CBGP
UPM-INIA

Bioinformatics approaches to understand RNA biology in the rice blast fungus

Marco Marconi

Center for Plant Biotechnology and Genomics
Technical University of Madrid

POLITÉCNICA
"Ingeniamos el futuro"

CAMPUS
DE EXCELENCIA
INTERNACIONAL

Magnaporthe oryzae is an ascomycete fungus that infects rice

panicle

node

leaf

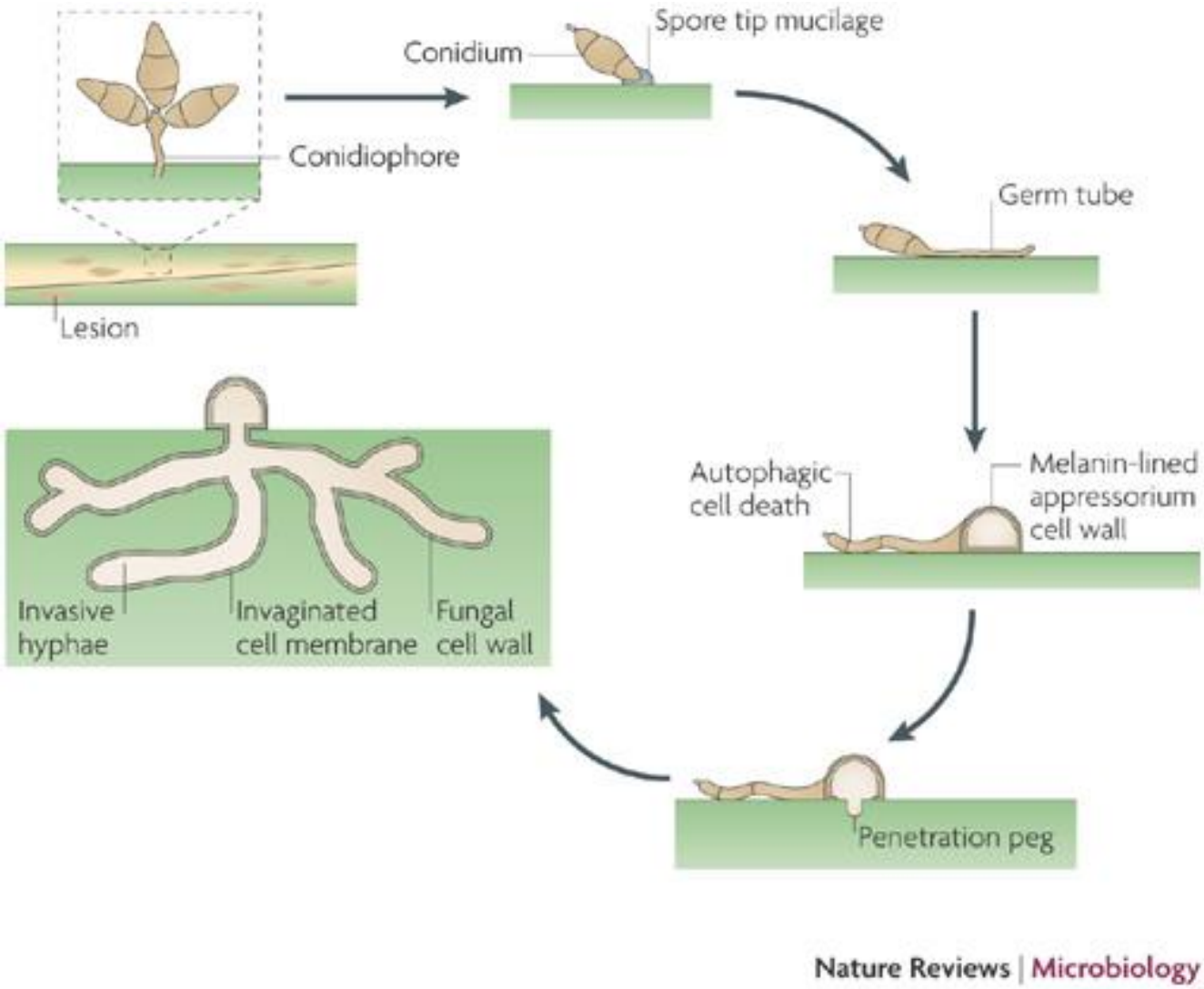
root

Sesma&Osbourn
Nature (2004)

neck

collar

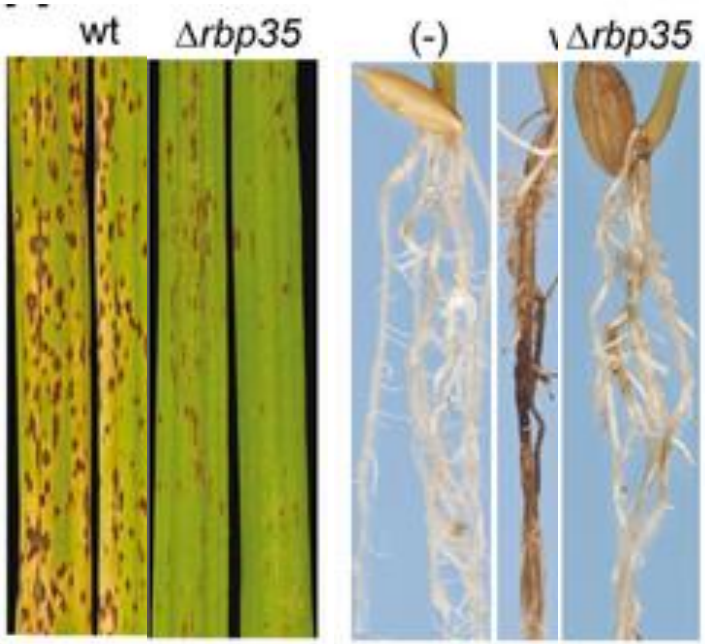
Magnaporthe oryzae life cycle



Richard A. Wilson & Nicholas J. TalbotNature Reviews Microbiology 7, 185-195 (March 2009)

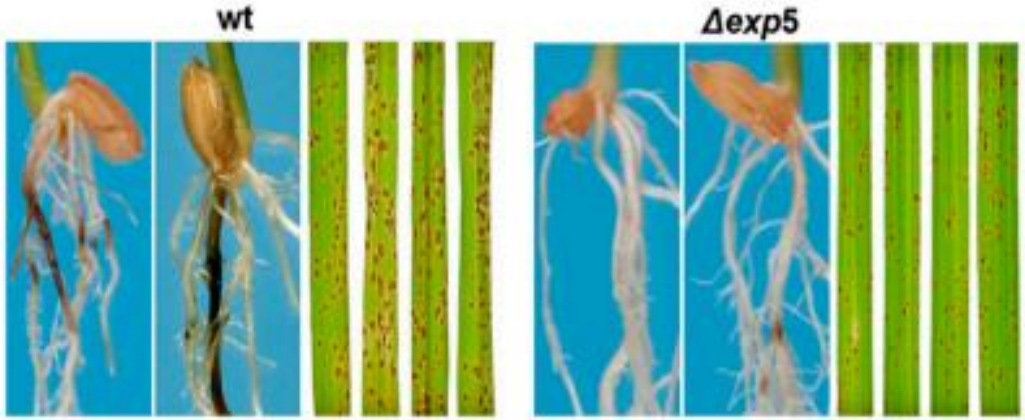
We are interested in two pathogenic-defective mutants…

Δrbp35



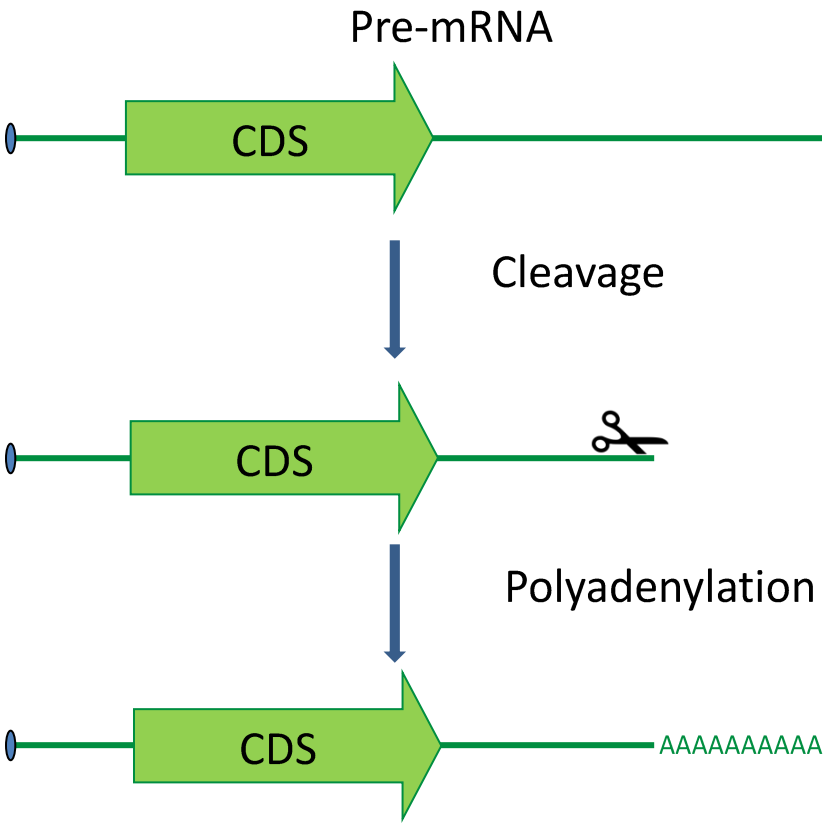
Marina Franceschetti , Emilio Bueno , Richard A. Wilson, Sara L. Tucker, Concepción Gómez-Mena, Grant Calder, Ane Sesma 2011

Δexp5

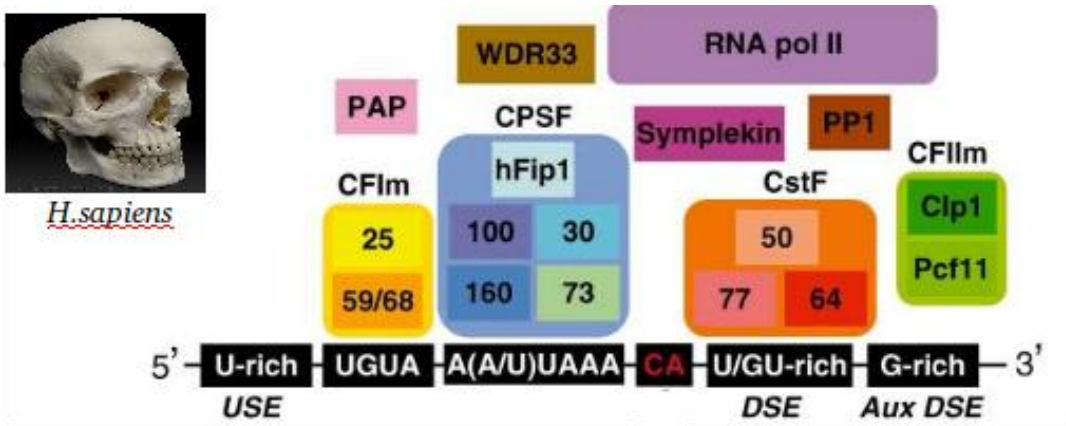


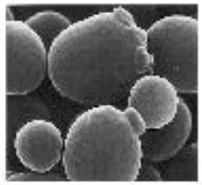
Tucker SL1, Besi MI, Galhano R, Franceschetti M, Goetz S, Lenhert S, Osbourn A, Sesma A. 2010

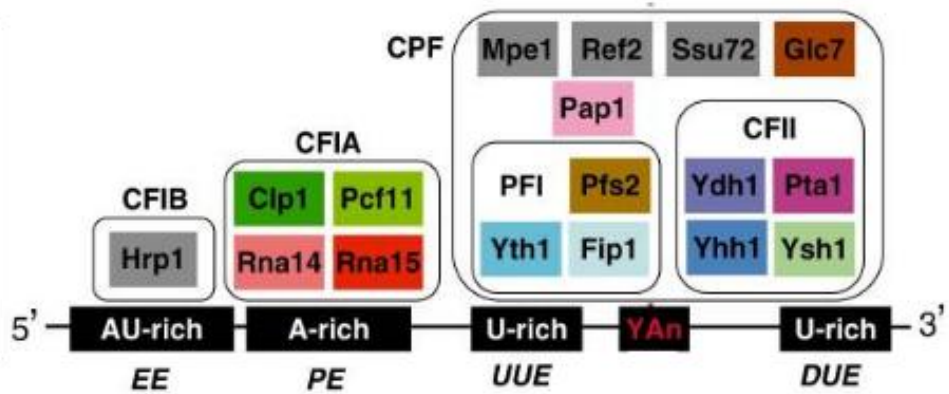
Rbp35 is a member of the polyadenylation machinery

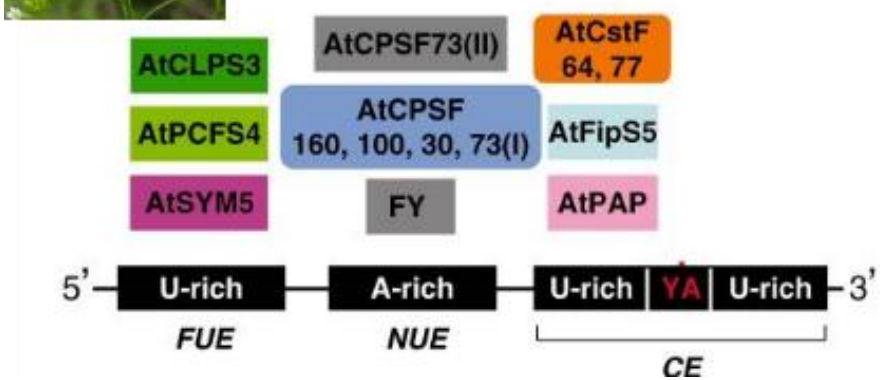


Rbp35 is a member of the polyadenylation machinery





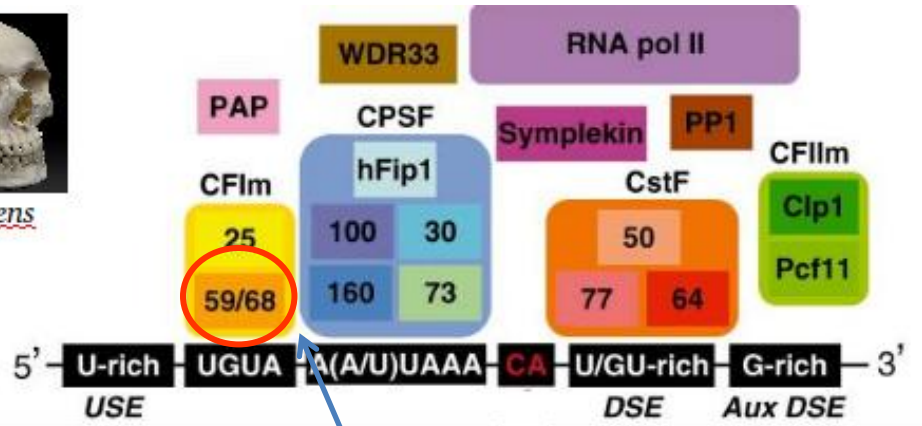




S Millevoi et al. 2010

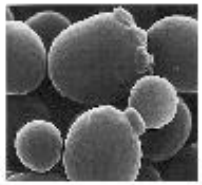
Rbp35 is a member of the polyadenylation machinery

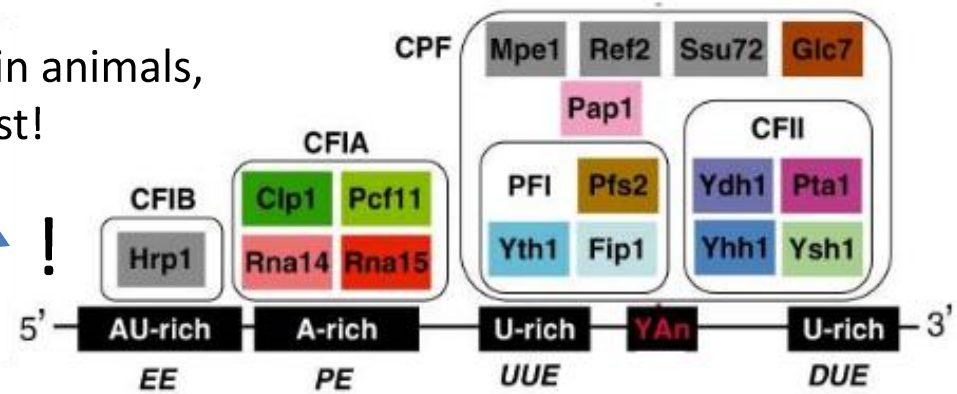
H.sapiens



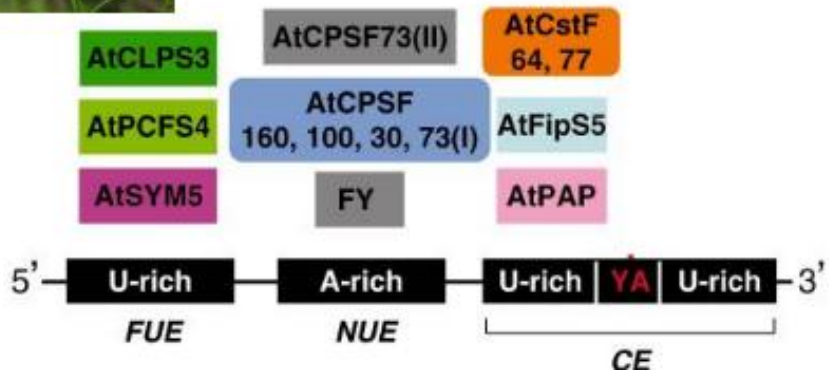
RBP35 ortholog found in animals, but not in yeast!

