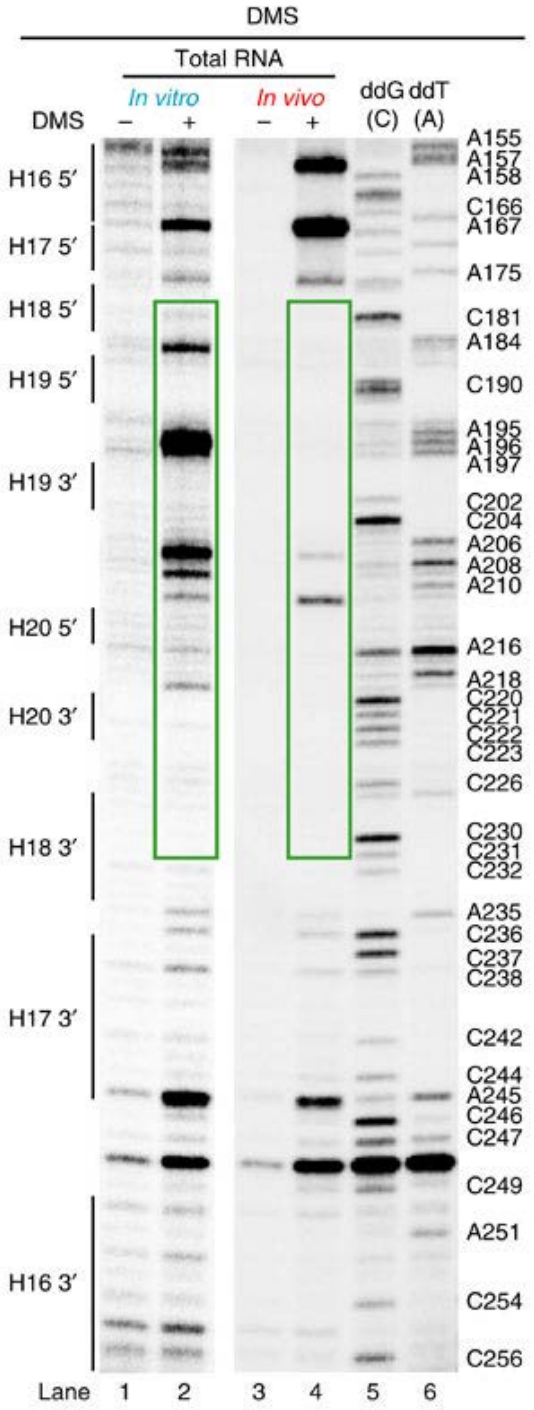
- Goal: to probe low-abundance RNAs (i.e. most mRNAs, ncRNA..)

1. DMS/SHAPE-RT in vivo in A. thaliana for high abundance RNAs (25S rRNA, 5.8S rRNA, chloroplast mRNA)
2. DMS/SHAPE-LMPCR in vivo for low abundance RNA (GRP3S mRNA, protein binding & ncRNA U12, a small nuclear RNA [snRNA])

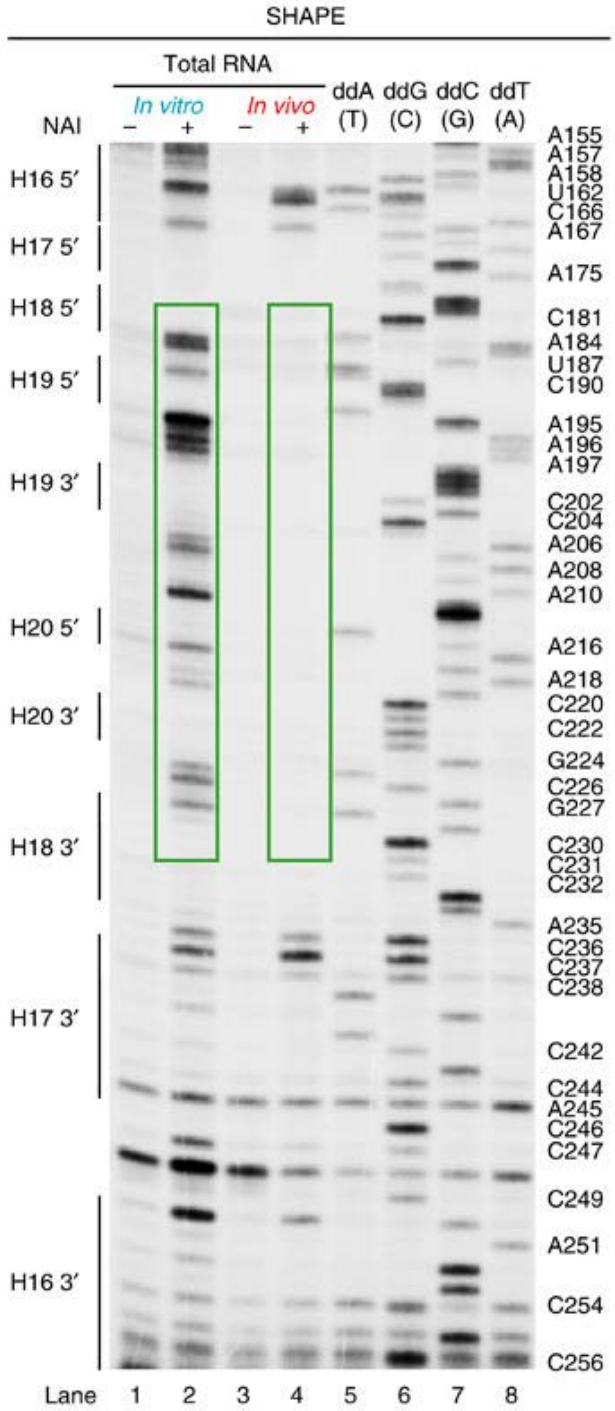
DMS/SHAPE-RT // DMS/SHAPE-LMPCR



- *In vitro* vs *in vivo* probing of helices H16-H20 of 25s rRNA of A. thaliana - DMS