Yeasts



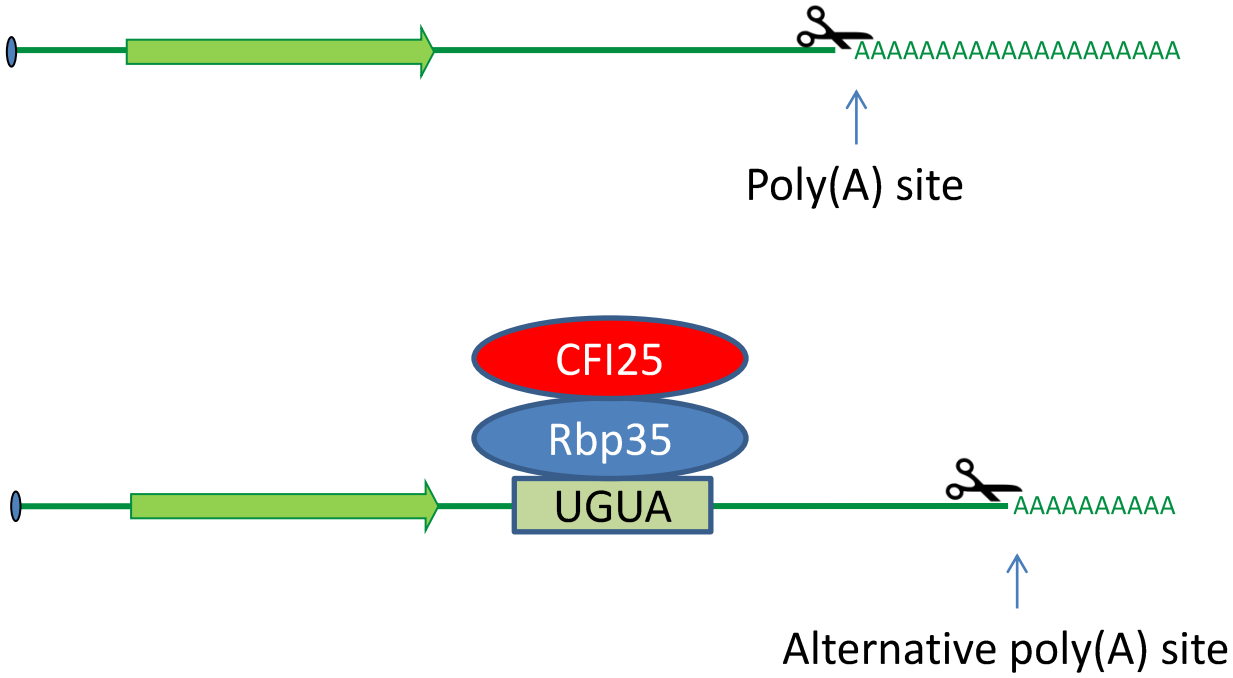


Plants

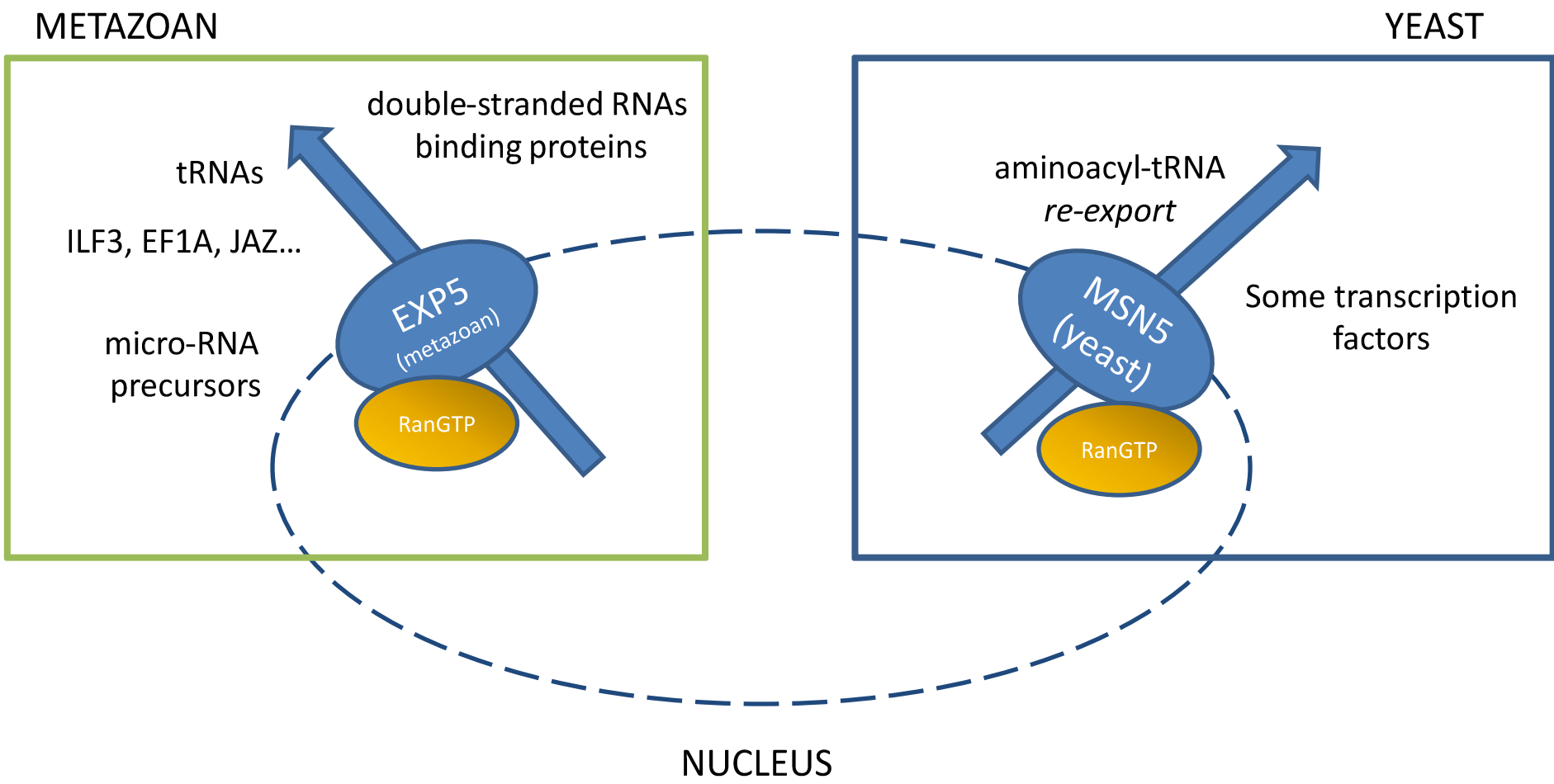


S Millevoi et al. 2010

Human Rbp35 orthologue binds to UGUA element to regulate the selection of alternative polyadenylation sites (APA)

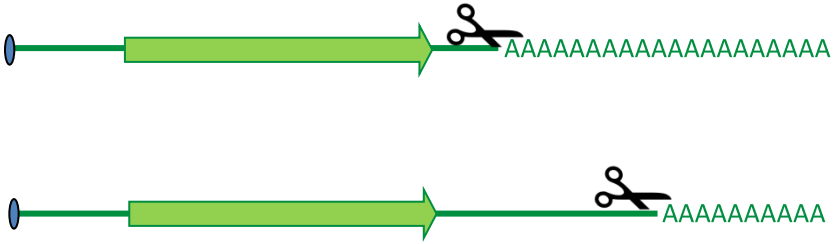


EXP5 is a kariopherin involved in the nuclear import/export of proteins and RNAs



We want to shed light on the role of RBP35 and EXP5 in *M.oryzae:*

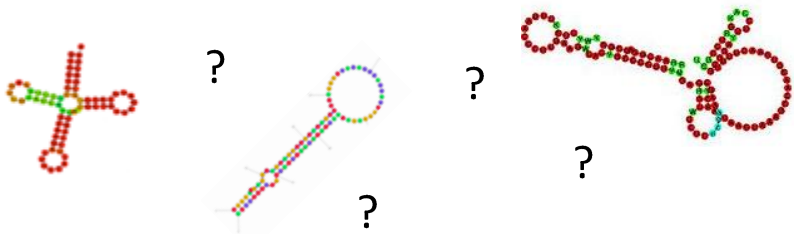
- RNA-seq on polyadenylation sites



We want to shed light on the role of RBP35 and EXP5 in *M.oryzae:*

- RNA-seq on polyadenylation sites

- Small RNA-sequencing



We want to shed light on the role of RBP35 and EXP5 in *M.oryzae:*

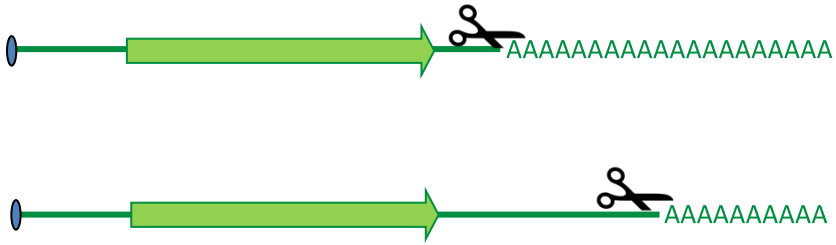
- RNA-seq on polyadenylation sites

- Small RNA-seq

- Evolutionary analysis of polyadenylation/splicing/translation machineries

We want to shed light on the role of RBP35 and EXP5 in *M.oryzae:*

- RNA-seq on polyadenylation sites



Sequencing details

Julio Rodríguez-Romero

Two strains:
WT
Δrbp35

Four nutrient conditions:
Complete medium
Minimal medium
Nitrogen starvation
Carbon starvation

Typical on-plant conditions

3 Replicates:
#1, #2, #3

Coverage:
400x poly(A) site

(Wilkening et al., Nucleic Acids Res. 2013 , 1–8. An efficient method for genome-wide polyadenylation site mapping and RNA quantification)

Poly(A) site mapping analysis

Results in wild-type

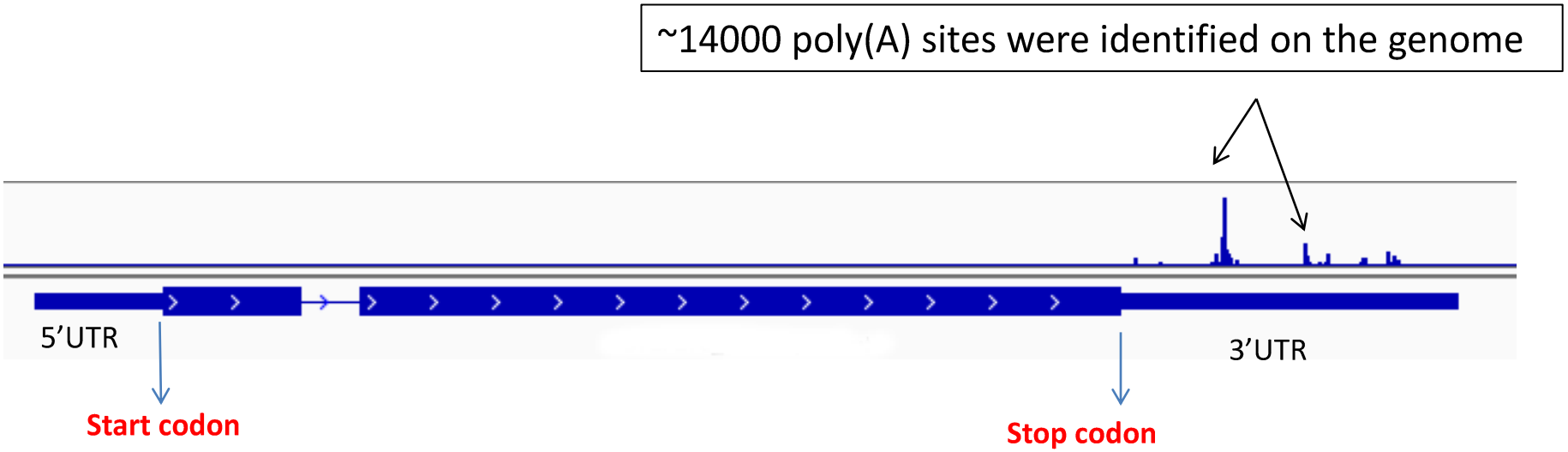
Results in the Δ*rbp35* mutant

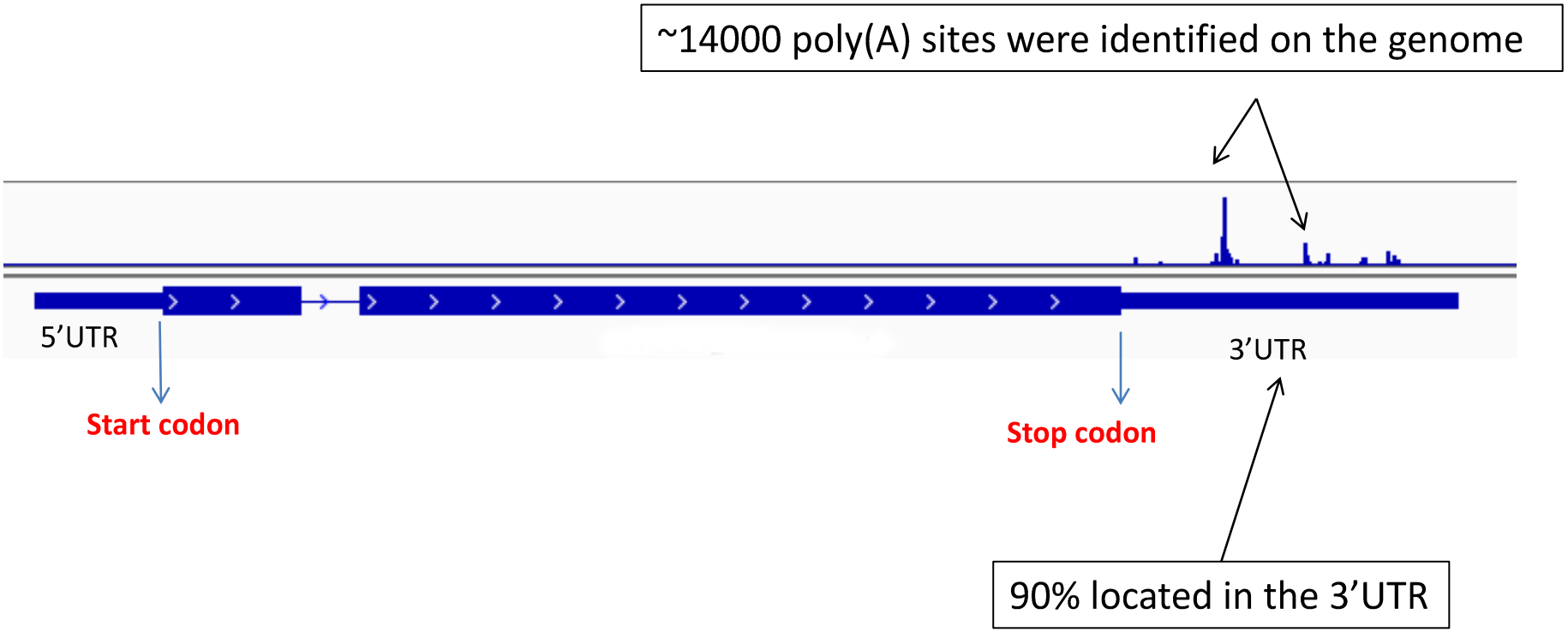
Poly(A) site mapping analysis

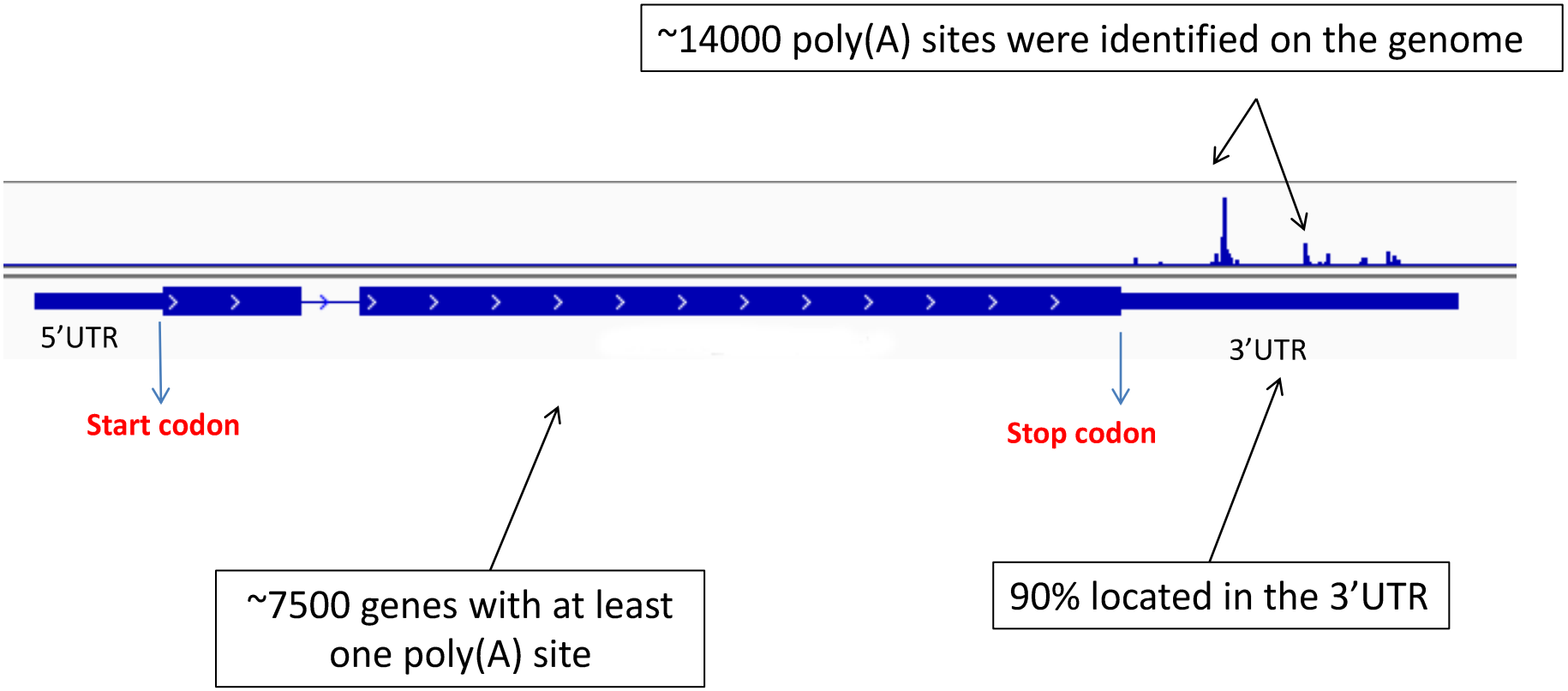
Results in wild-type

Results in the Δ*rbp35* mutant

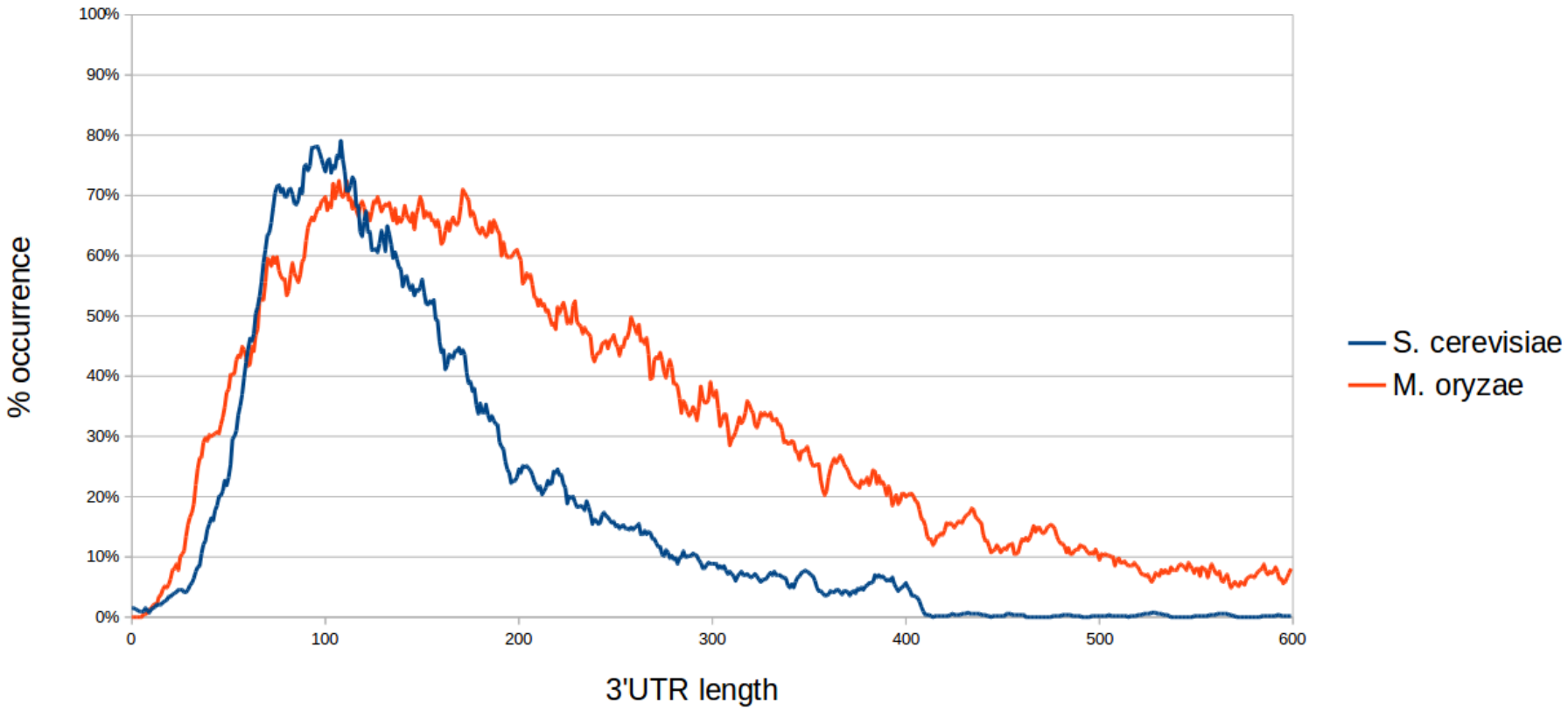
5'UTR
3'UTR
Start codon
Stop codon


~14000 poly(A) sites were identified on the genome
5'UTR
3'UTR
Start codon
Stop codon


~14000 poly(A) sites were identified on the genome
5'UTR
Start codon
Stop codon
3'UTR
90% located in the 3'UTR

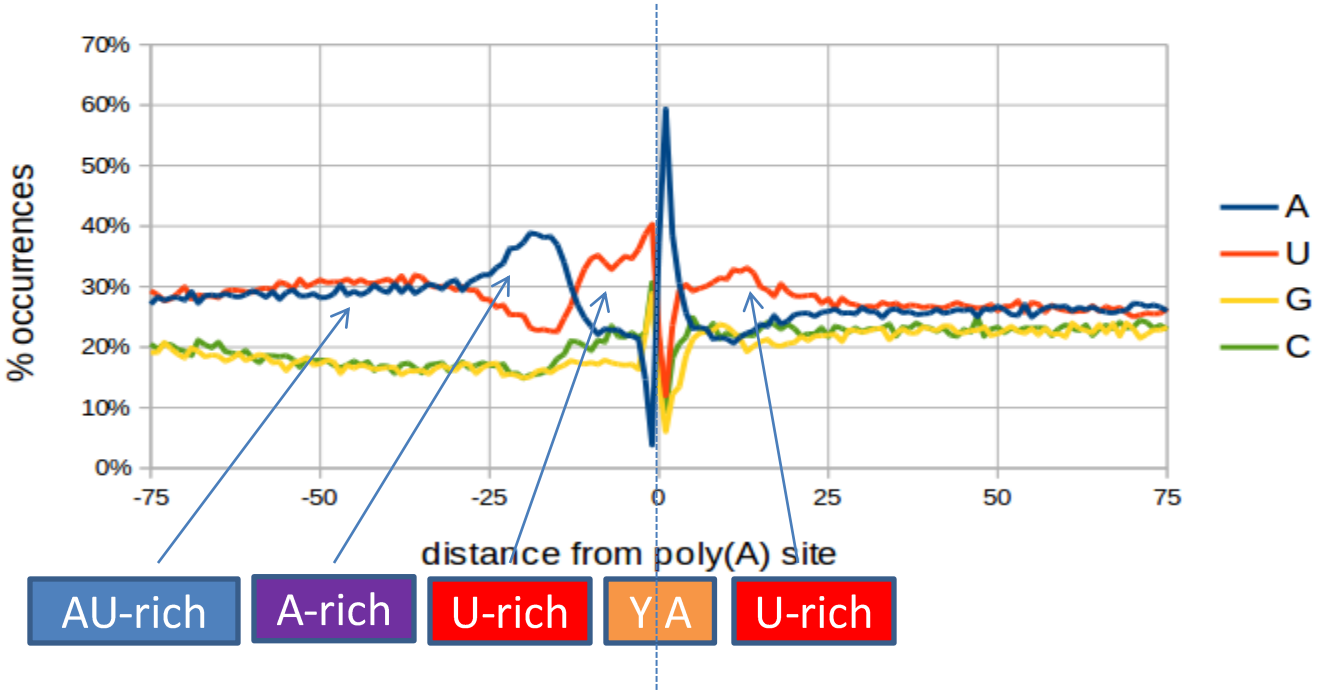

~14000 poly(A) sites were identified on the genome
5'UTR
3'UTR
Start codon
Stop codon
~7500 genes with at least one poly(A) site
90% located in the 3'UTR

Magnaporthe oryzae 3’ UTR lengths are usually longer than Yeast



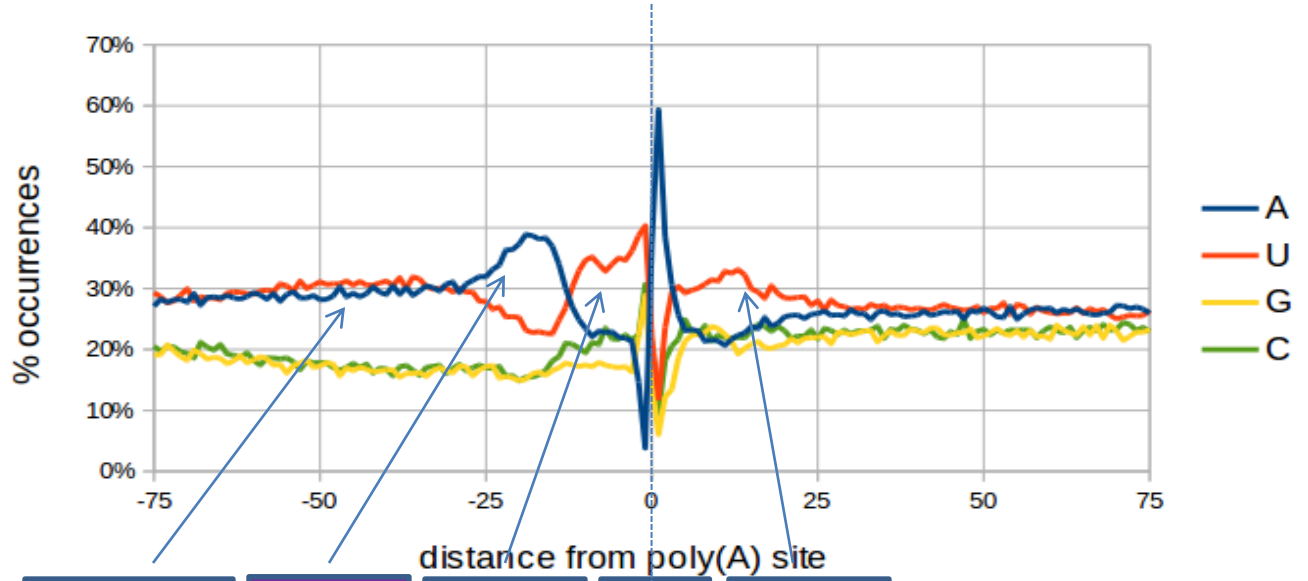
Nucleotide Profile of poly(A) site differs slightly from yeast

YEAST



Nucleotide Profile of poly(A) site differs slightly from yeast

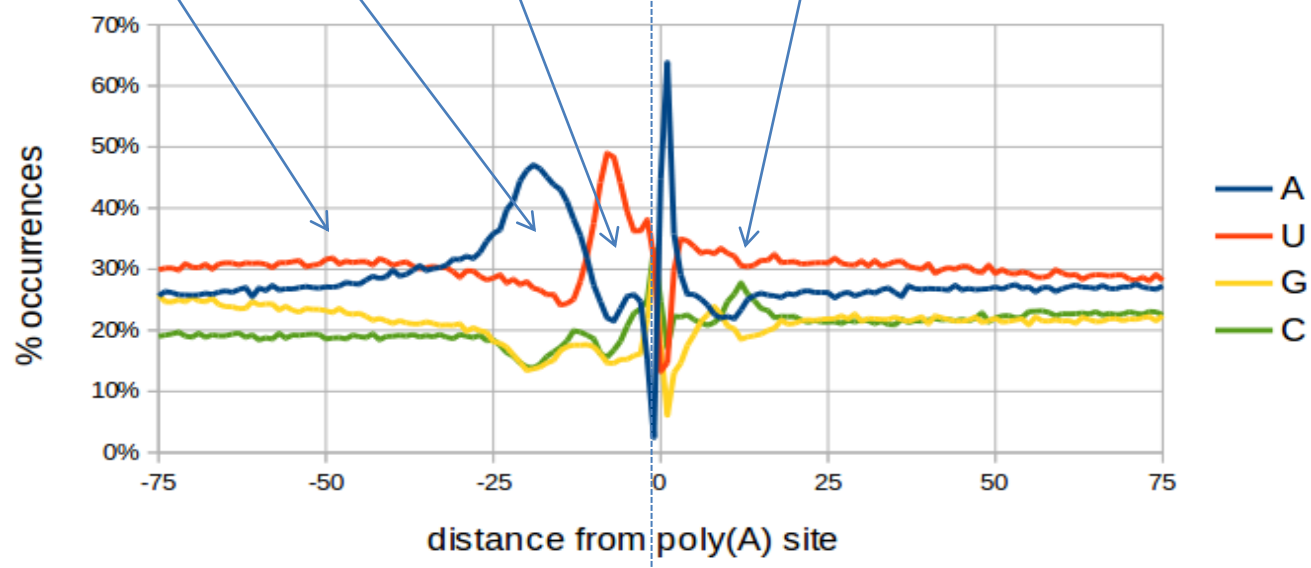
YEAST



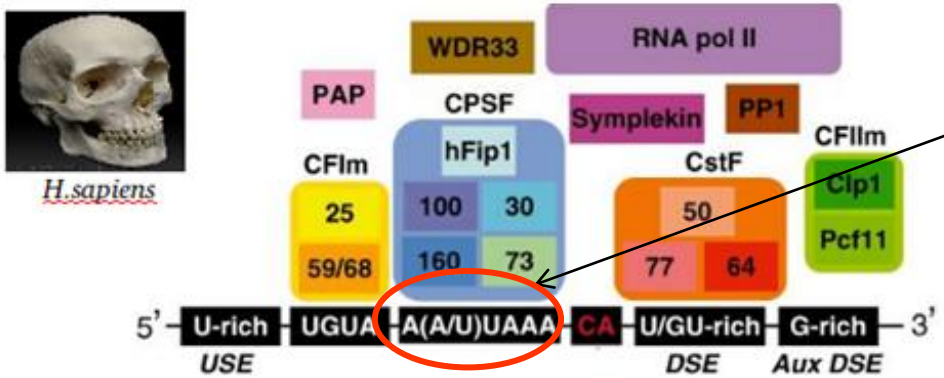
AU-rich | A-rich | U-rich | Y A | U-rich

UGUA | UAGA | A-rich | U-rich | S A | C peak

M.oryzae

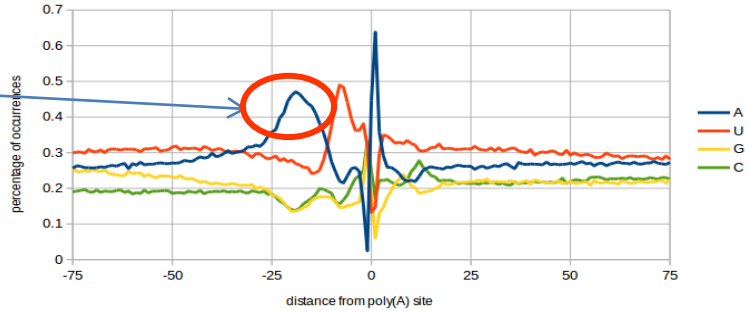


Contrary to metazoan, fungal canonical AAUAAA polyadenylation signal is highly degenerated

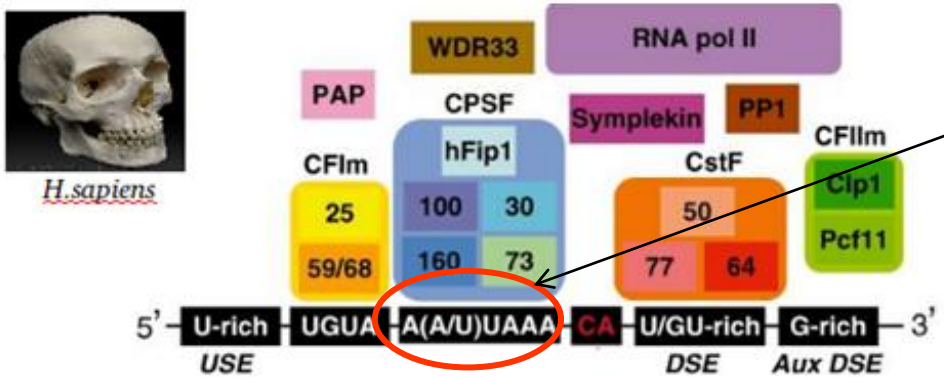


Very uniform in animals

A-rich

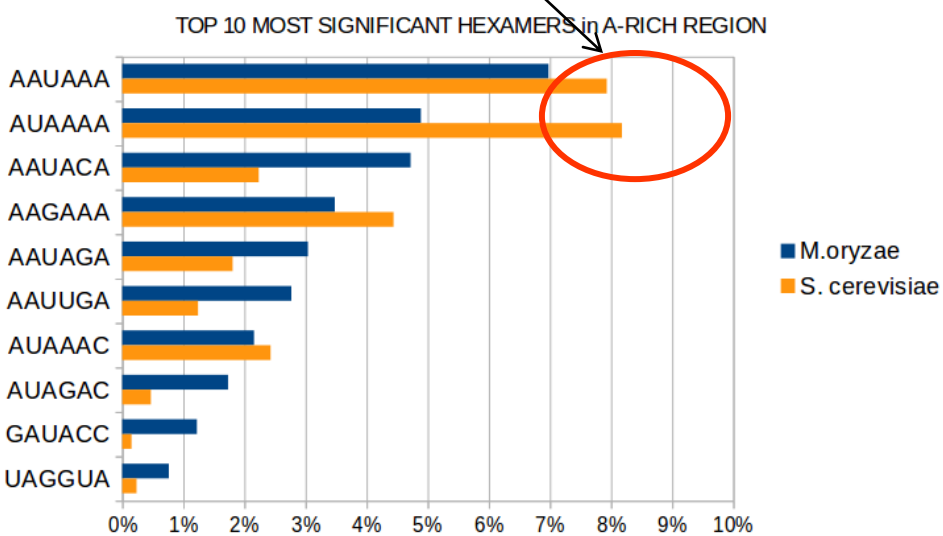


Contrary to metazoan, fungal canonical AAUAAA polyadenylation signal is highly degenerated

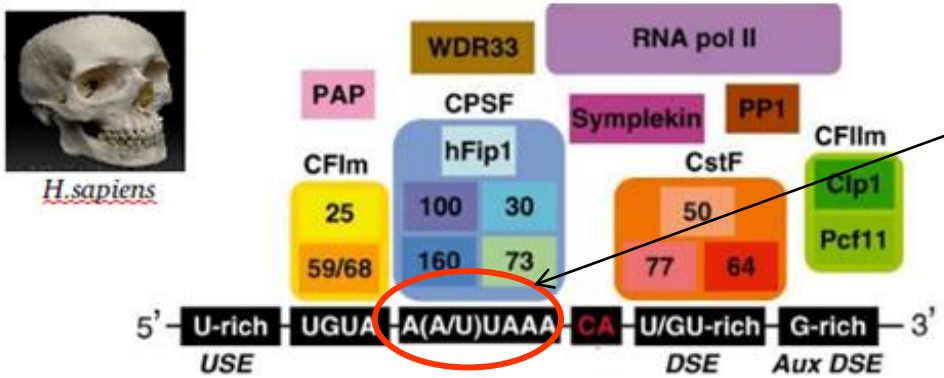


Very uniform in animals

Only 7% in fungi!



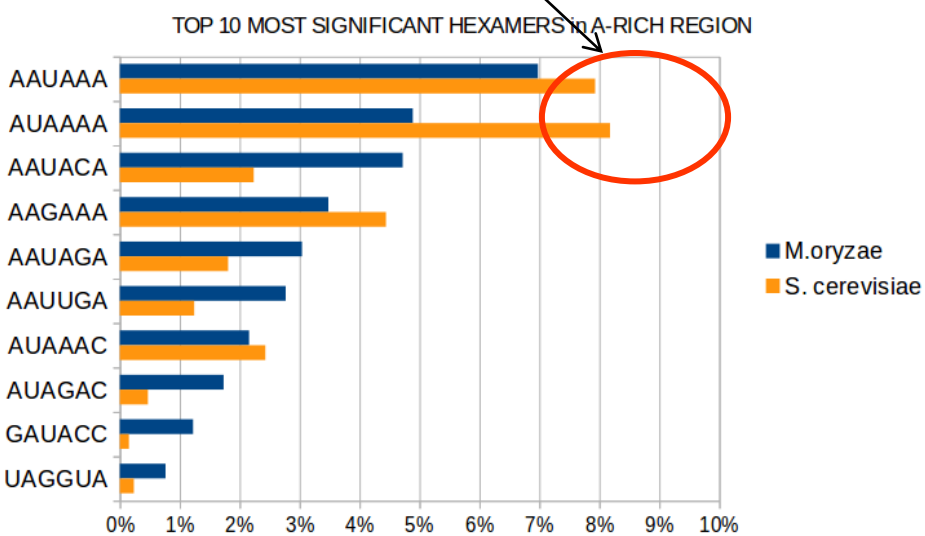
Contrary to metazoan, fungal canonical AAUAAA polyadenylation signal is highly degenerated

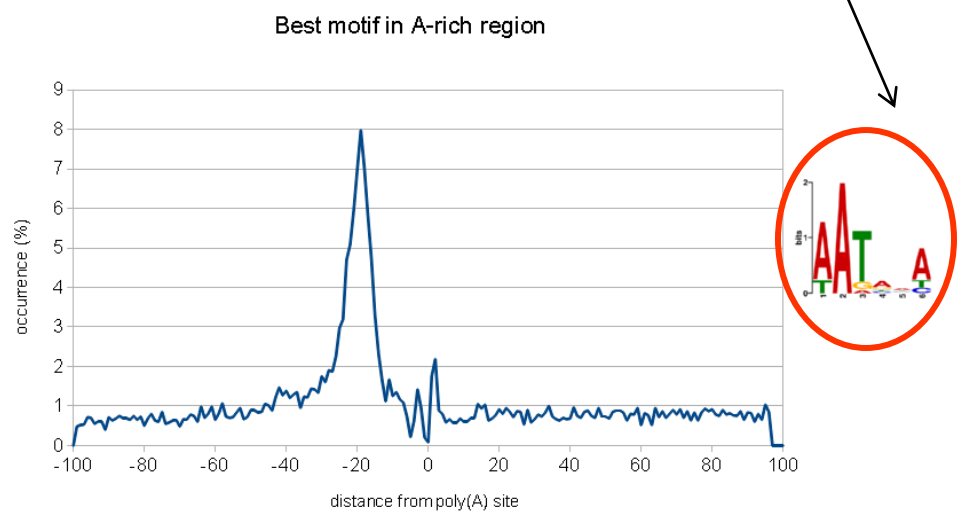


Very uniform in animals

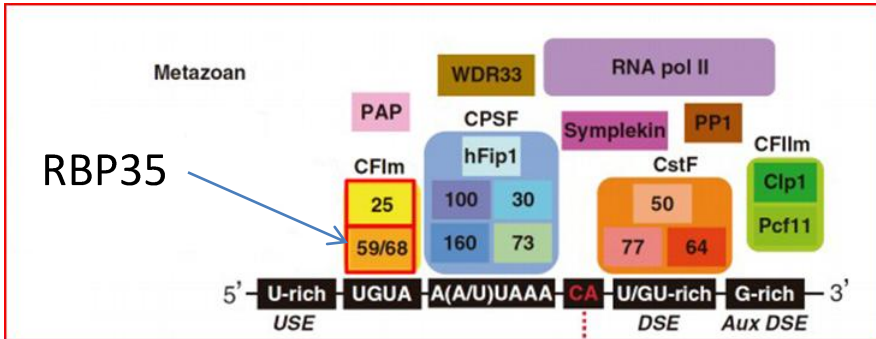
Only 7% in fungi!

Very degenerated!