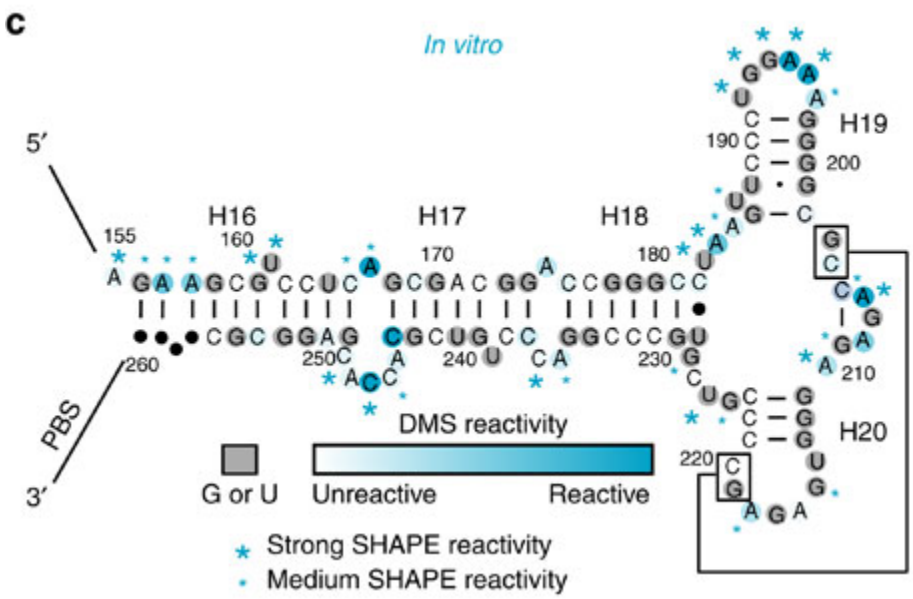


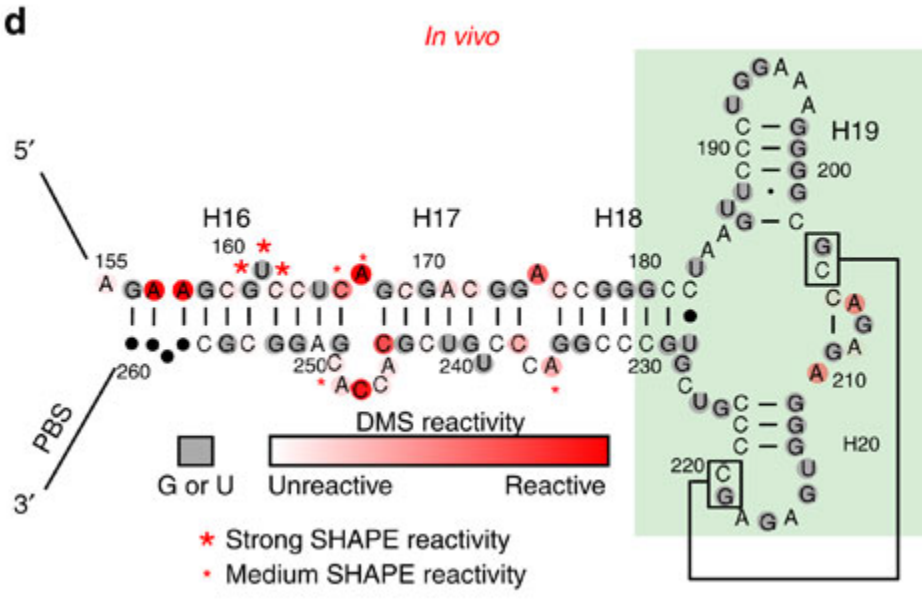
- *In vitro* vs *in vivo* probing of helices H16-H20 of 25s rRNA of A. thaliana - SHAPE



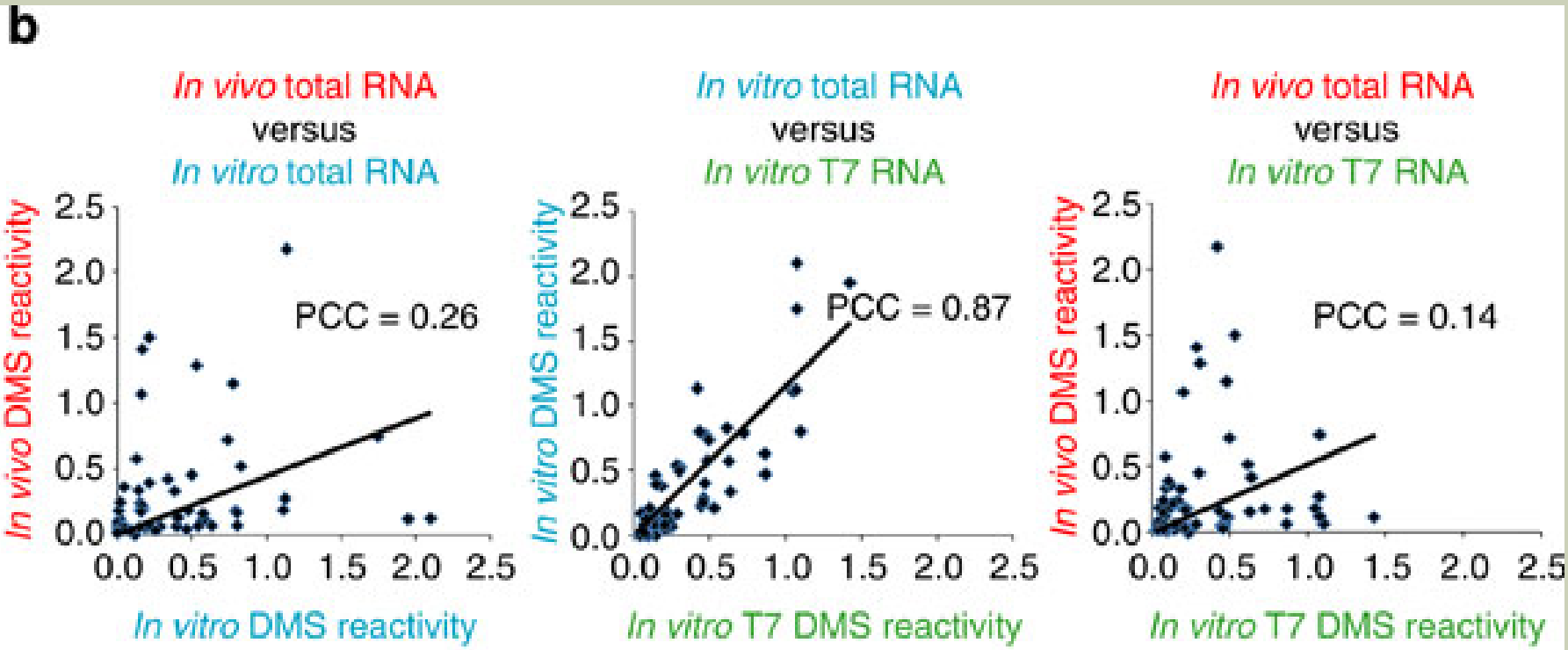
BINDING OF RIBOSOMAL PROTEINS EXPLAINS DIFFERENTIAL CONFORMATIONS *IN VITRO* AND *IN VIVO*

- Crystal structure of yeast ribosome was examined (conserved with A. thalania and other eukaryotes)
- H19-H20 (i.e. those helices with diverging DMS&SHAPE patterns) showed high binding of ribosomal proteins

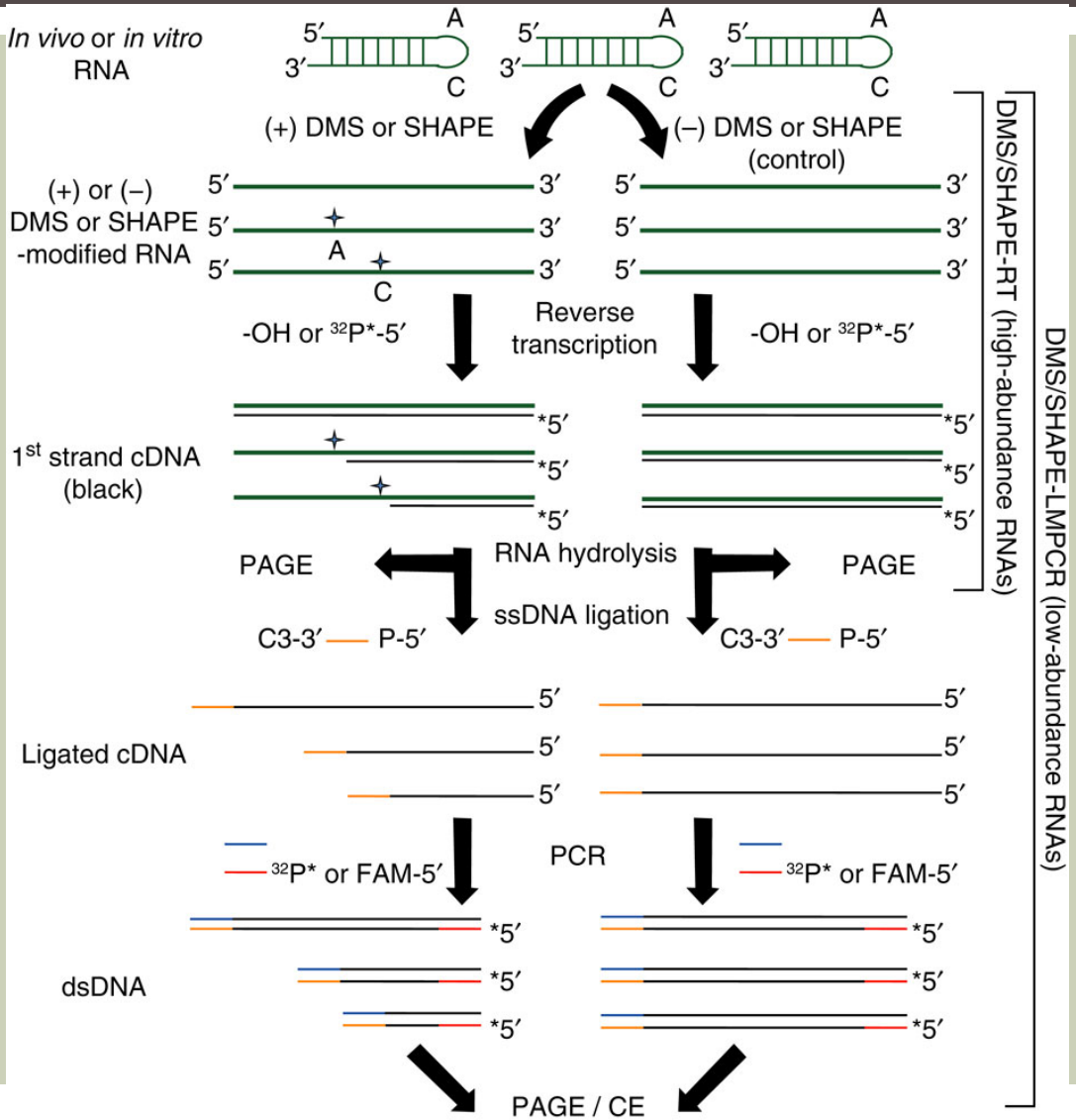




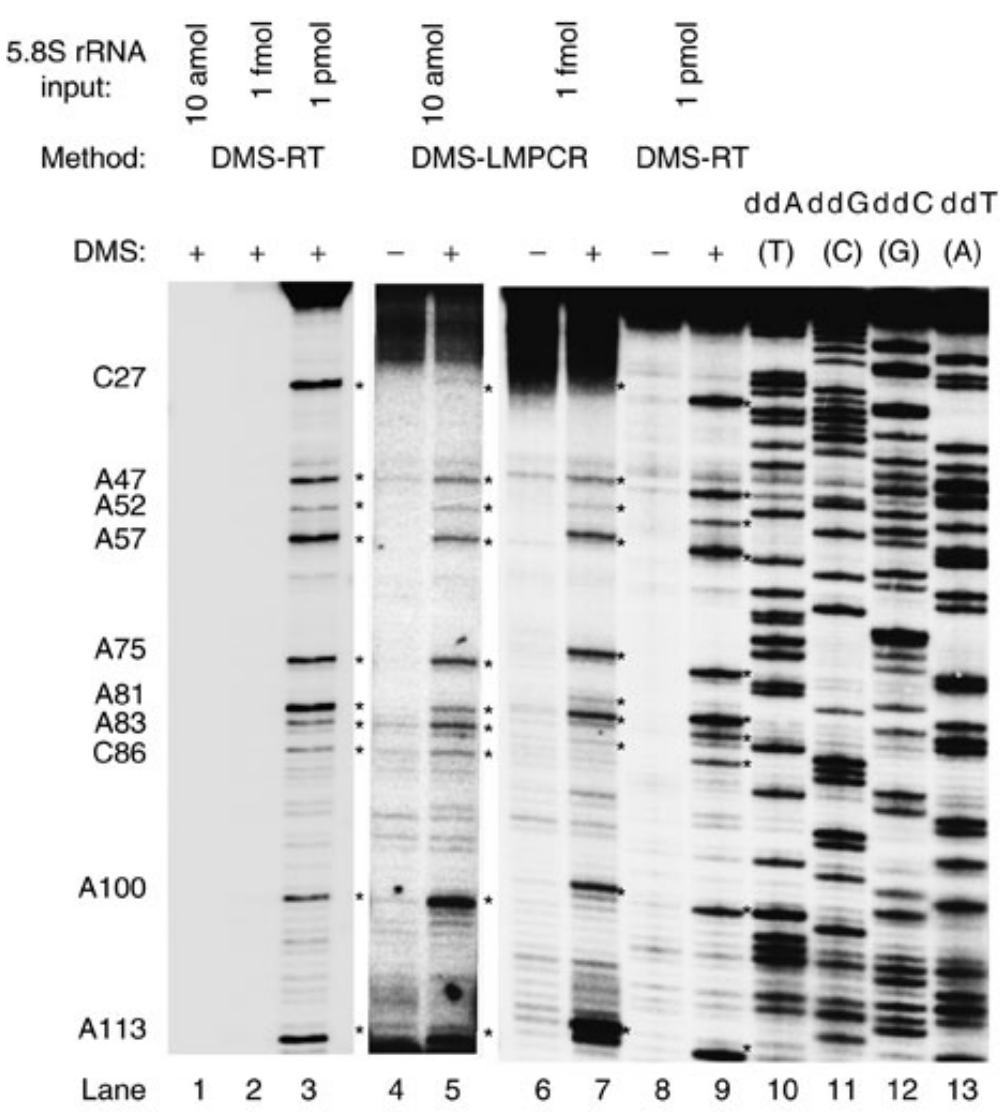
DMS PROBING OF 5.8S $_R$RNA ALSO YIELDS DIVERGING RNA STRUCTURES *IN VITRO* VS *IN VIVO*

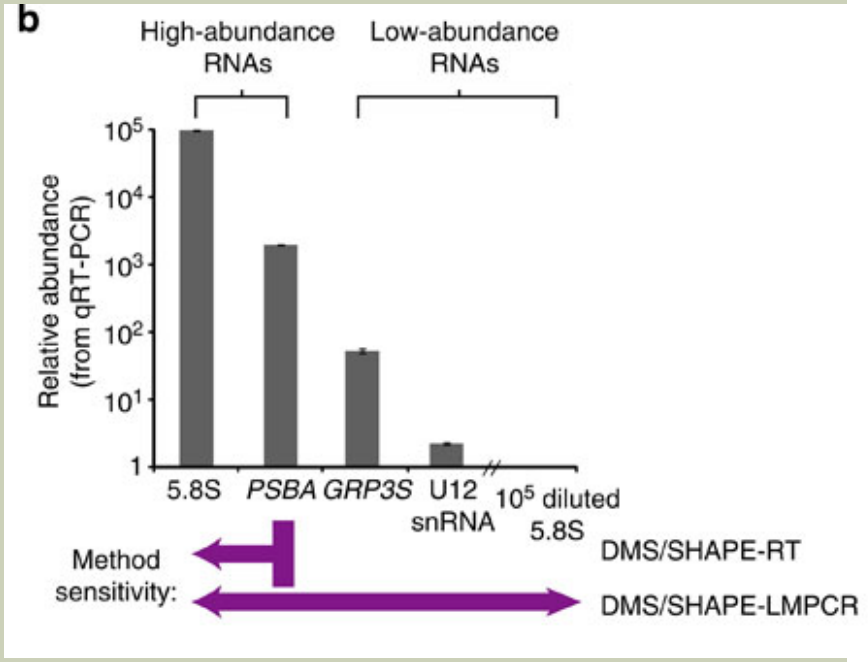


INTRODUCTION OF LMPCR REVEALS RNA STRUCTURES AT 10-ATTOMOLE LEVELS (10^{-17})