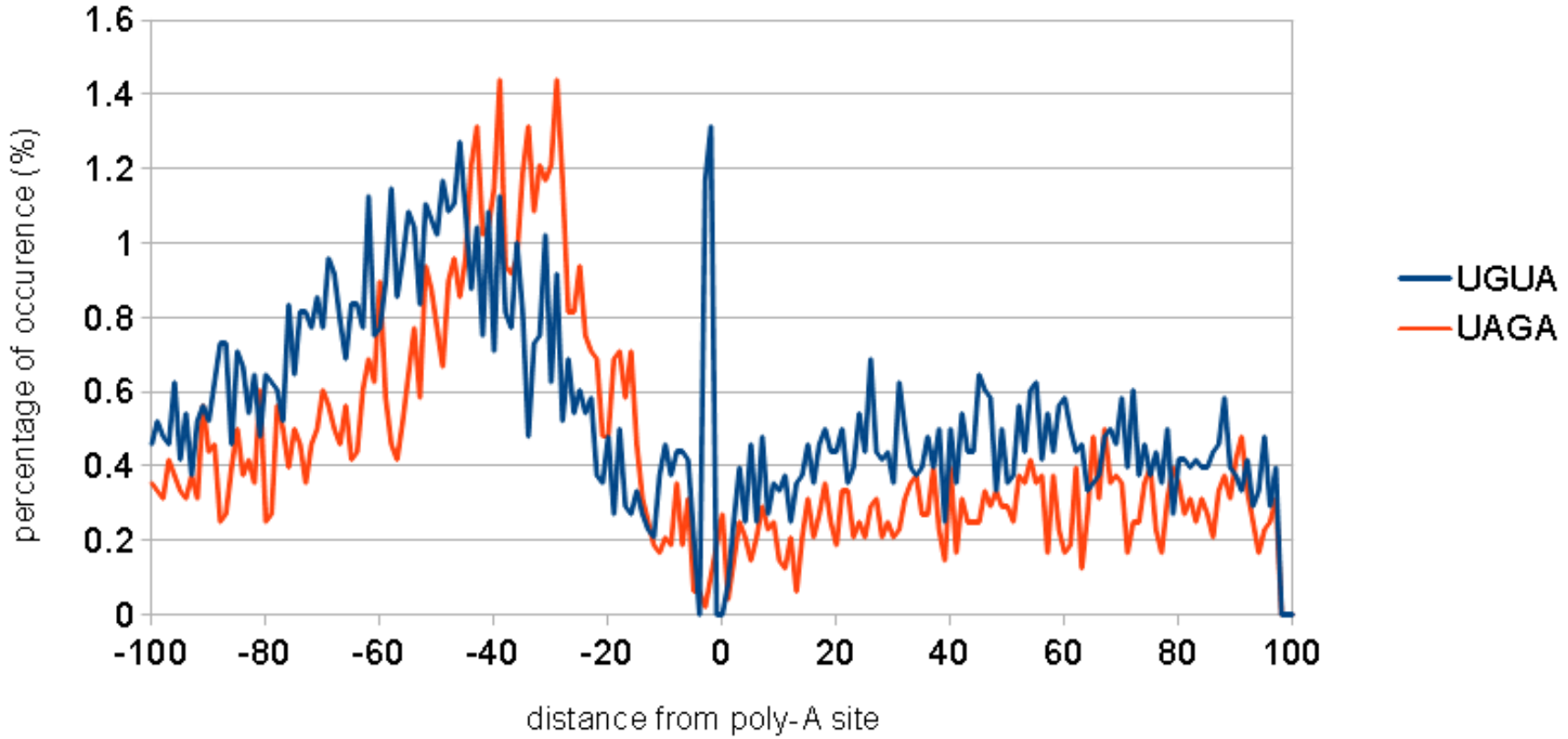


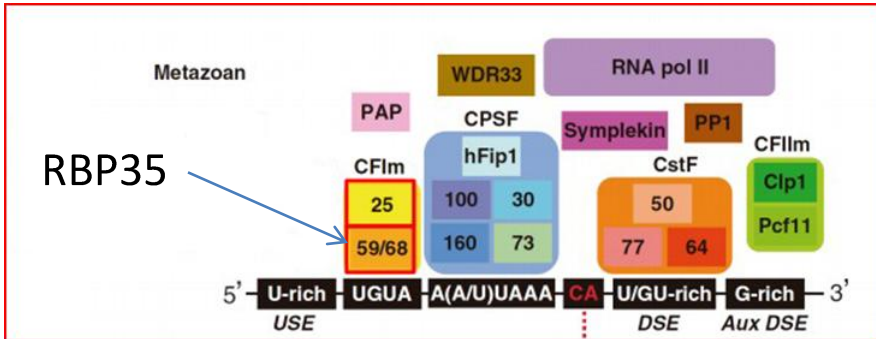


As in metazoan, *M. oryzae* has a UGUA motif, together with a new UAGA motif



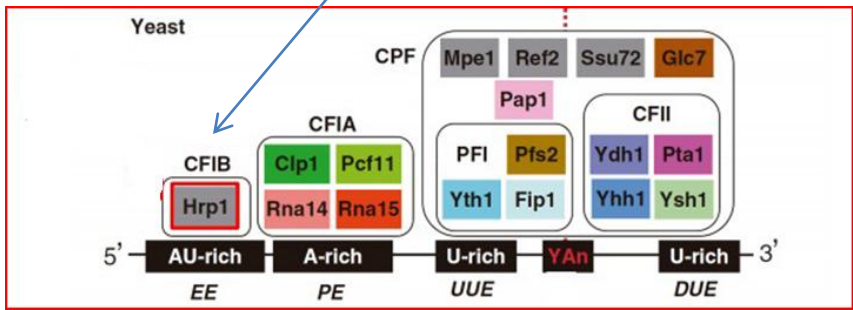
As in metazoan, *M. oryzae* has a UGUA motif, together with a new UAGA motif



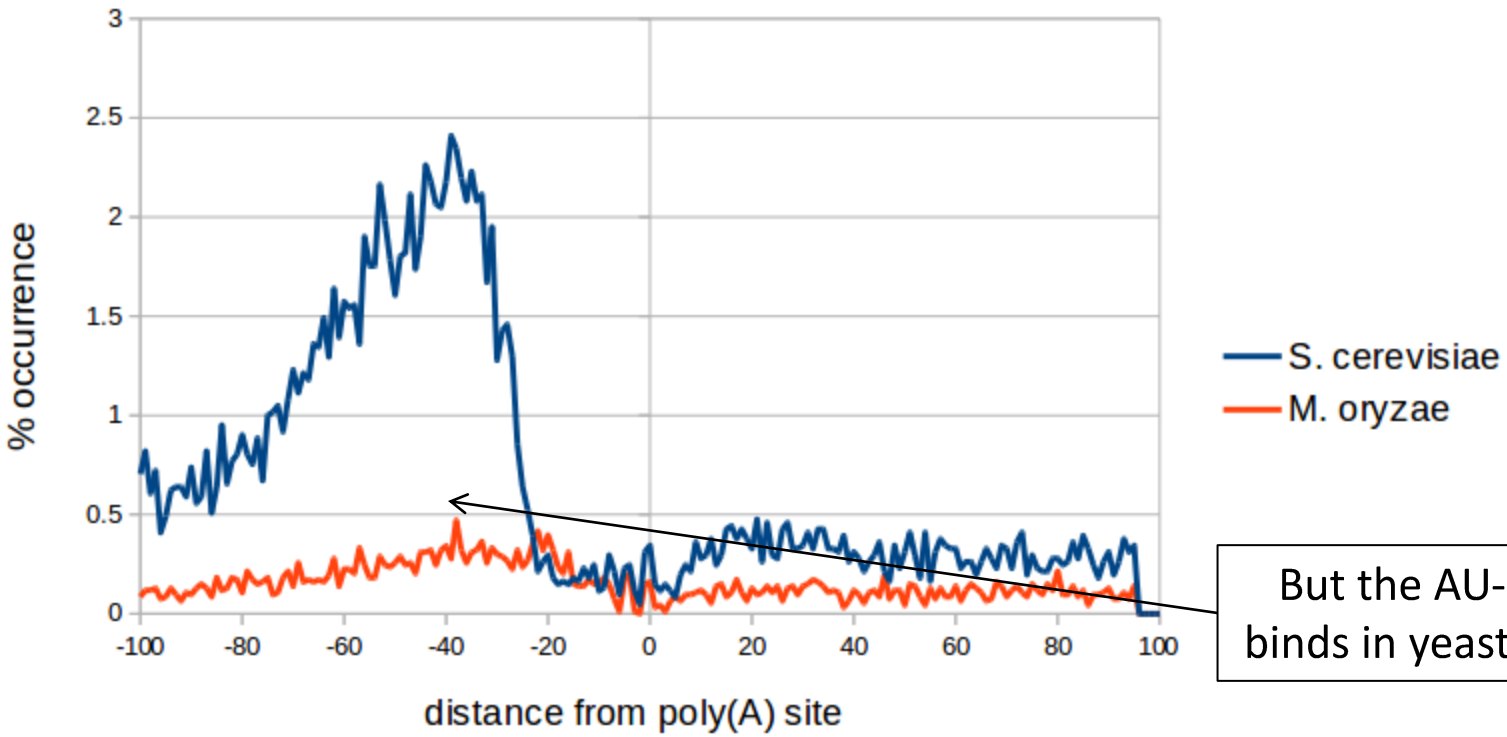


The motif bound by Hrp1 in yeast is missing is *M. oryzae*

M. oryzae **has** this protein, Hrp1

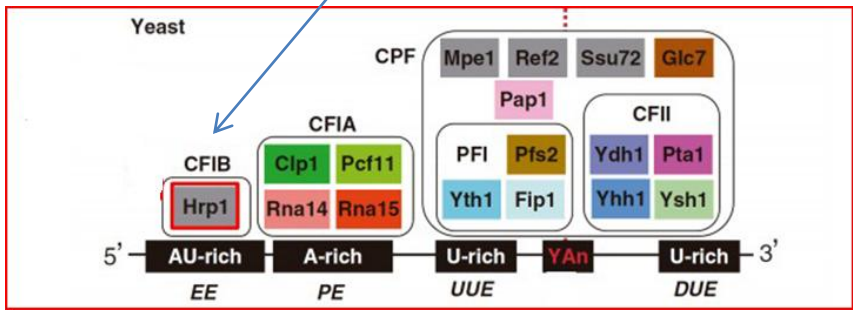


The motif bound by Hrp1 in yeast is missing is *M. oryzae*

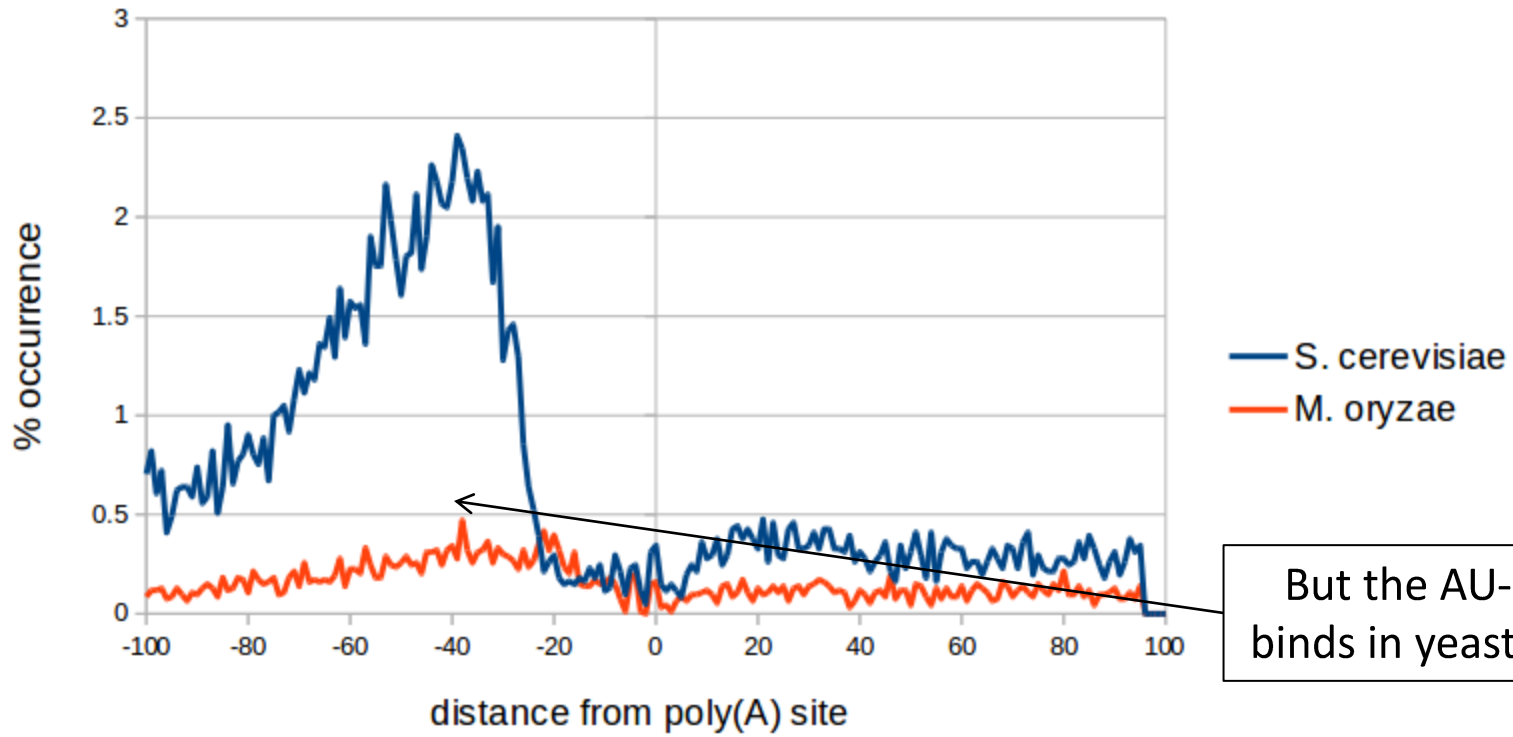


But the AU-rich element it binds in yeast is missing here!

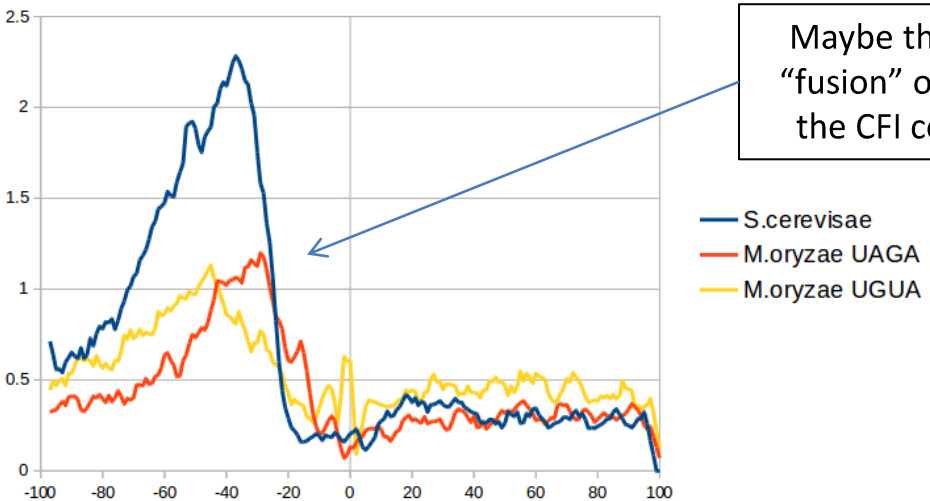
M. oryzae **has** this protein, Hrp1



The motif bound by Hrp1 in yeast is missing is *M. oryzae*

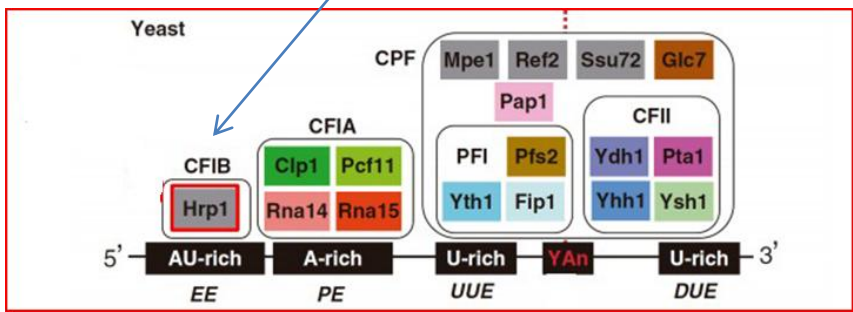


But the AU-rich element it binds in yeast is missing here!

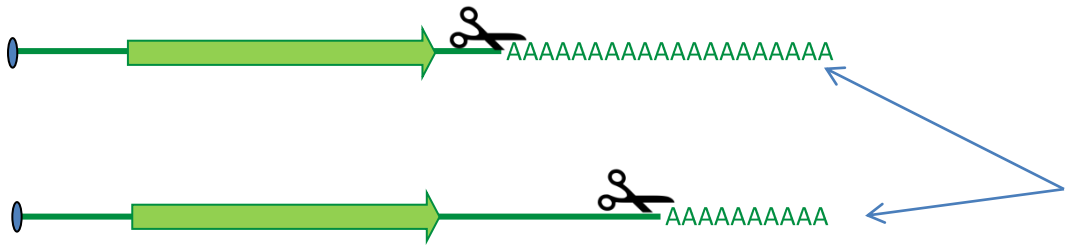


Maybe the AU-rich element is the "fusion" of UGUA and UAGA, when the CFI complex was lost in yeast

M. oryzae **has** this protein, Hrp1

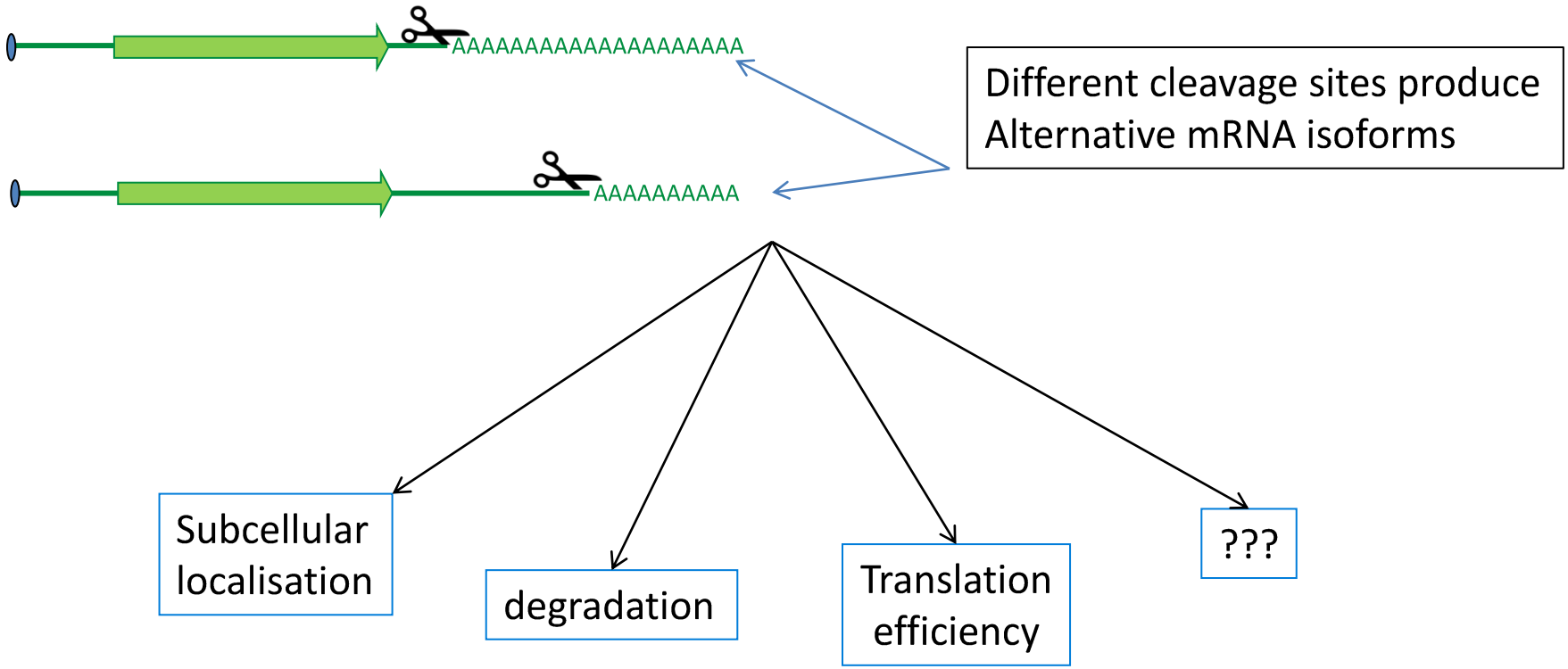


What about alternative polyadenylation?

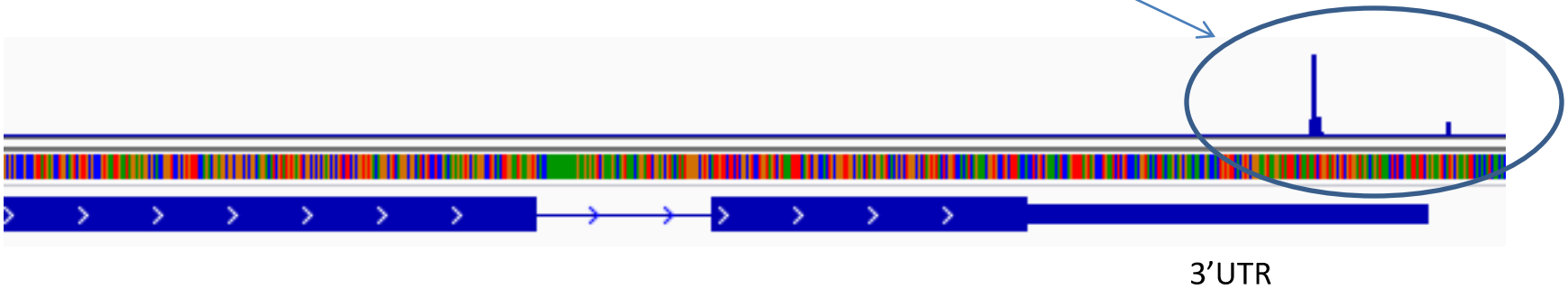


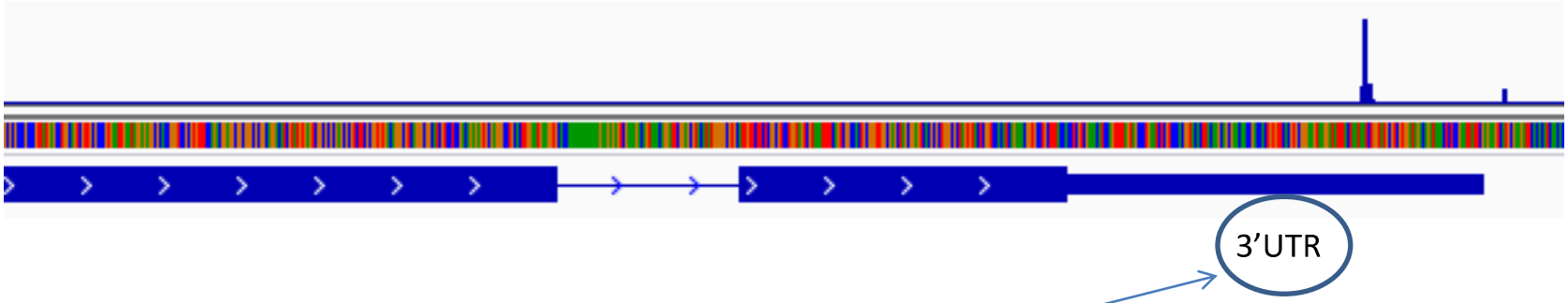
Different cleavage sites produce Alternative mRNA isoforms

What about alternative polyadenylation?