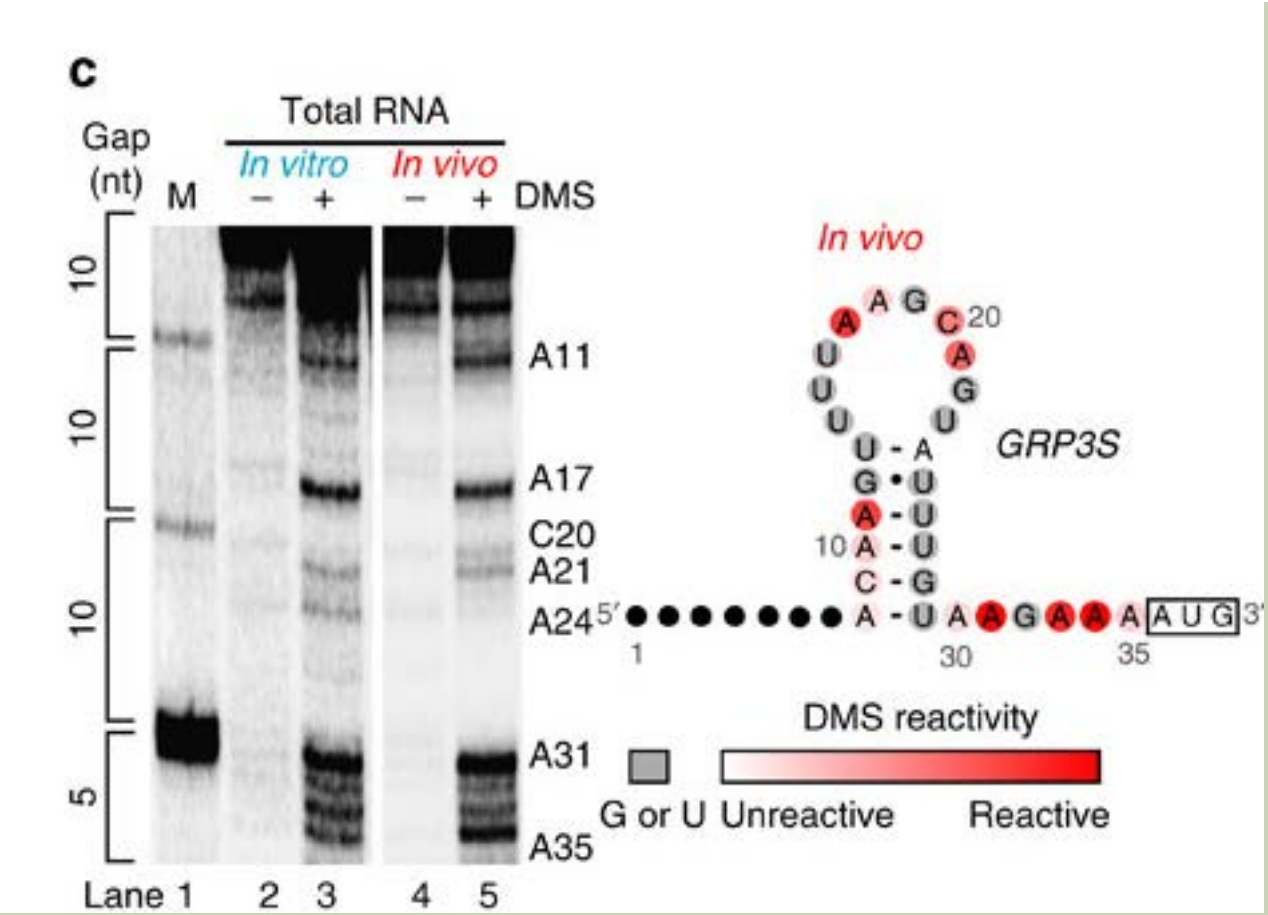


INCREASING DMS PROBING SENSITIVITY BY 5 LOGS





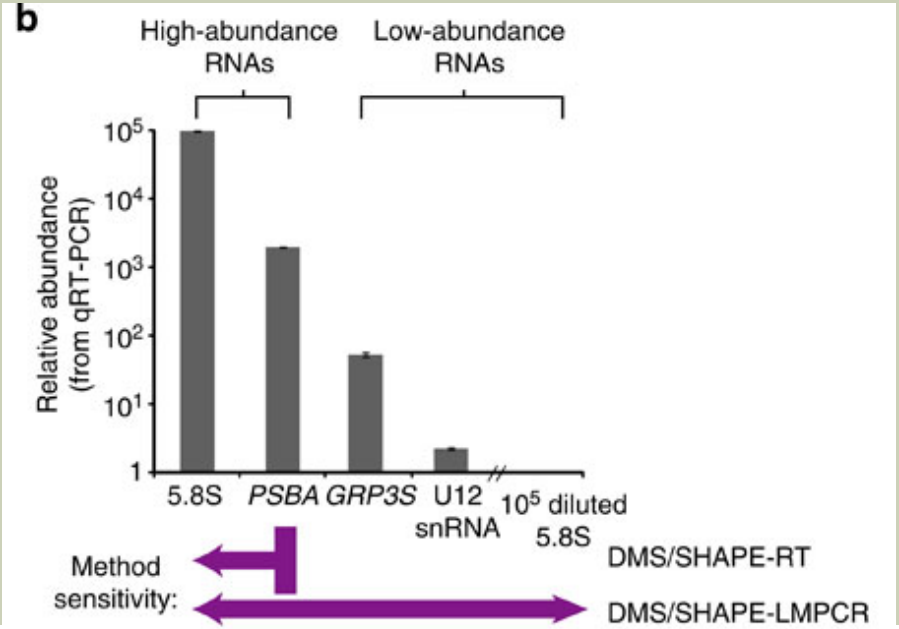
DMS PROBING OF GRP3S
(1.900-FOLD LOWER CONC. THAN 5.8S $_R$RNA)



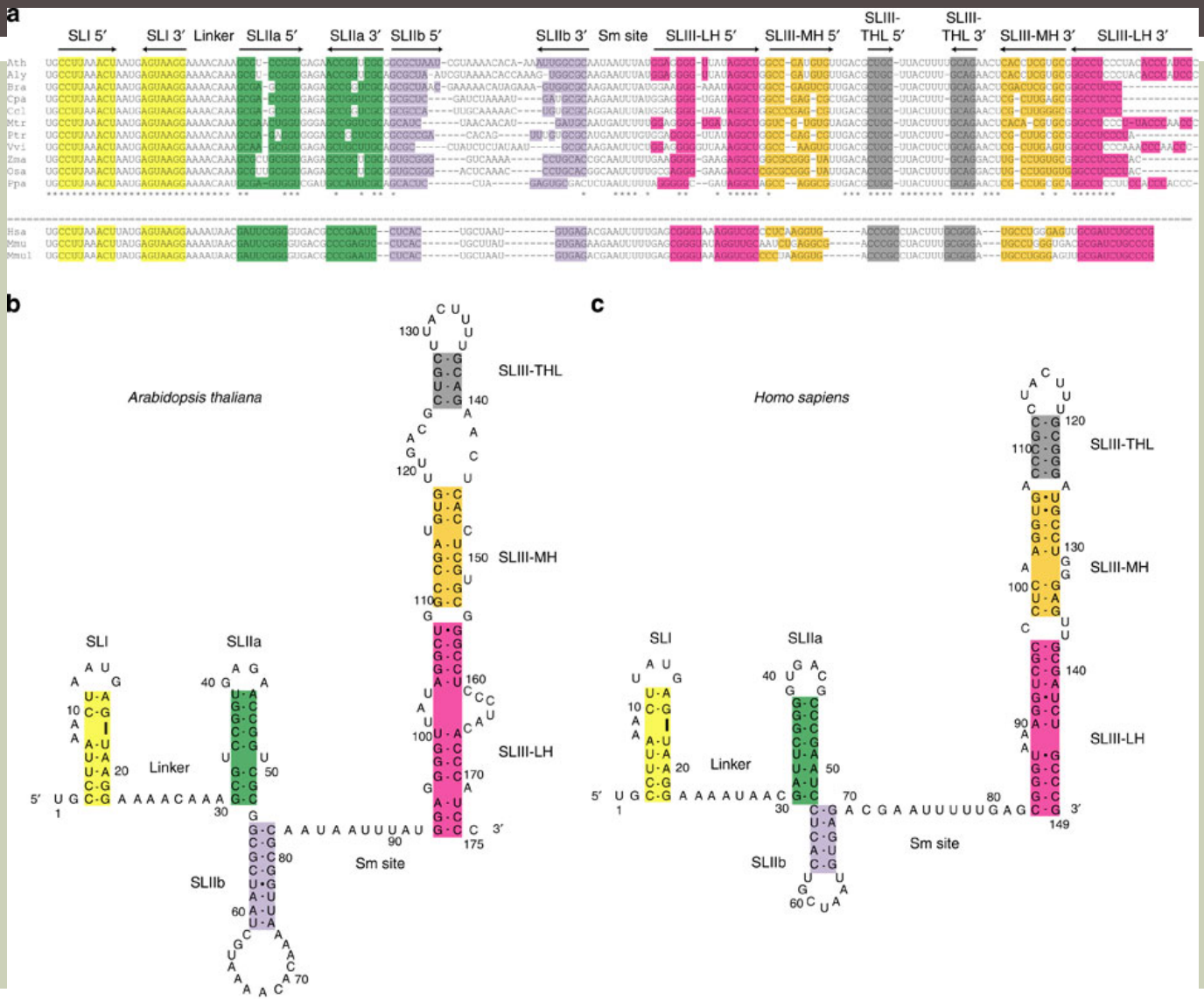
Pearson's correlation coefficient (in vitro vs in vivo)=0.84

DMS/SHAPE-LMPCR PROBING OF U12

- U12 is a nc snRNA of the minor spliceosomal complex, responsible for the splicing of a divergent class of pre-mRNA introns
- Levels of U12 snRNA in A. thaliana are ~45.000 lower than those of 5.8S rRNA

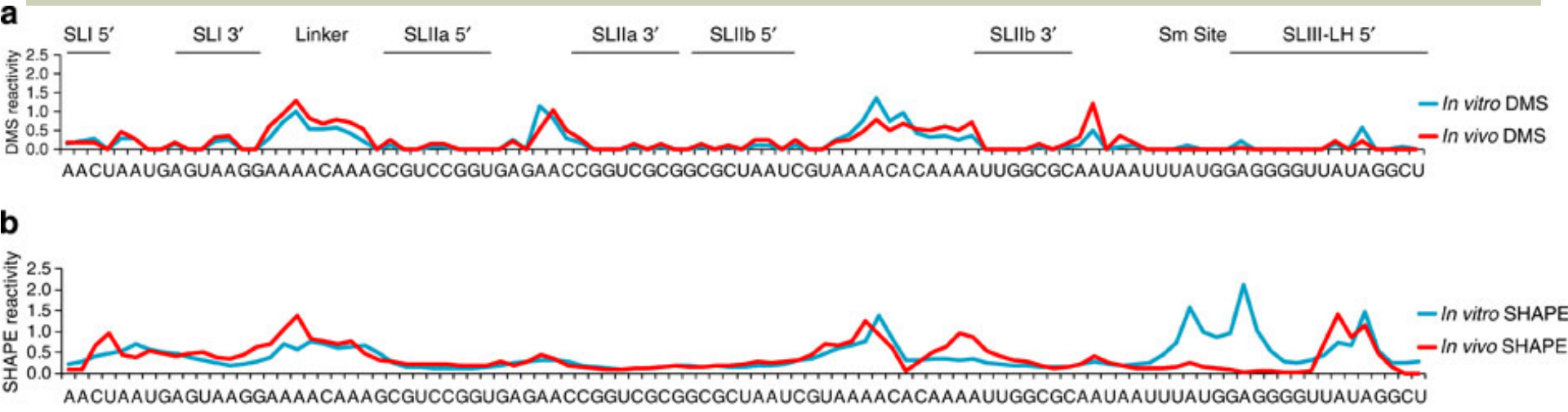


CONSERVATION OF U12 2ND STRUCTURE THROUGH COMPARATIVE SEQUENCE ANALYSIS



DMS/SHAPE LMPCR REVEALS IN VIVO PROTEIN BINDING SITE OF LOW-ABUNDANCE $_{NC}$RNA U12

- Sm site has been reported to be important for stable small nuclear ribonucleoparticle formation