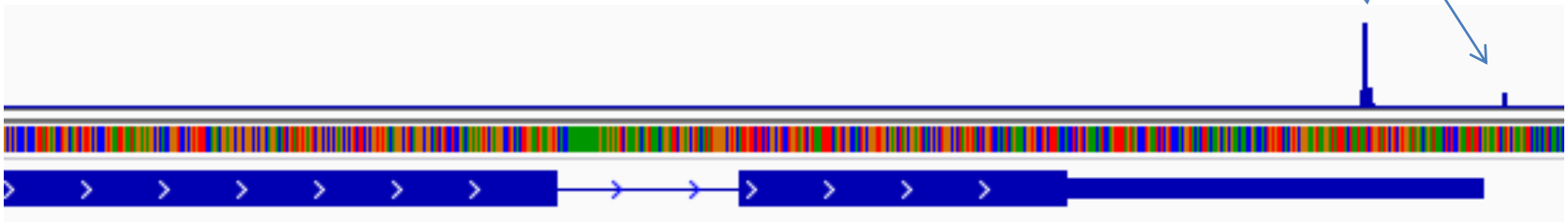


In *M. oryzae* ~30% of genes are alternatively polyadenylated




3’UTR
90% of APA events
occur in the 3’UTR

80% of APA events are made of 2 alternatives only

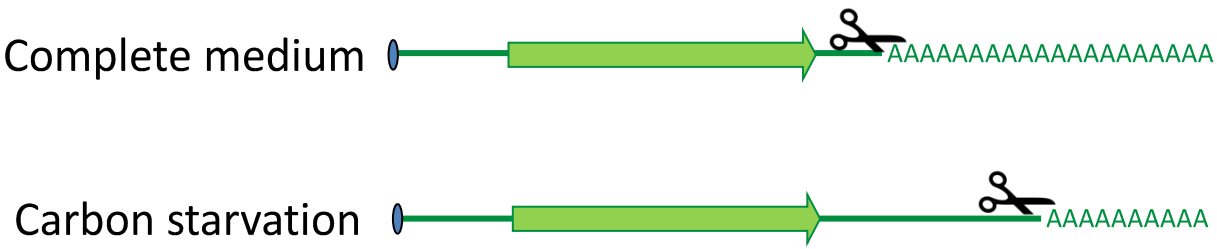


3’UTR

Alternative polyadenylation seems to be involved in specific biological processes

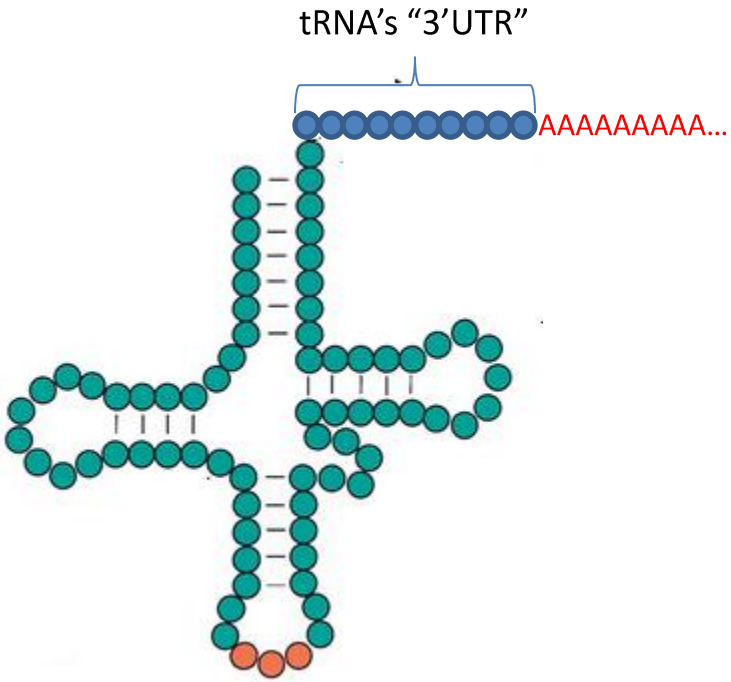
"REVIGO summarizes and visualizes long lists of Gene Ontology terms“ PLoS ONE 2011

Carbon Starvation affects poly(A) selection in 455 genes, preferring ***distal*** poly(A) sites

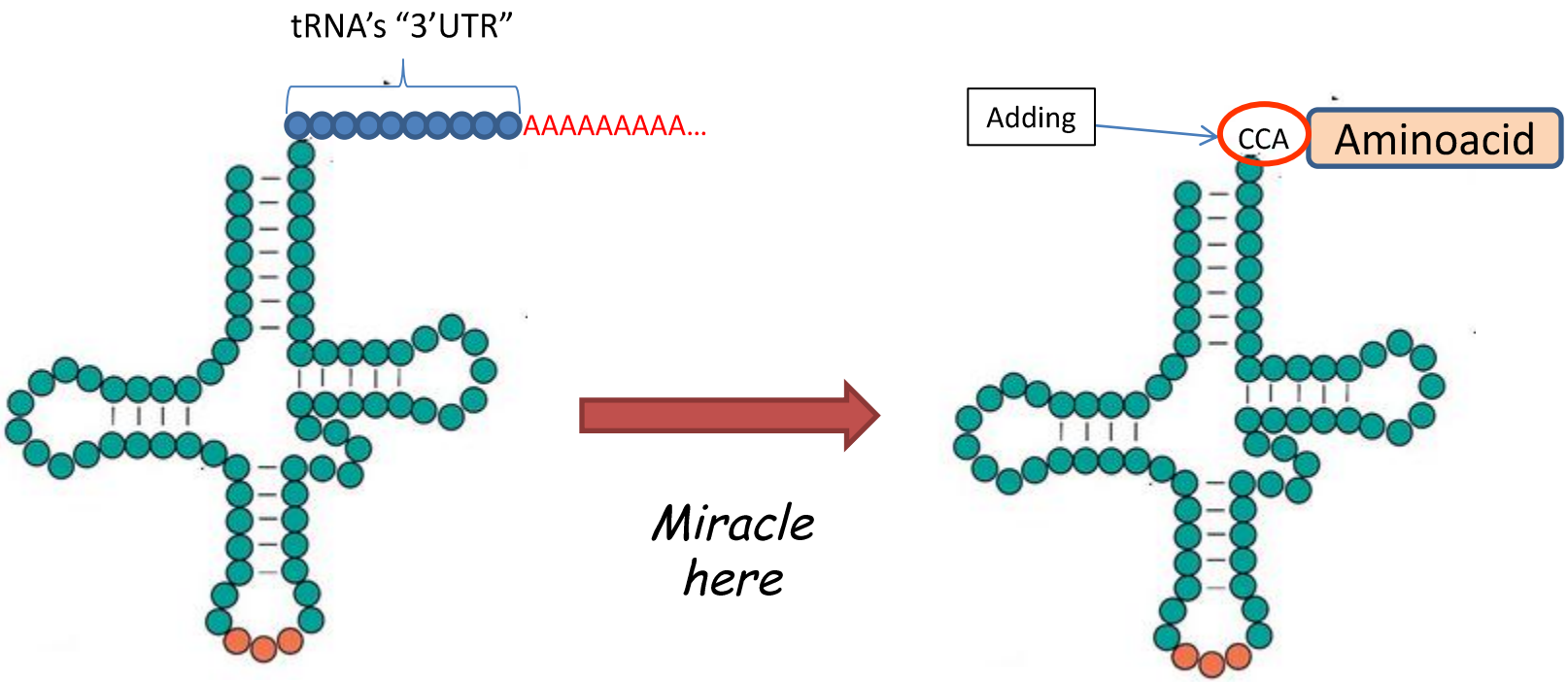


In carbon starvation longer UTRs are usually produced

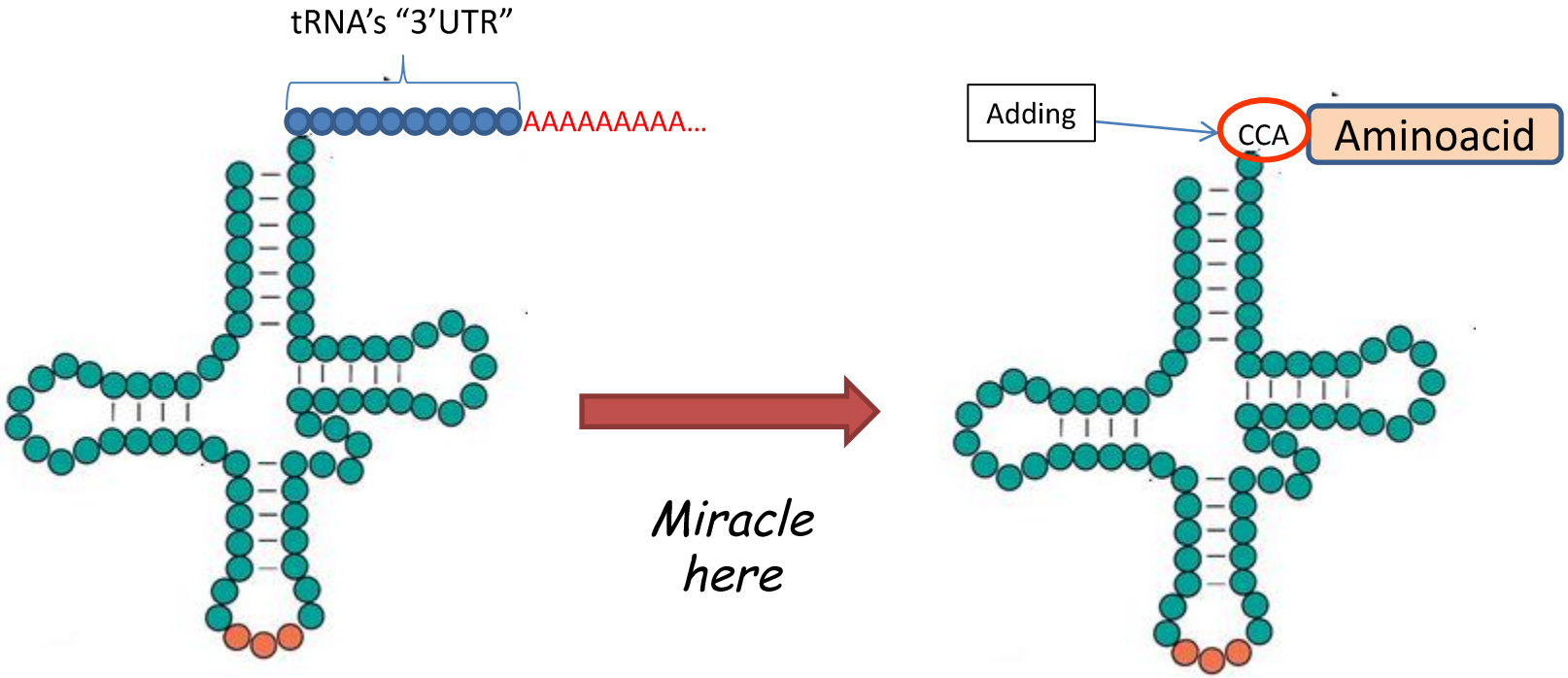
tRNAs are also polyadenylated

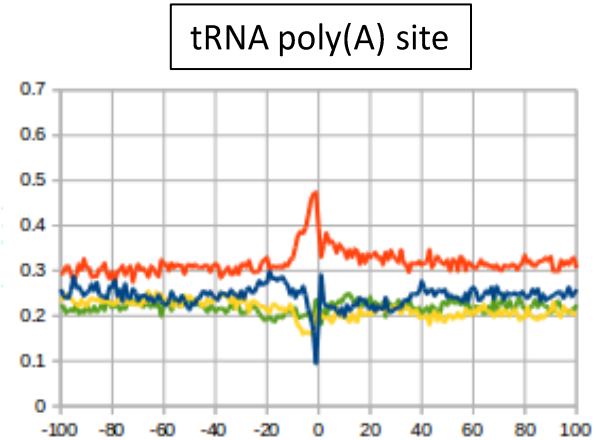


tRNAs are also polyadenylated



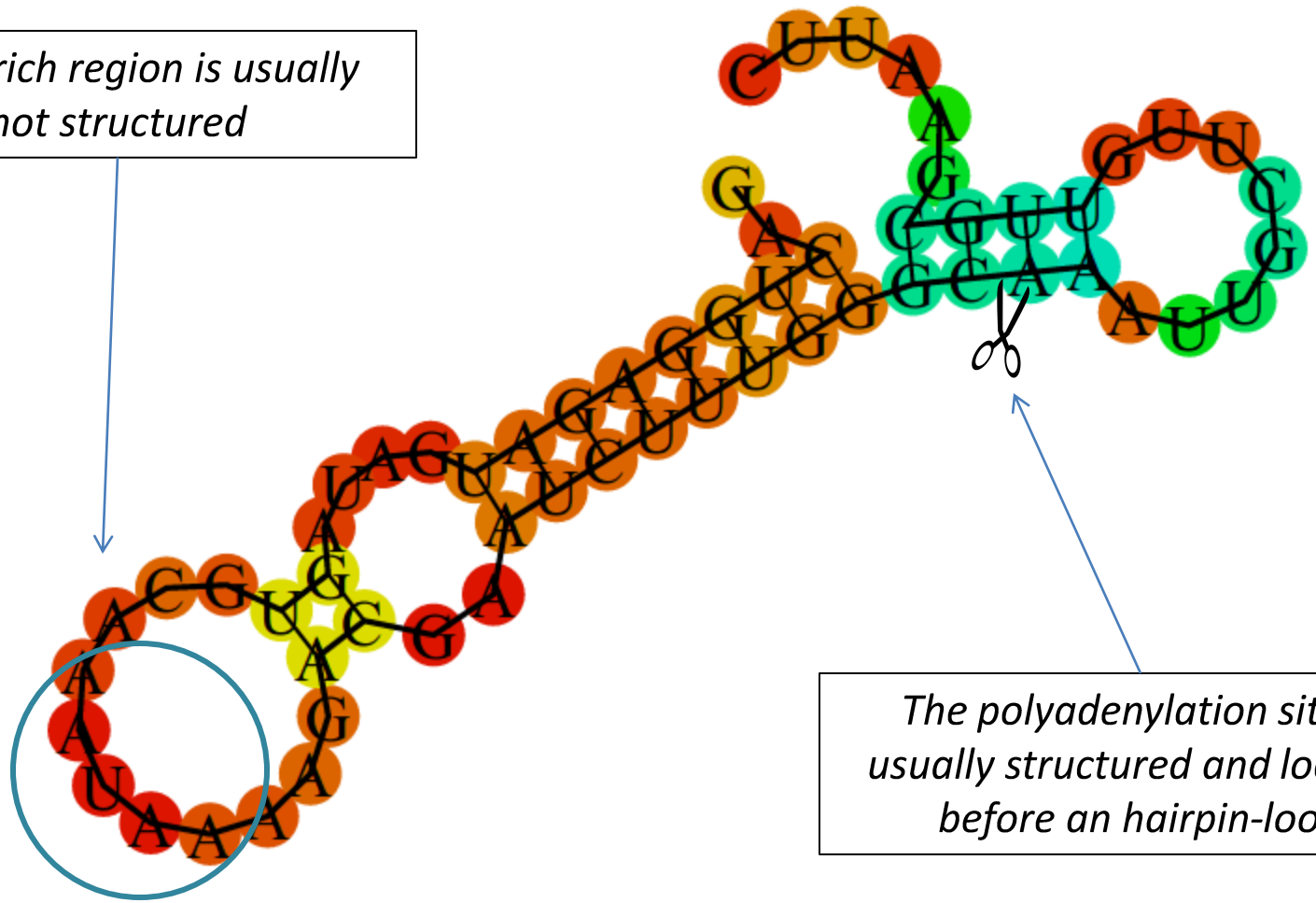
tRNAs are also polyadenylated





The polyadenylation site region has a defined secondary structure

The A-rich region is usually not structured



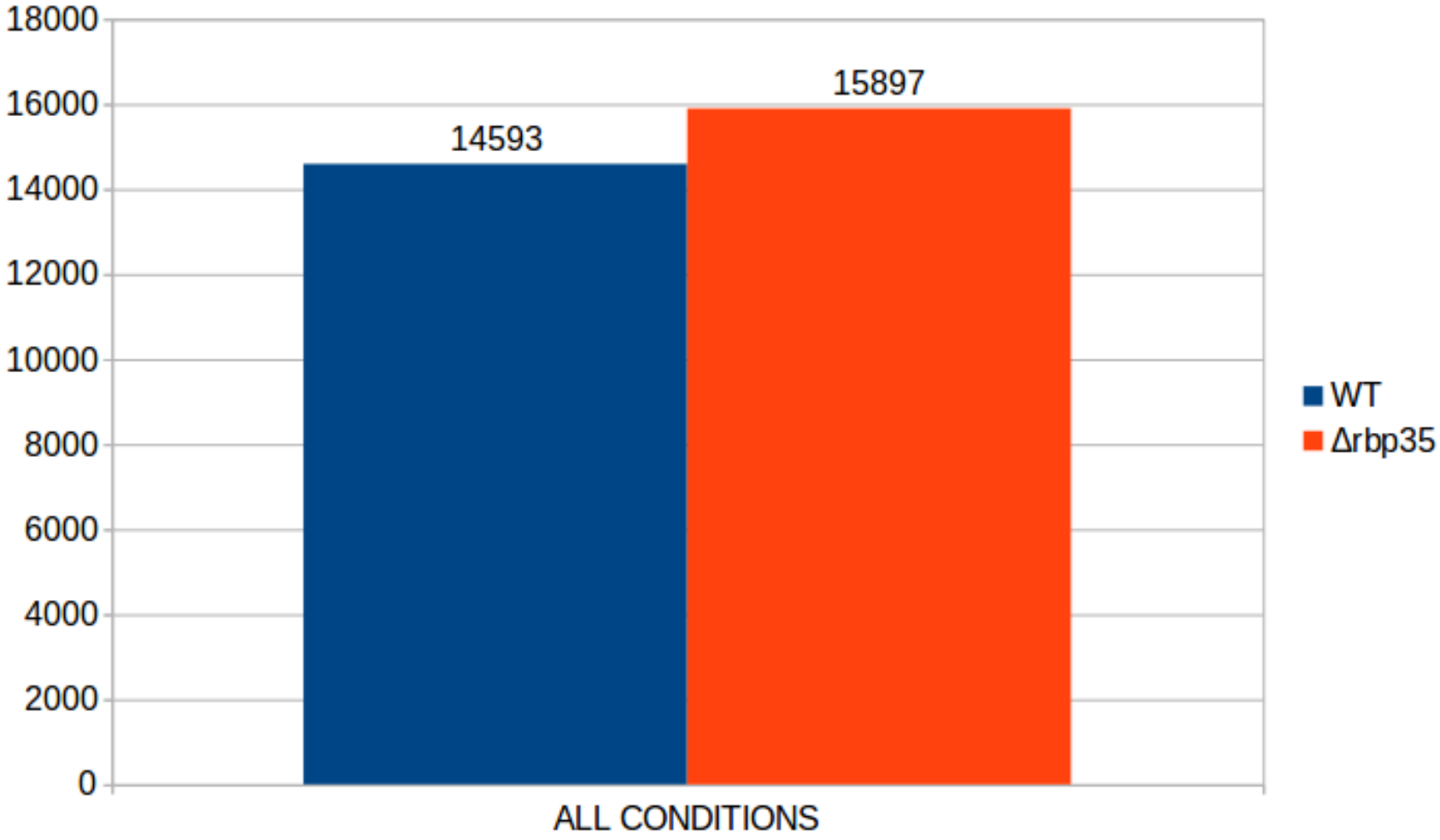
The polyadenylation site is usually structured and located before an hairpin-loop