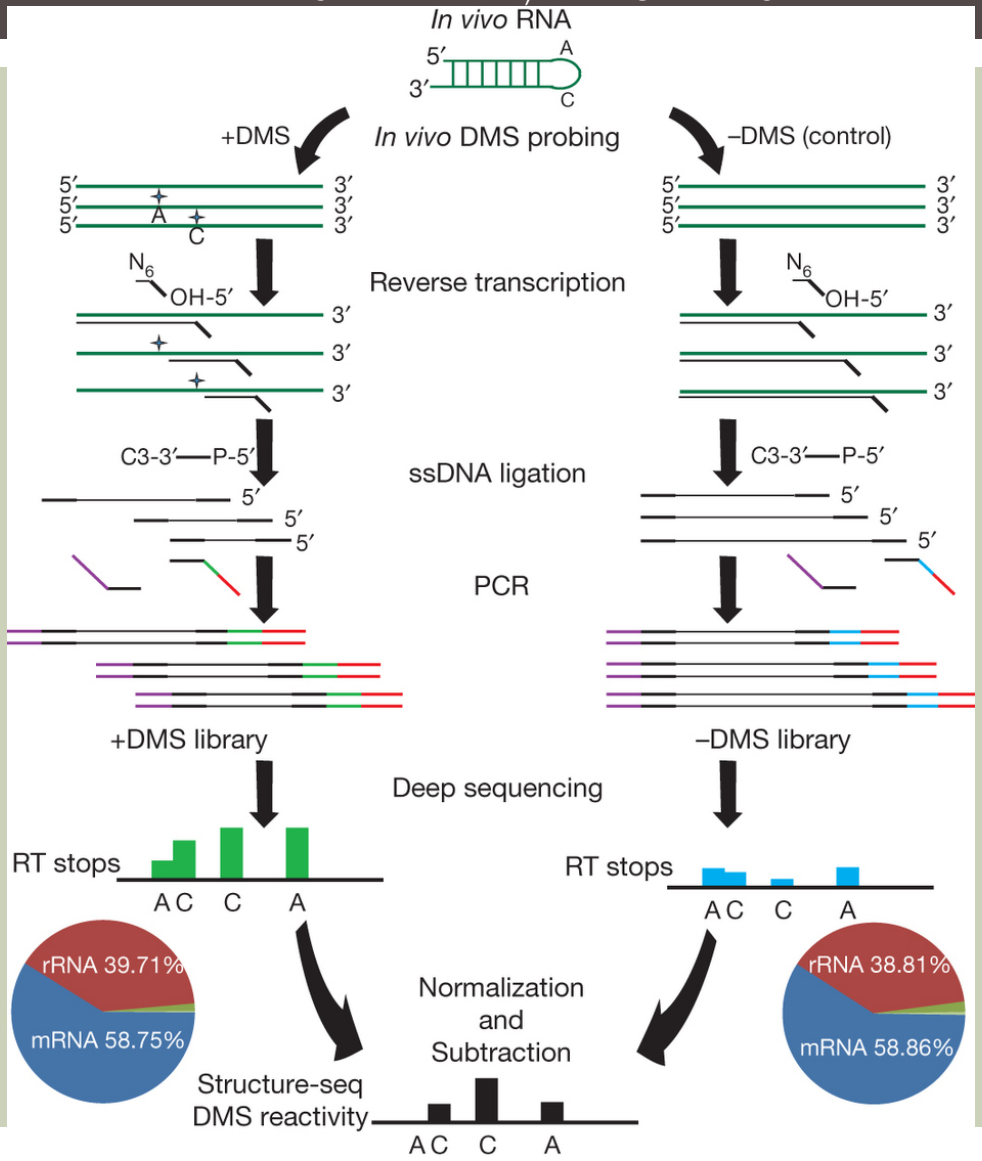


CONCLUSIONS

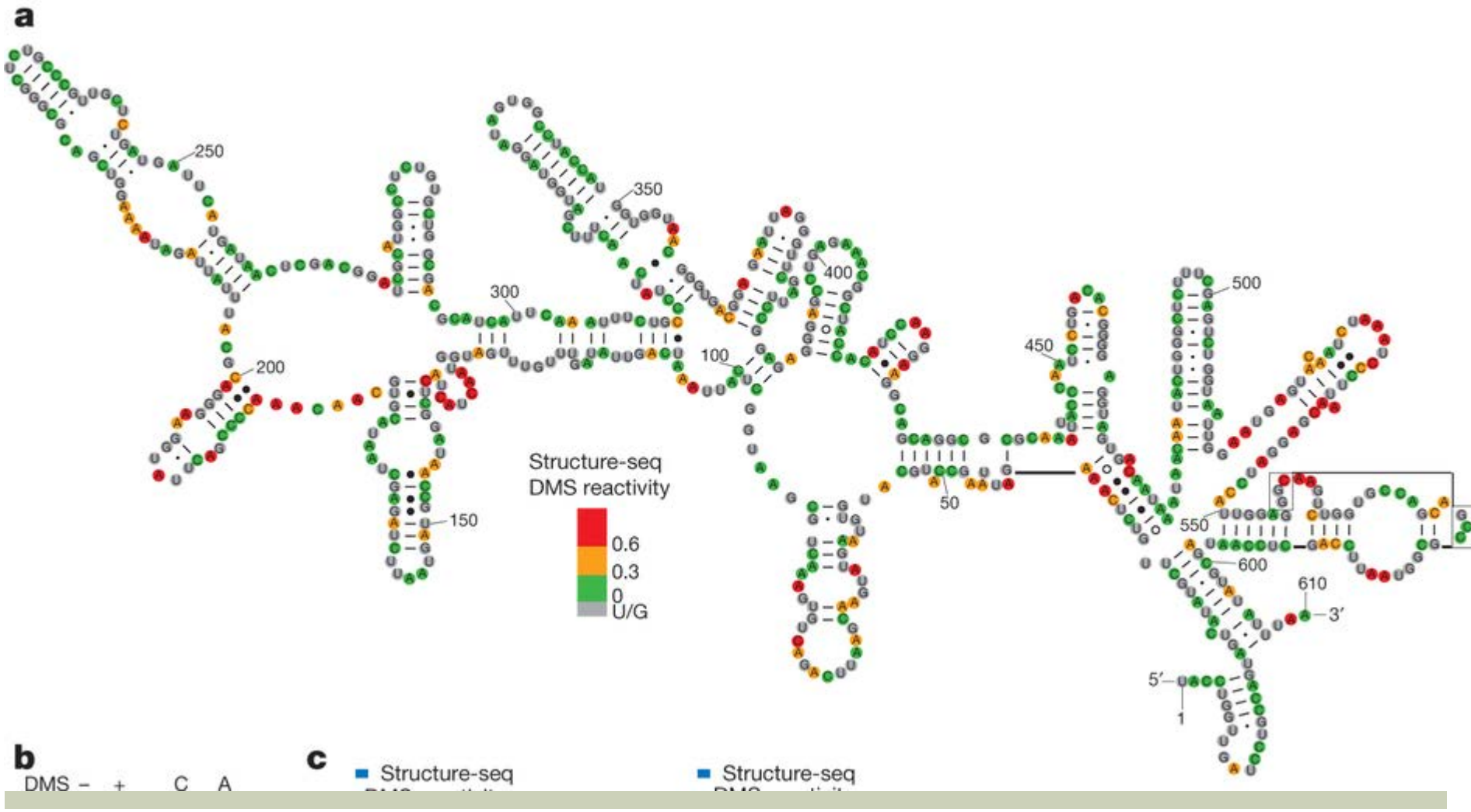
1. For detection of low-abundance RNAs, no depletion of high-abundance RNAs is needed (through LMPCR) – or, conversely, only simple RNA handling and preparation is needed
2. Coupling of DMS/SHAPE probing with LMPCR increases the sensitivity of transcripts by 5 logs
3. DMS/SHAPE-LMPCR reveals differences between structural data gained *in vitro* and in *vivo*