Poly(A) site mapping analysis

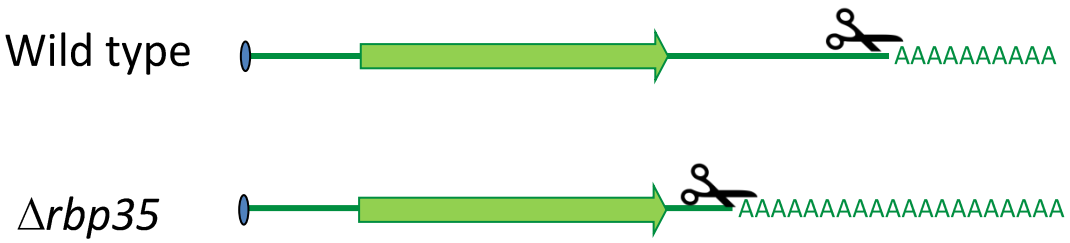
Results in wild-type

Results in the Δ*rbp35* mutant

Δ*rbp35* increases the number of Poly-A sites

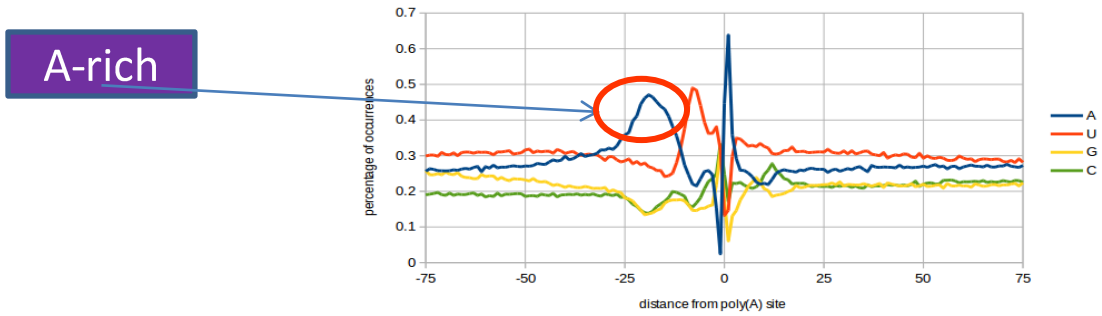


~25% of *M. oryzae*'s genes have a different poly(A) site usage in Δ*rbp35*

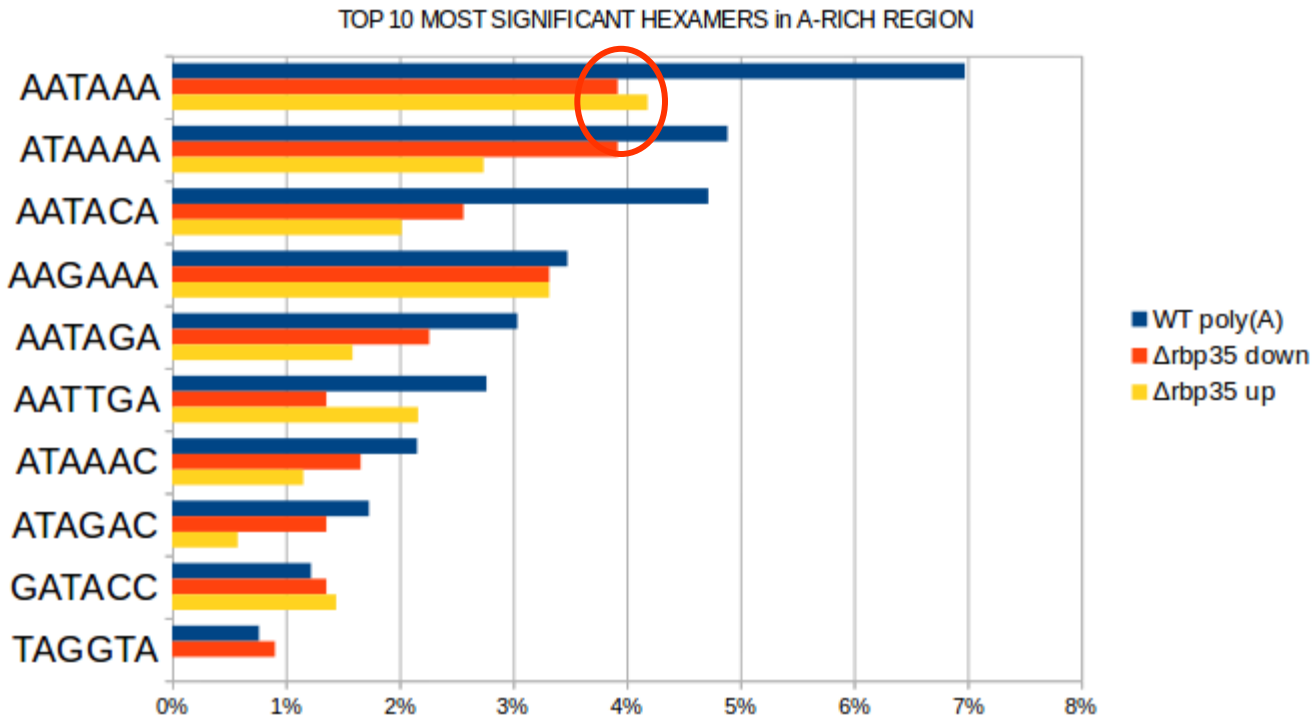


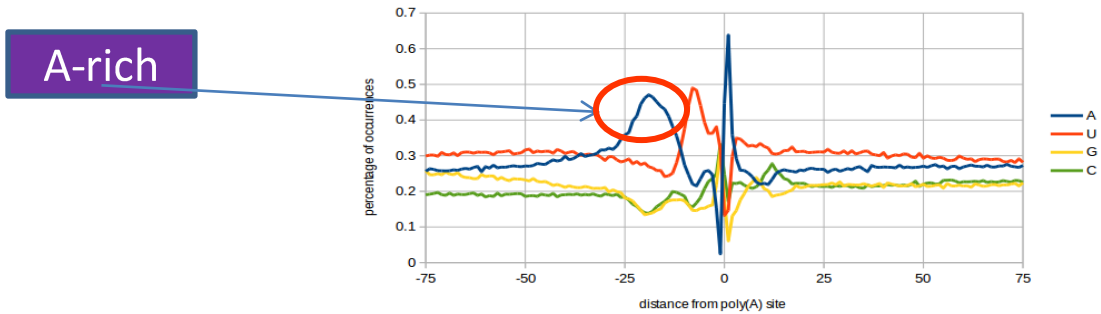
In the mutant, shorter UTRs are usually preferred

RBP35-dependent poly(A) sites are less likely to possess the canonical AAUAAA sequence

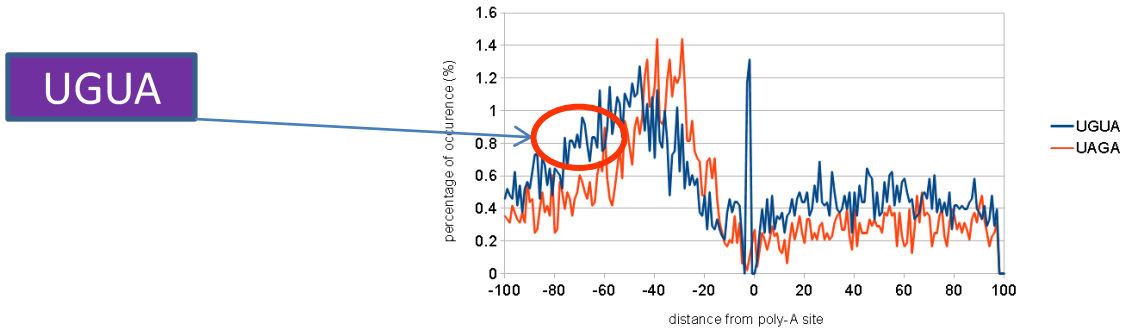


RBP35-dependent poly(A) sites are less likely to possess the canonical AAUAAA sequence

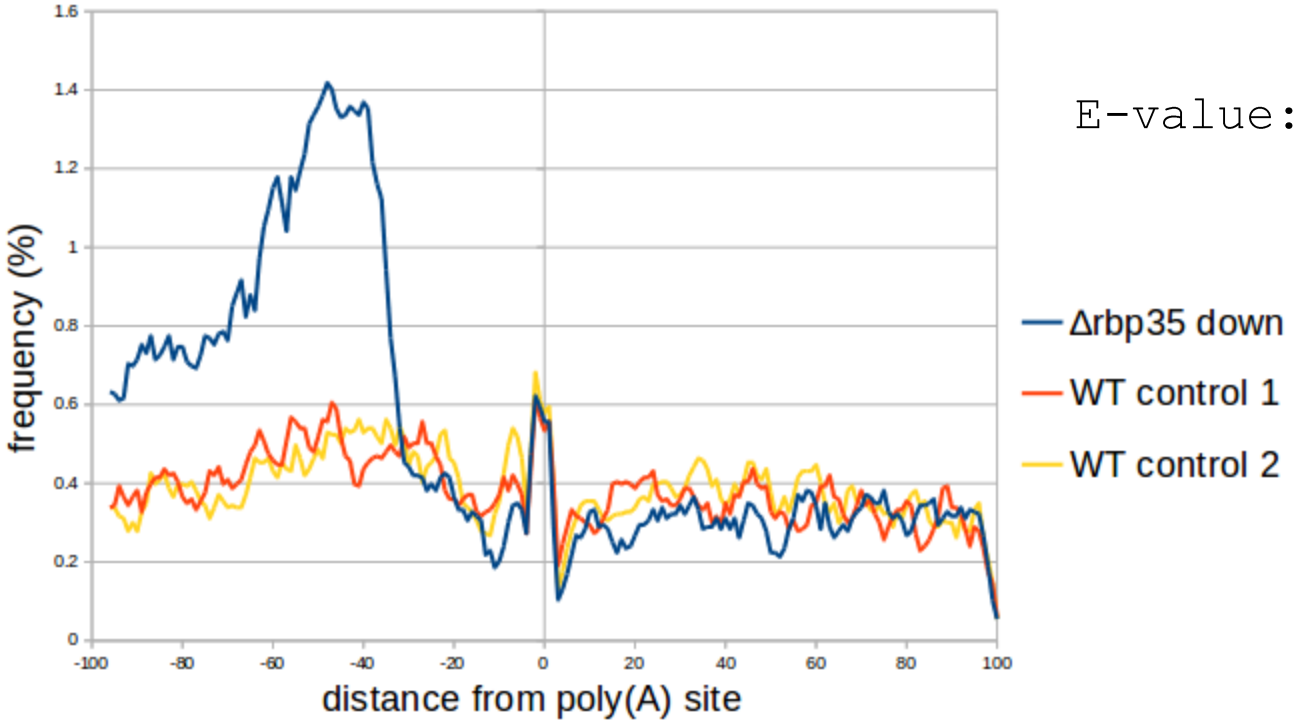




UGUA motif is very enriched in down-regulated poly(A) sites in Δ*rbp35*

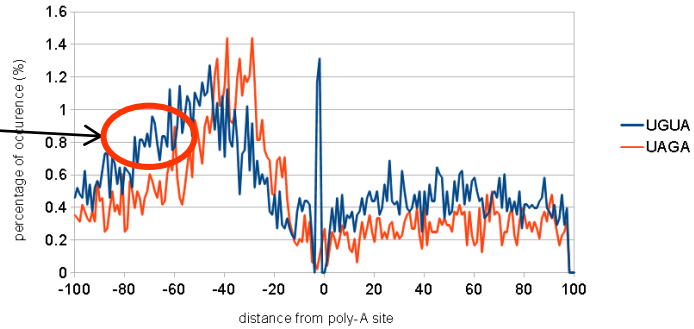


UGUA motif is very enriched in down-regulated poly(A) sites in Δ*rbp35*



E-value: 5.3e-103

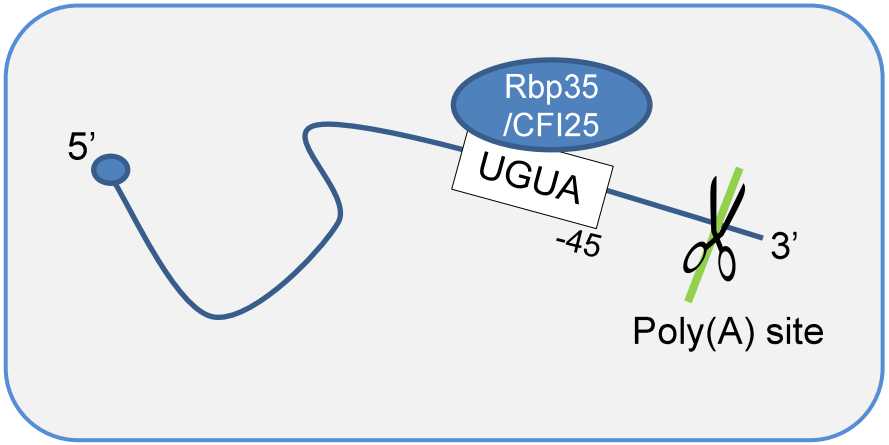
UGUA



Can we devise an hypothetical role for RBP35?

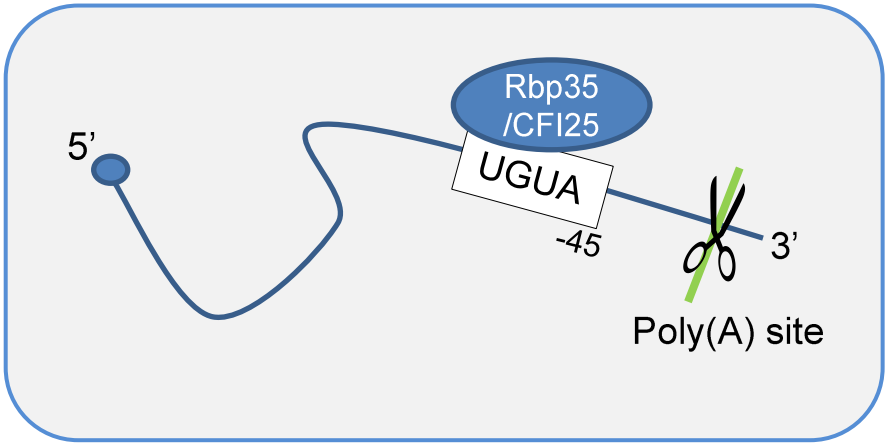
The recognition of a proper UGUA motif by Rbp35 is essential for the cleavage reaction.

Wild type strain (common situation)

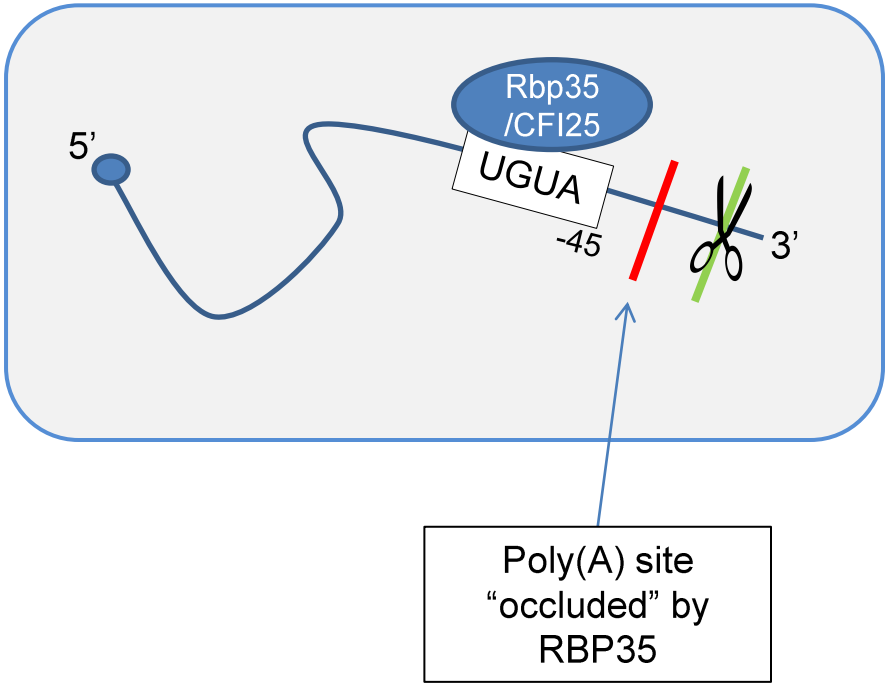


The recognition of a proper UGUA motif by Rbp35 is essential for the cleavage reaction.

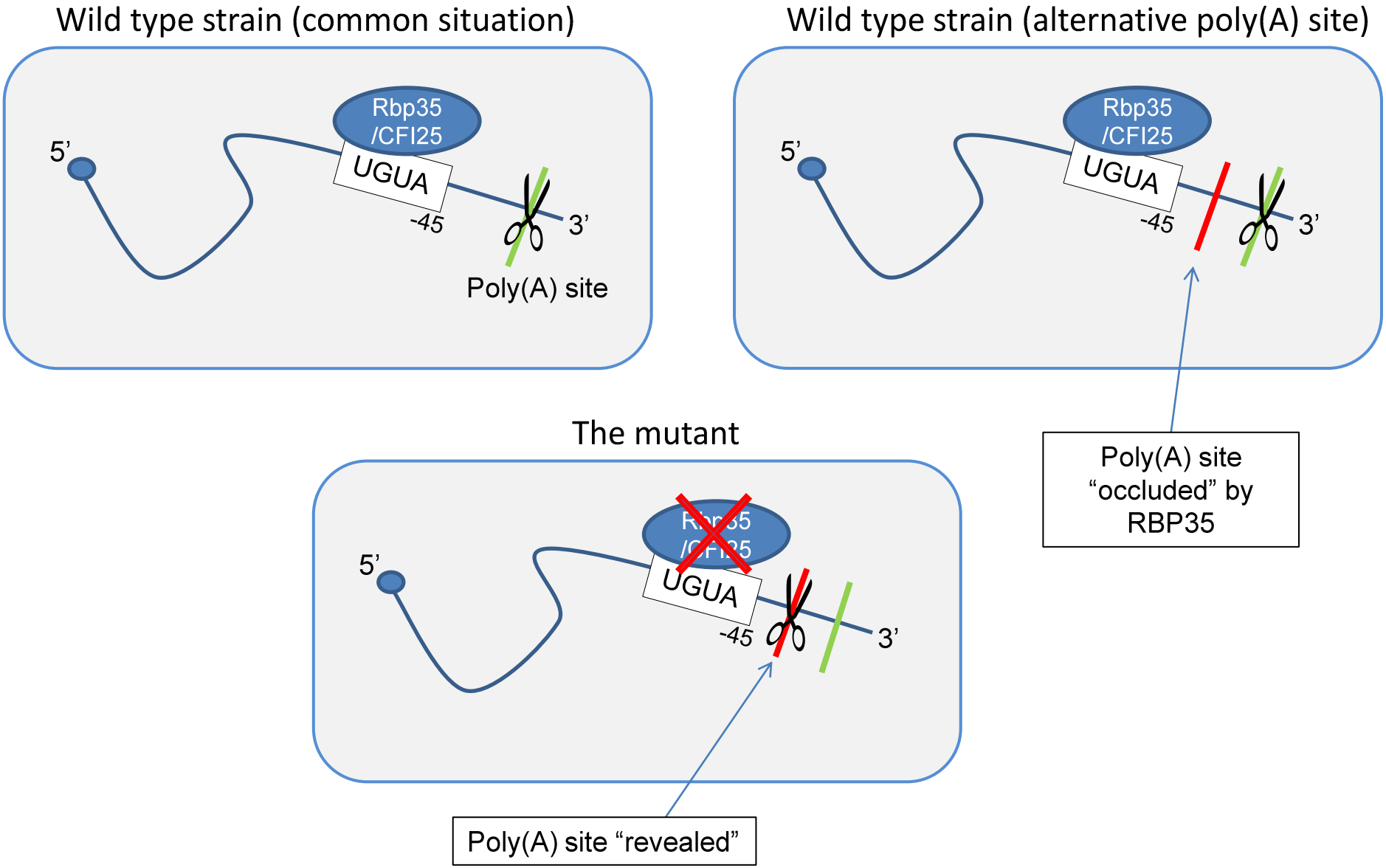
Wild type strain (common situation)



Wild type strain (alternative poly(A) site)



The recognition of a proper UGUA motif by Rbp35 is essential for the cleavage reaction.