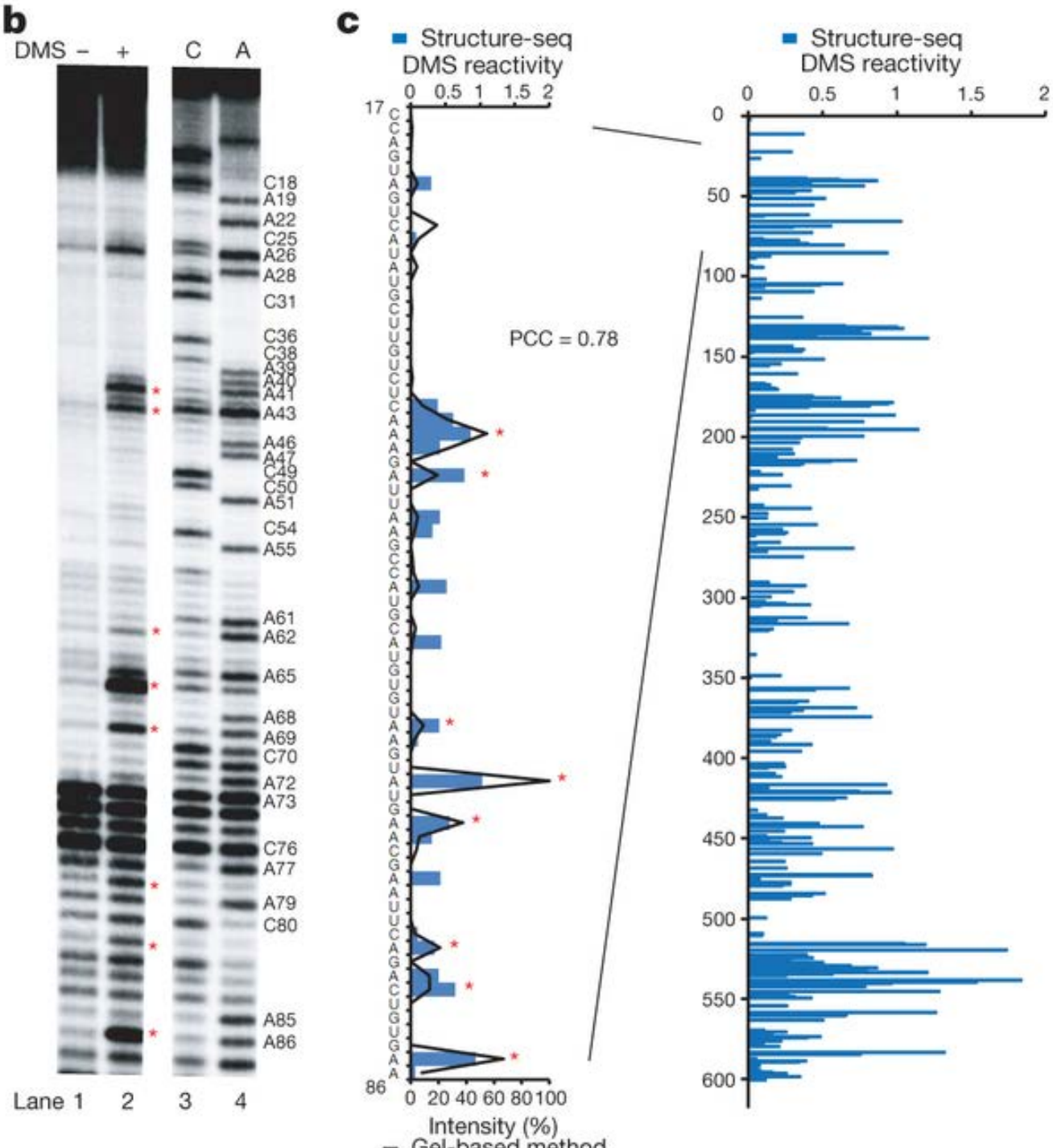
COMBINING NGS WITH IN VIVO RNA STRUCTURE DETERMINATION – STRUCTURE-SEQ

DING Y ET AL., NATURE 2014

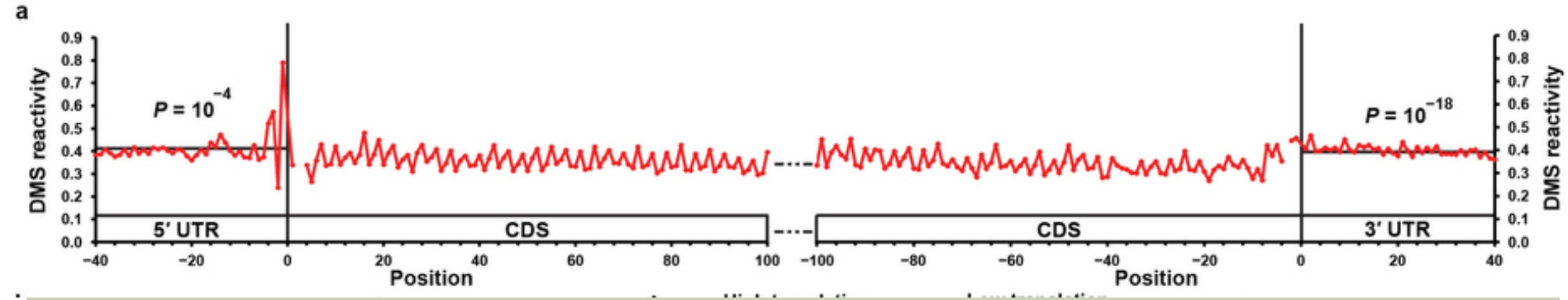


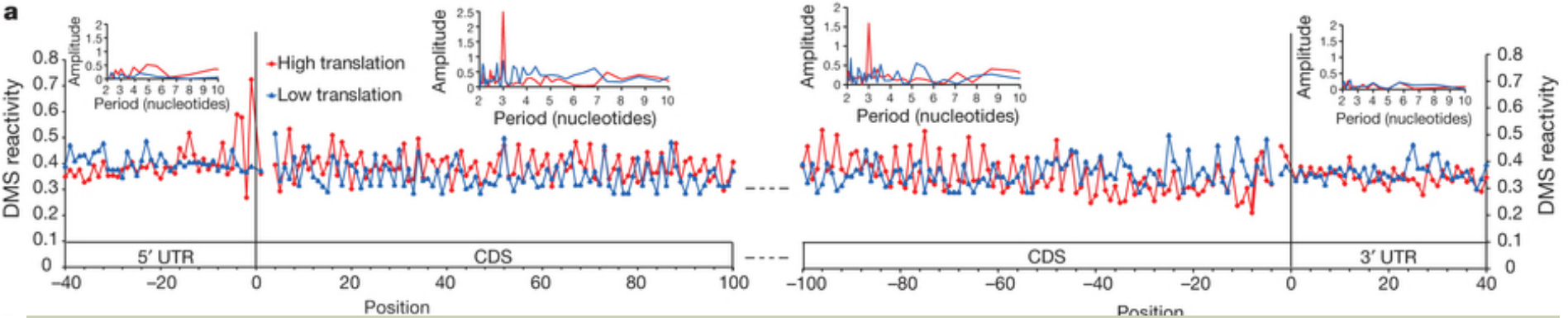
VALIDATION OF STRUCTURE-SEQ



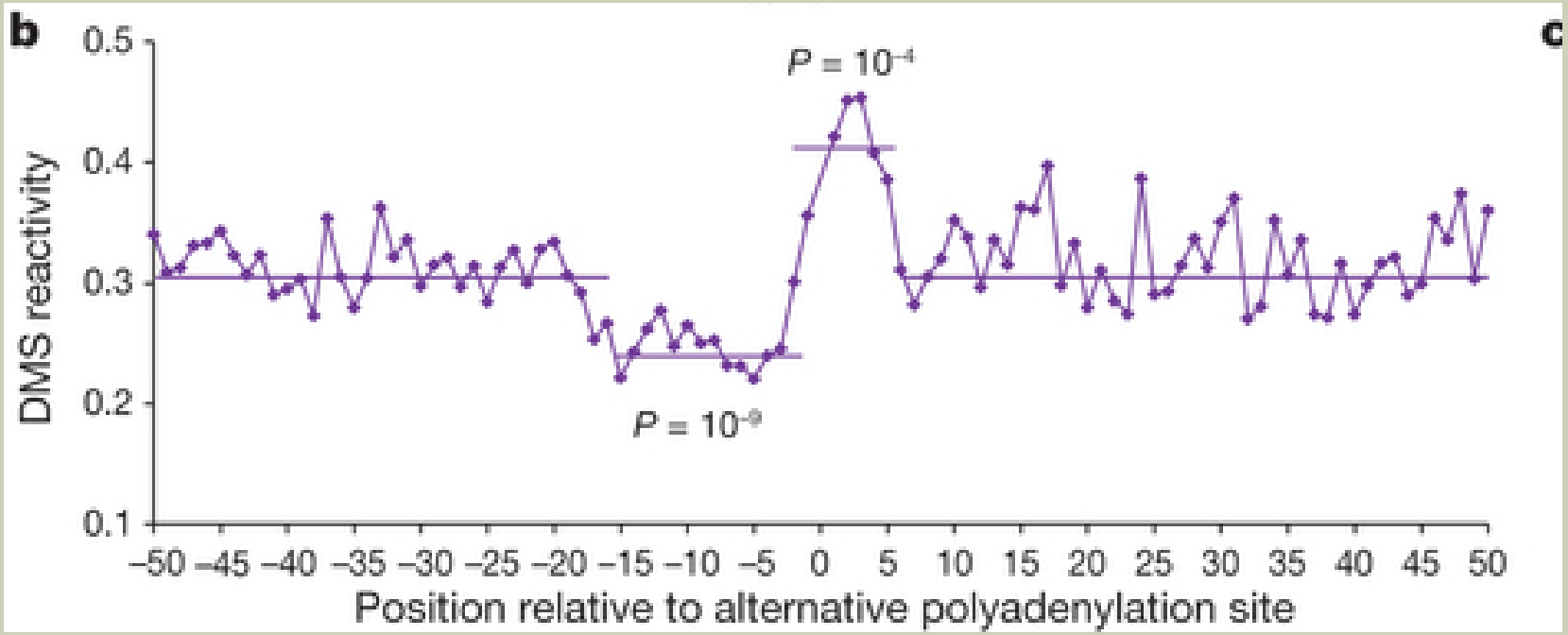

b
DMS −
+
C
A
C18
A19
A22
C25
A26
A28
C31
C36
C38
A39
A40
A41
A43
A46
A47
C49
C50
A51
C54
A55
A61
A62
A65
A68
A69
C70
A72
A73
C76
A77
A79
C80
A85
A86
Lane 1
2
3
4
c
Structure-seq
DMS reactivity
0 0.5 1 1.5 2
17
86
PCC = 0.78
0 20 40 60 80 100
Intensity (%)
Structure-seq
DMS reactivity
0 0.5 1 1.5 2
0
50
100
150
200
250
300
350
400
450
500
550
600