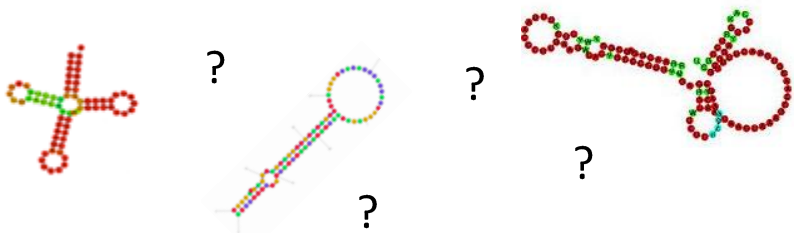
Part 1 - Conclusions

- Polyadenylation in *M. oryzae* differs from S. cerevisiae
- Alternative polyadenylation in *M. oryzae* is common and related with specific functional groups
- Carbon starvation induces longer 3’UTRs
- RBP35 (probably) interacts with the UGUA element, to select distal poly(A) sites

We want to shed light on the role of RBP35 and EXP5 in *M.oryzae:*

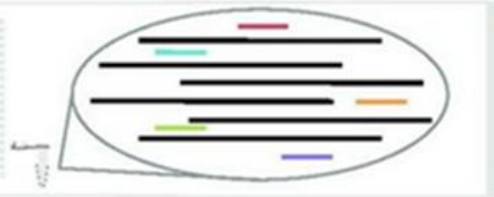
- RNA-seq on polyadenylation sites

- Small RNA-sequencing



Small RNA library preparation and sequencing

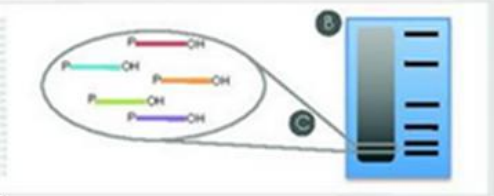
Total RNA isolation



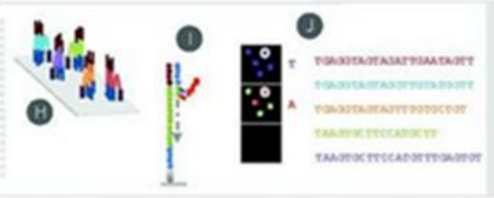
HD Adaptor ligation

RT & PCR

Size fractionation



Sequencing



Adriana Illana Otero, in Norwich

Sequencing details

<u>Three strains:</u>

WT

Δexp5

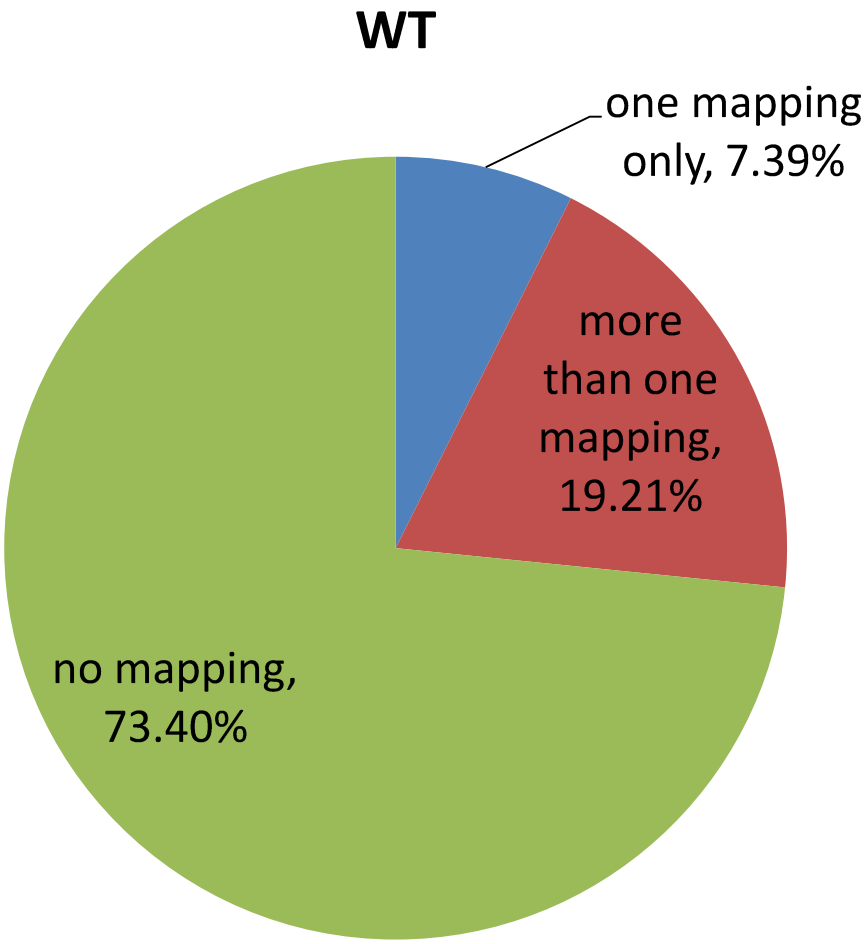
Δrbp35

<u>3 Replicates:</u>

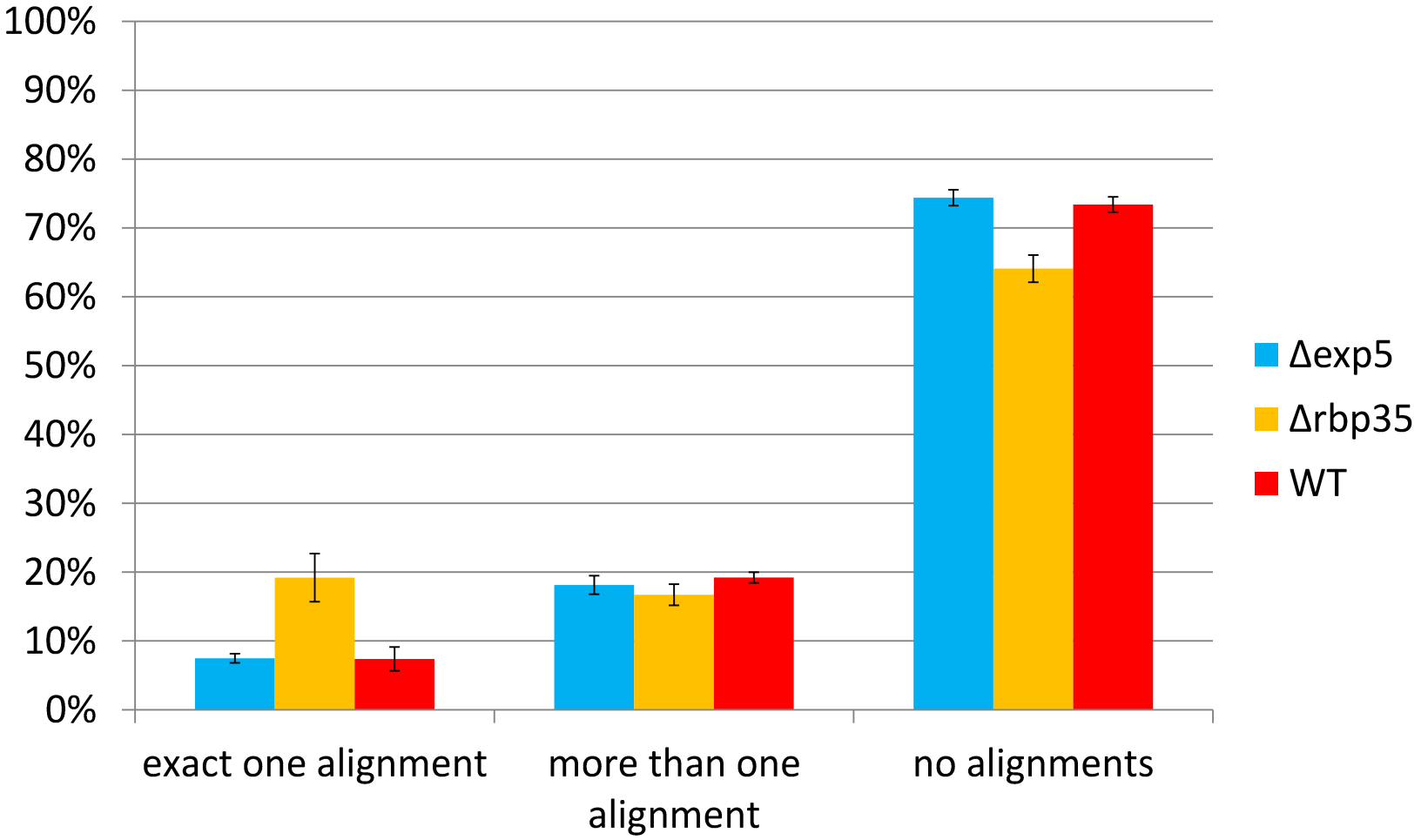
#1, #2, #3

Let’s see what we have on the annotated genome…..

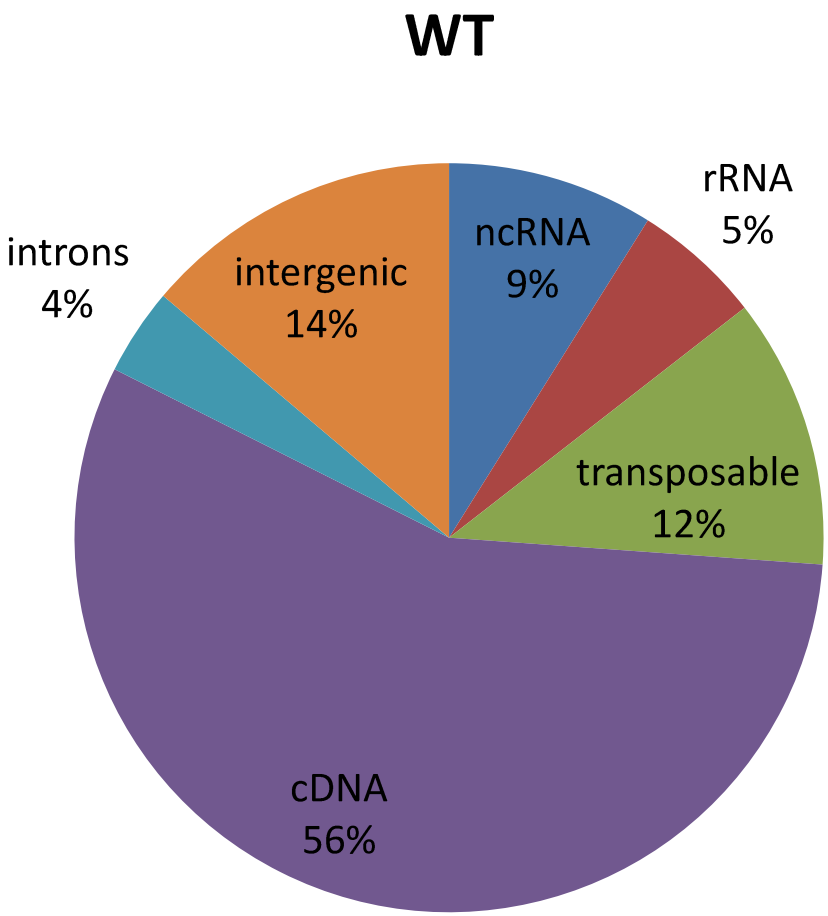
Only 25% of reads map on the genome



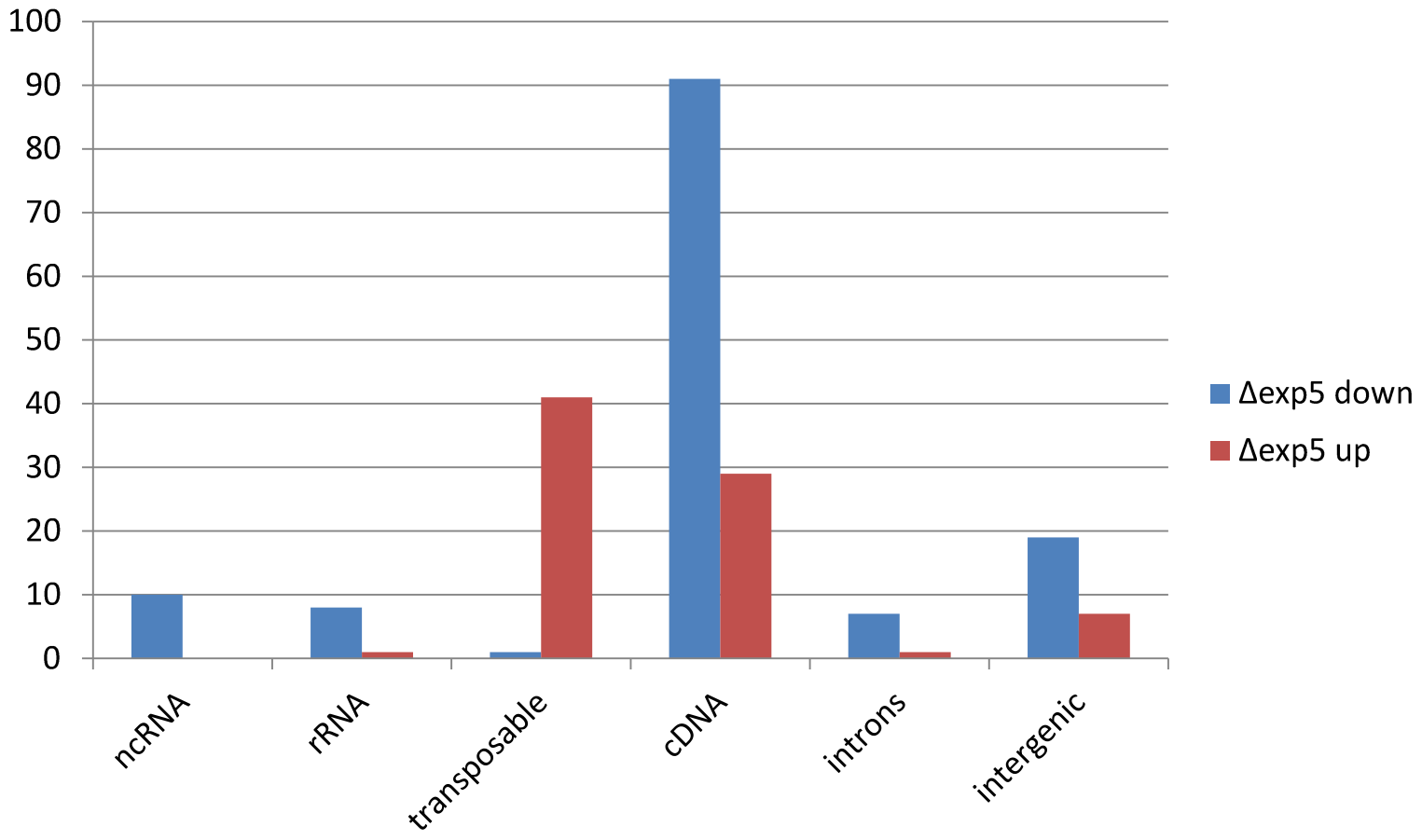
Δrbp35 reads align slighty more



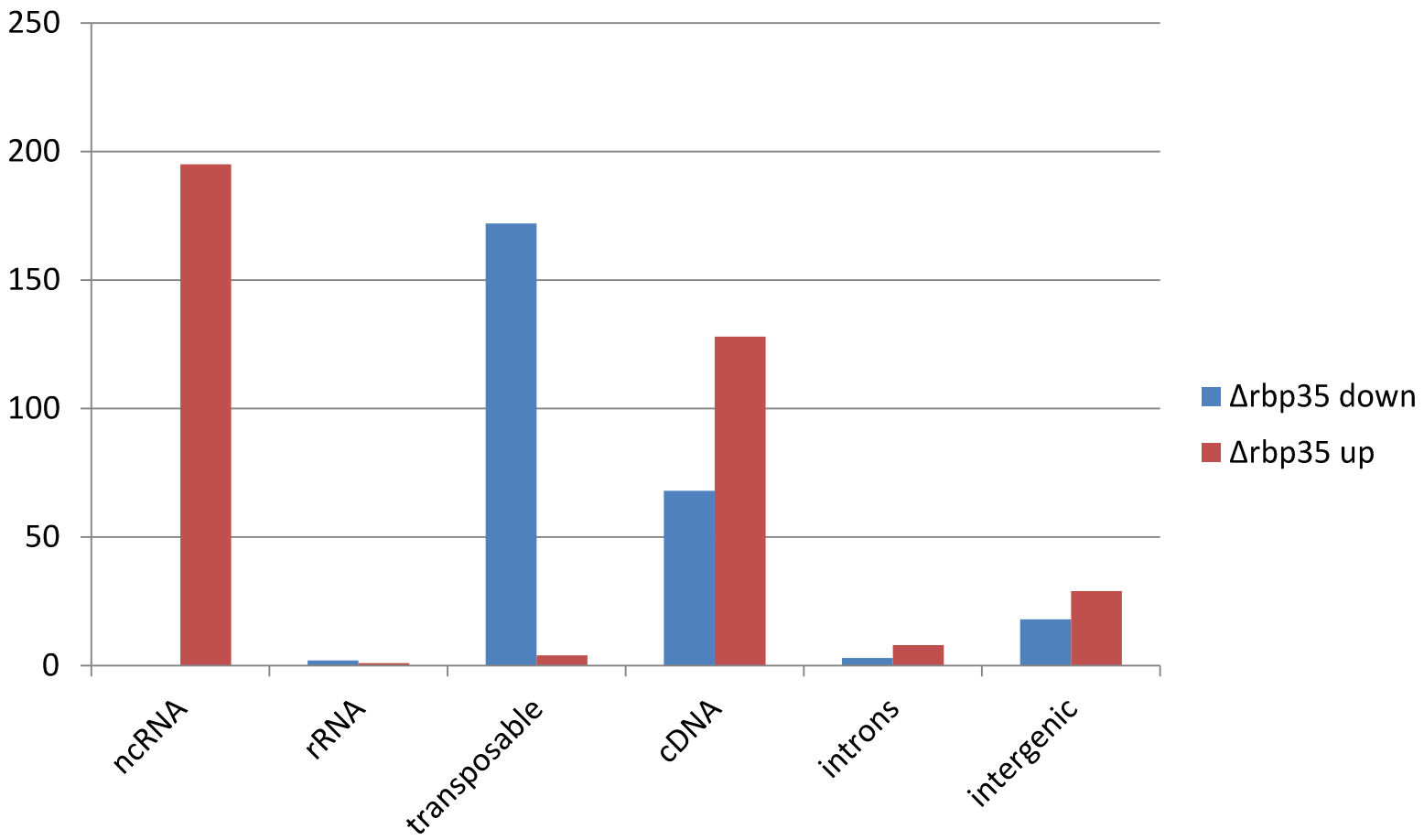
Most of genome sRNA loci (4131) are located in mRNAs



In *Δexp5*, a lot of protein-coding related loci are down regulated

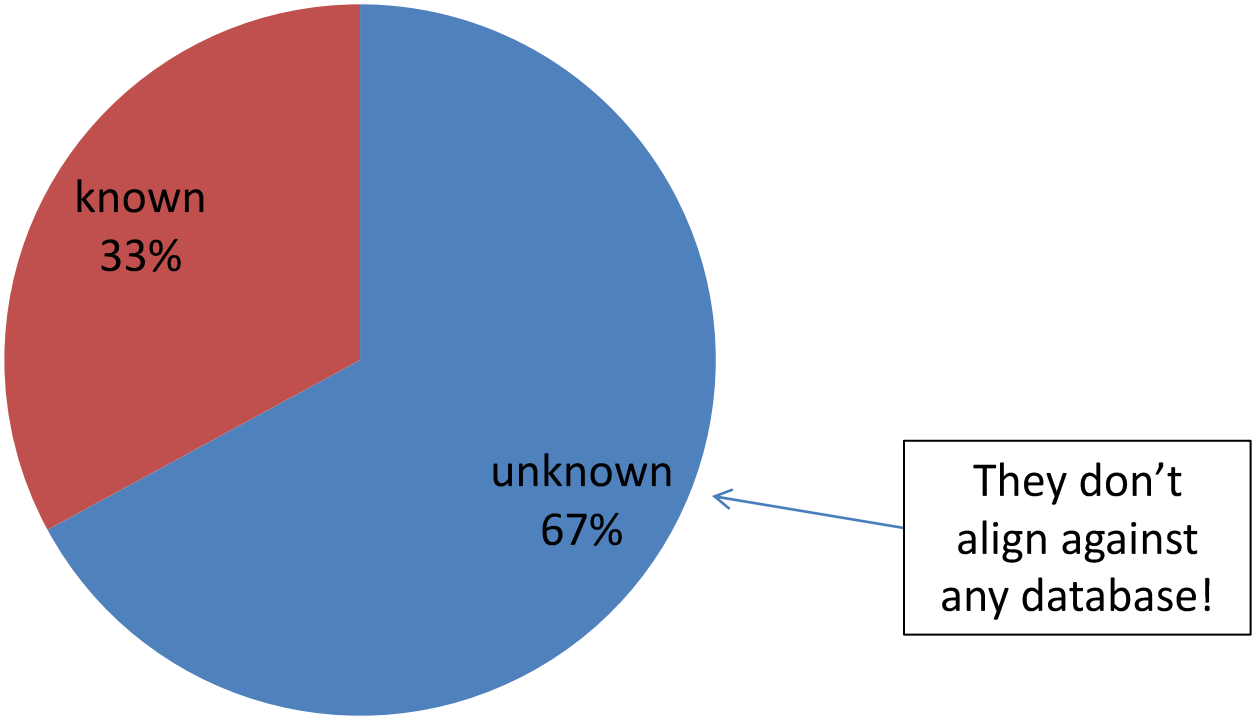


In *Δrbp35*, a lot of transposable elements related loci are down regulated, and many ncRNA related loci are up regulated