UNSTRUCTURED REGION UPSTREAM OF START CODON HAS ENRICHED DMS-REACTIVITY IN HIGH TRANSLATION EFFICIACY TRANSCRIPS

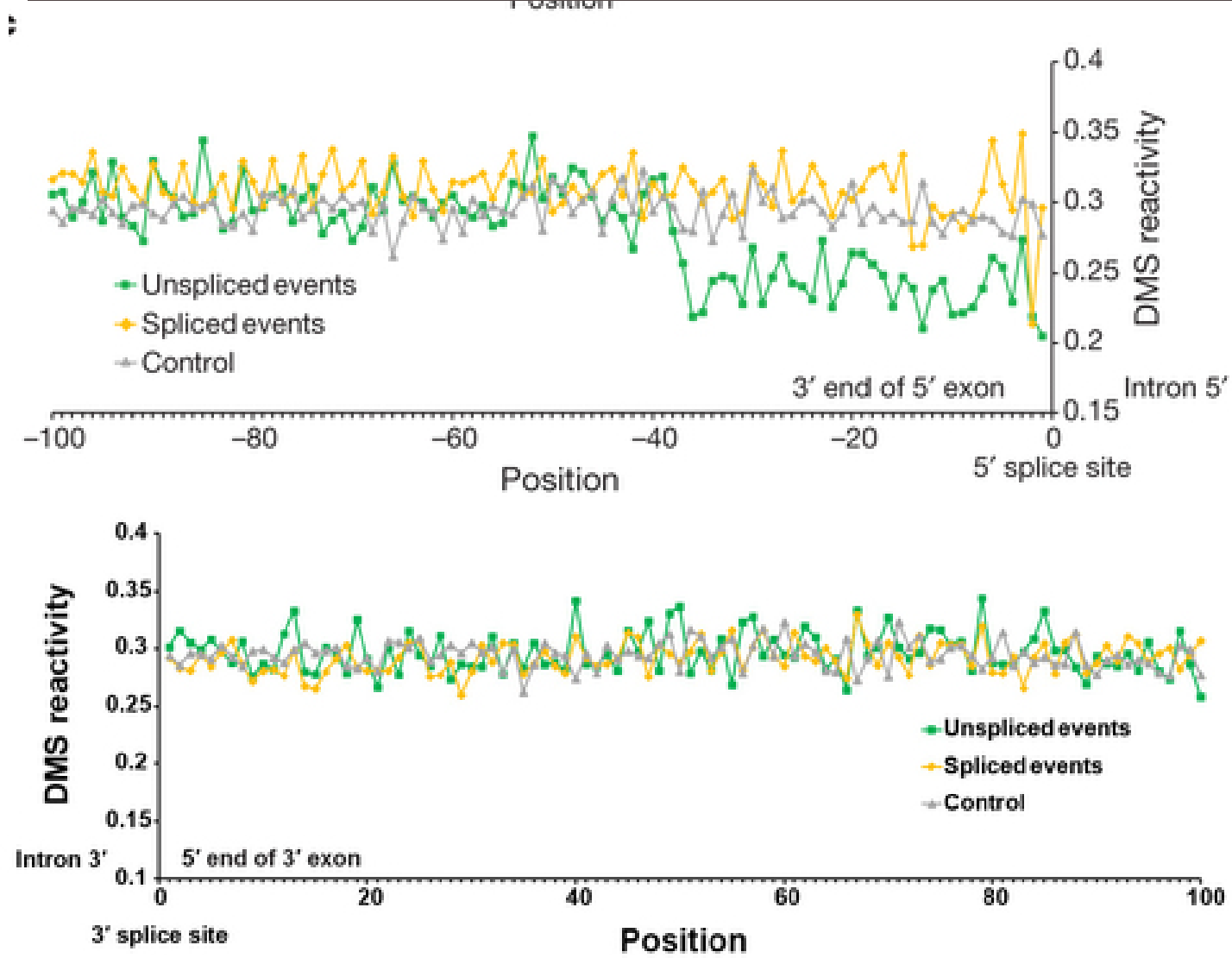




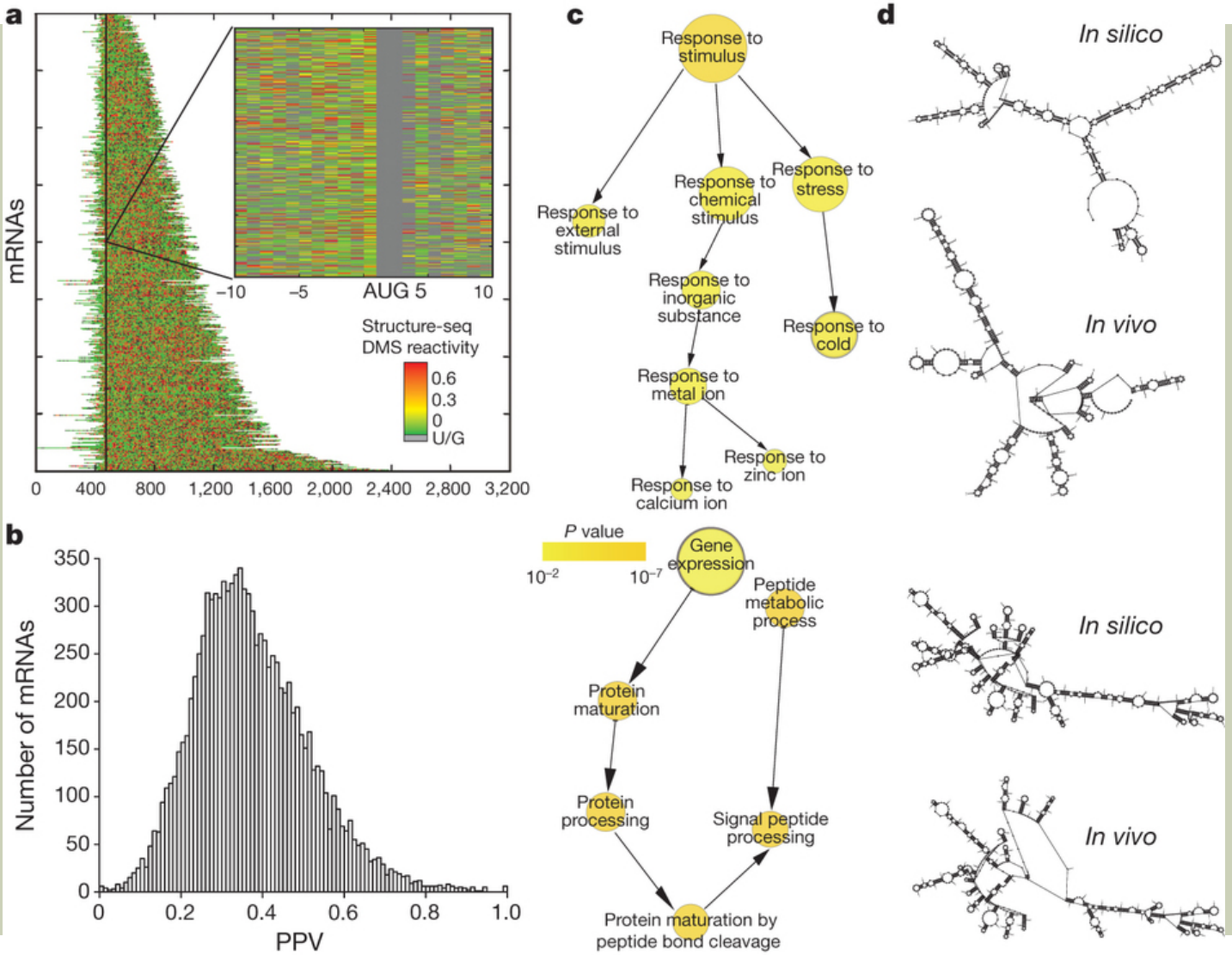
DIFFERENTIAL DMS REACTIVITY AROUND CLEAVAGE SITES OF REGIONS WITH ALTERNATIVE POLYADENYLATION



LOWER DMS REACTIVITY AT ~40 BP UPSTREAM OF 5' ALTERNATIVE SPLICE SITES SHOWS 2ND STRUCTURE REGULATORY MECHANISM



POOR 2ND STRUCTURE CORRELATION BETWEEN IN VIVO AND IN VITRO/IN SILICO DATASETS



CONCLUSION

- Combining RNA secondary structure probing with NGS yields new insights into RNA structure characteristics at high resolution and accuracy