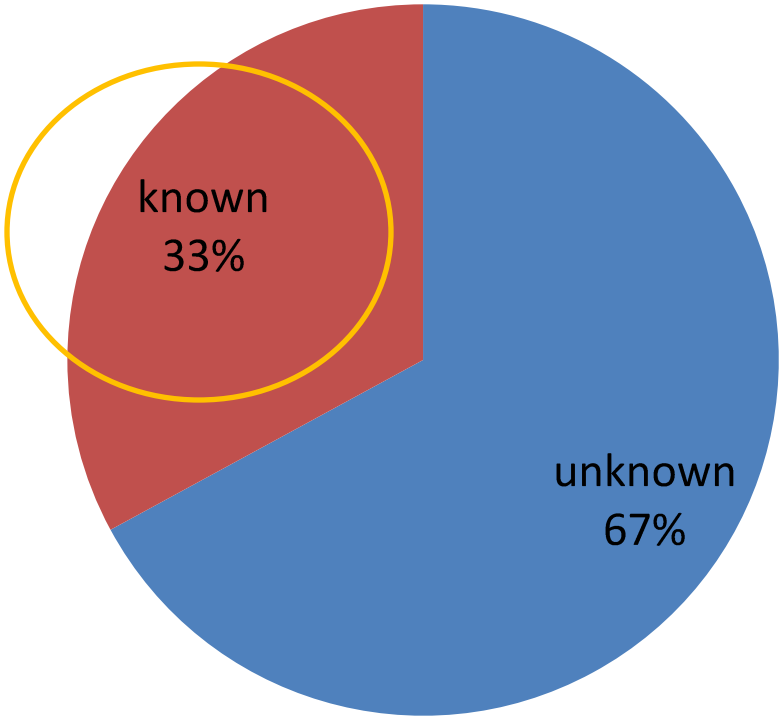


…and this was 25% of the story, so let’s forget about the annotated genome for a while…

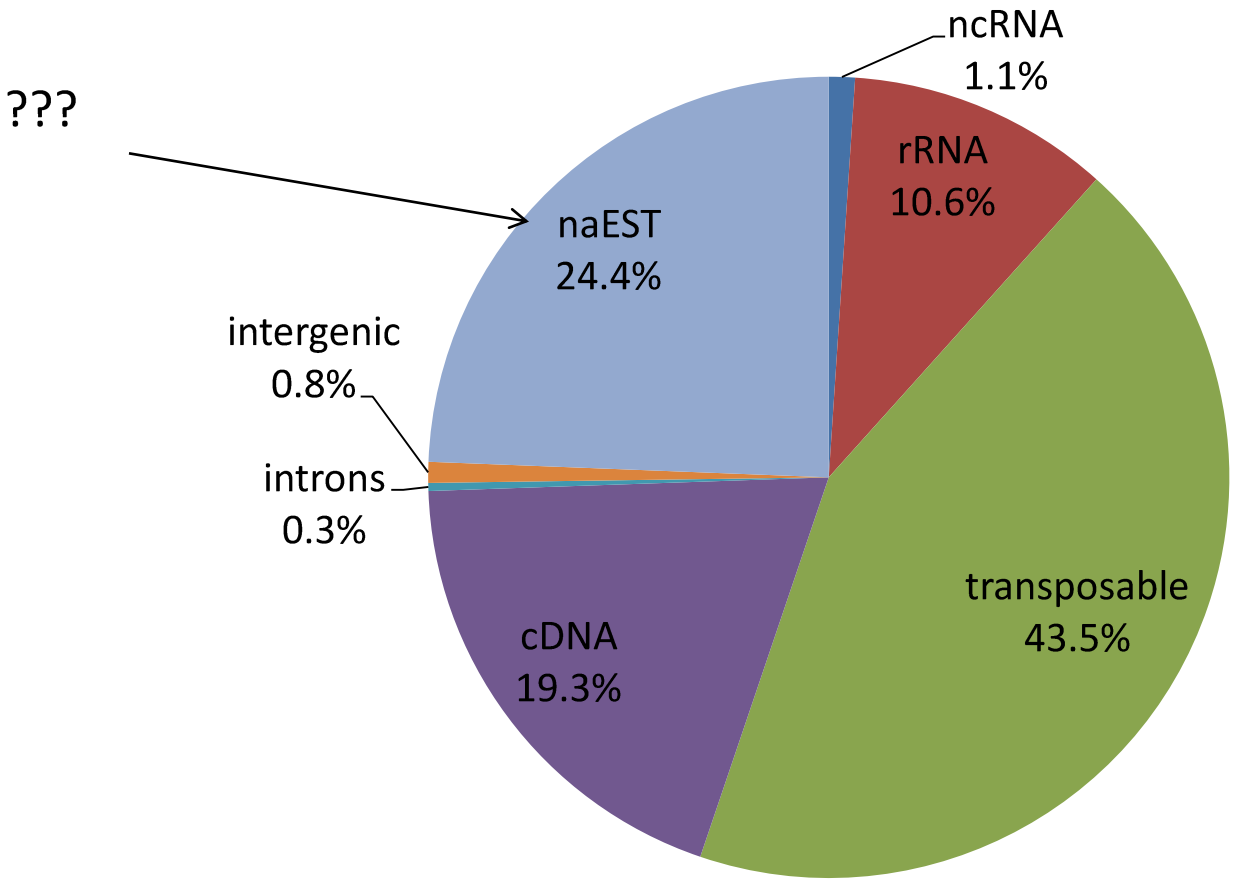
The majority of reads are of unknown origin



The majority of reads are of unknown origin

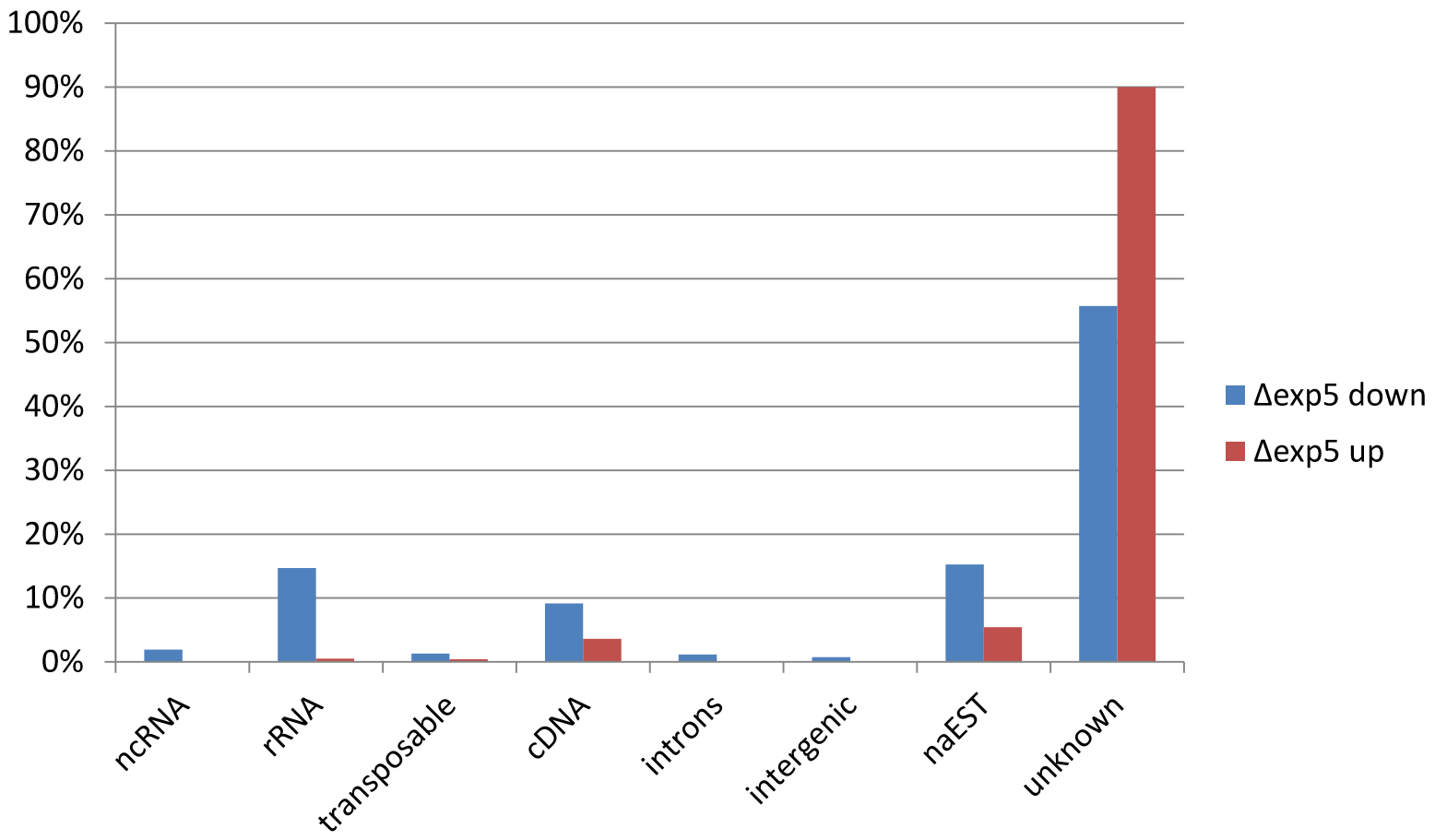


Many reads map to unknown ESTs non-aligning on the genome!

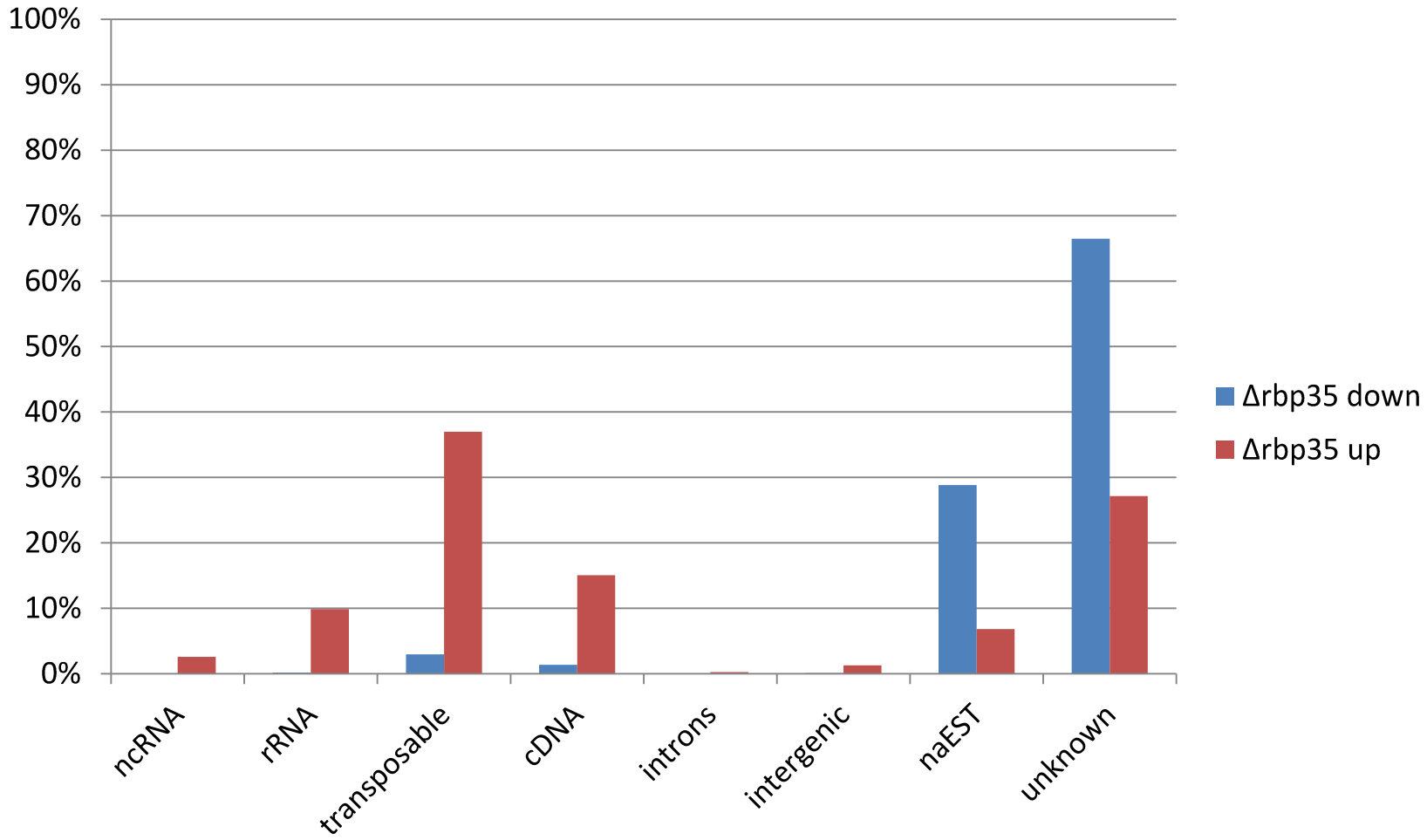


So, what’s happening in the mutants?

We cannot classify most of the changes in *Δexp5*

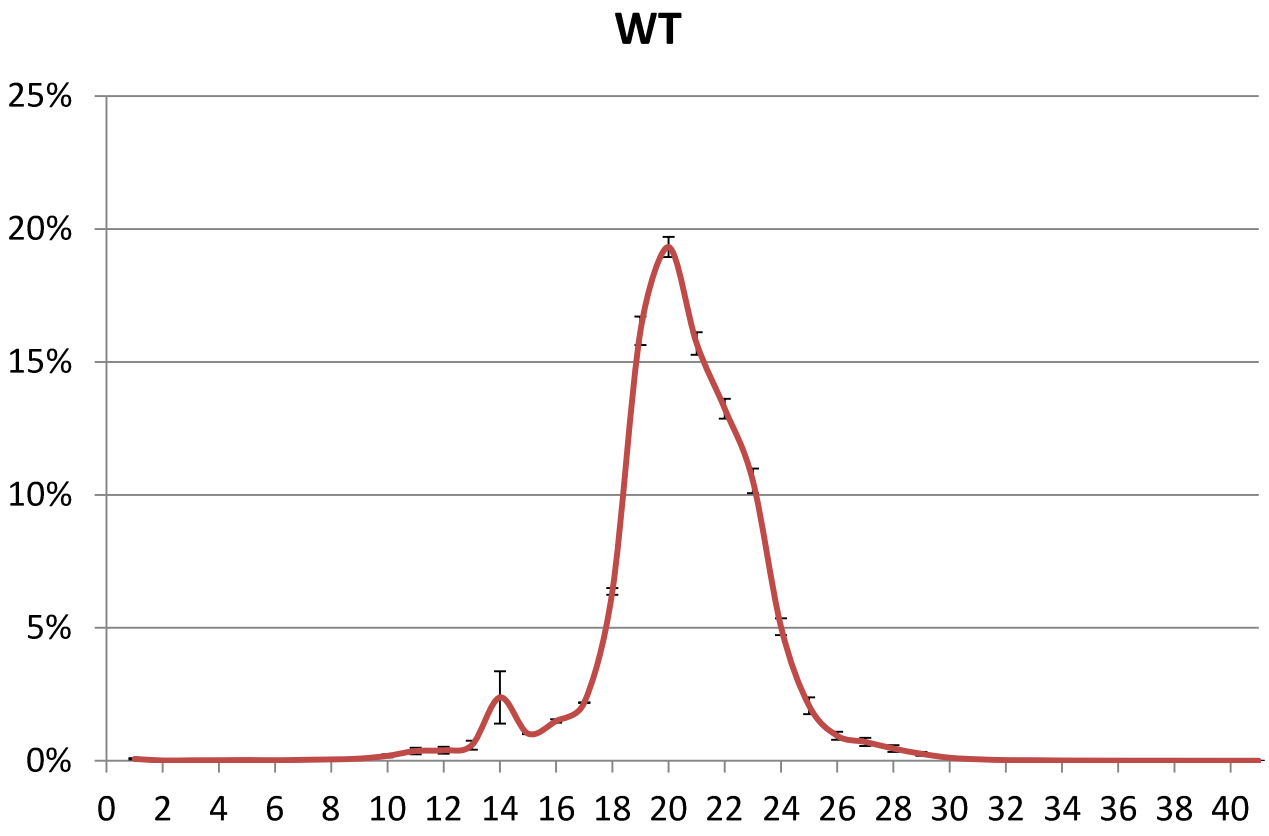


We have more luck with *Δrbp35*

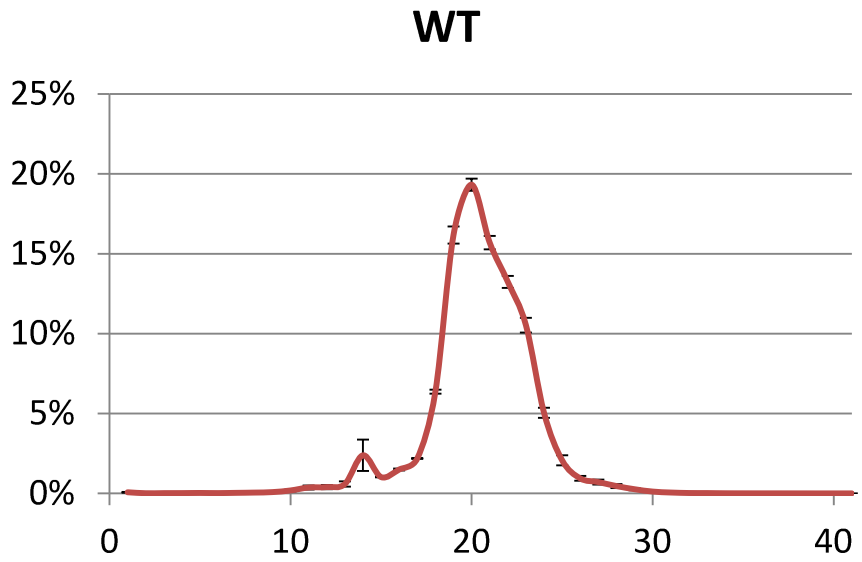


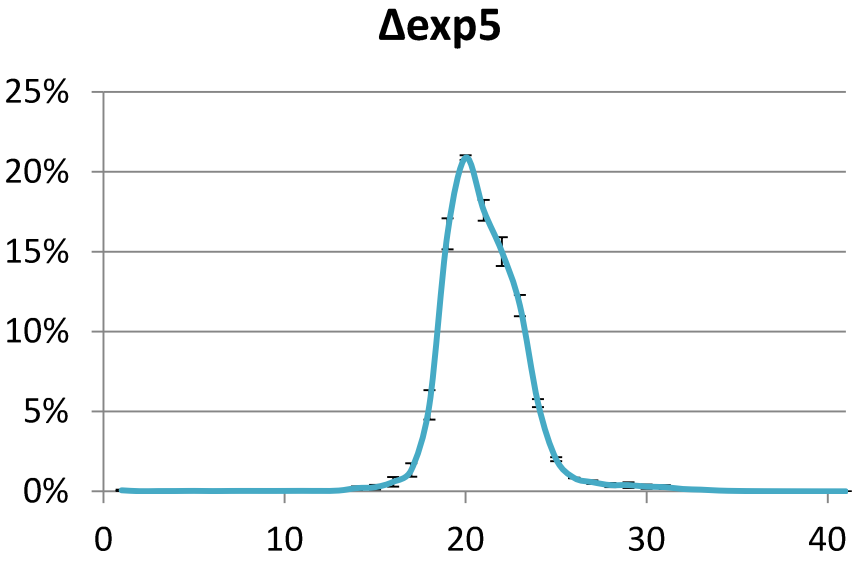
We don’t know what they are…
but at least, how do they look like?

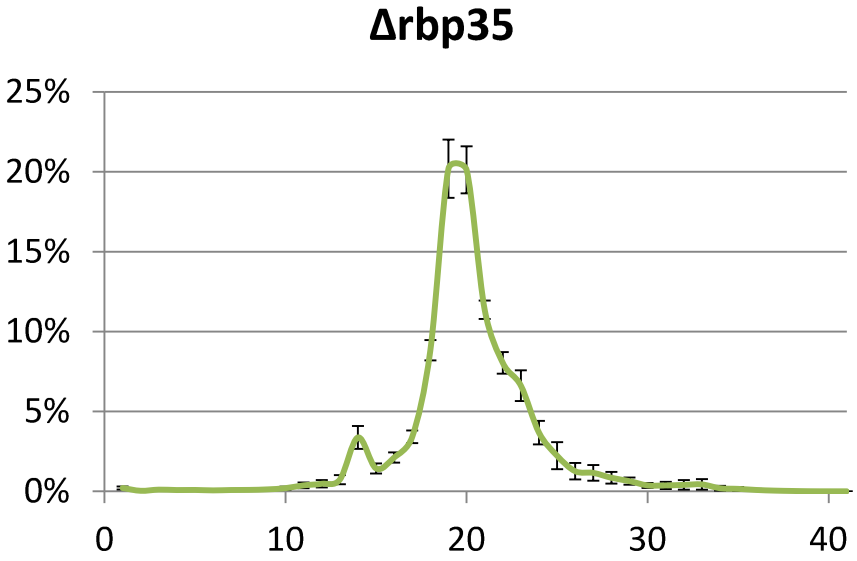
We obtained reads in the range 18-25 bp



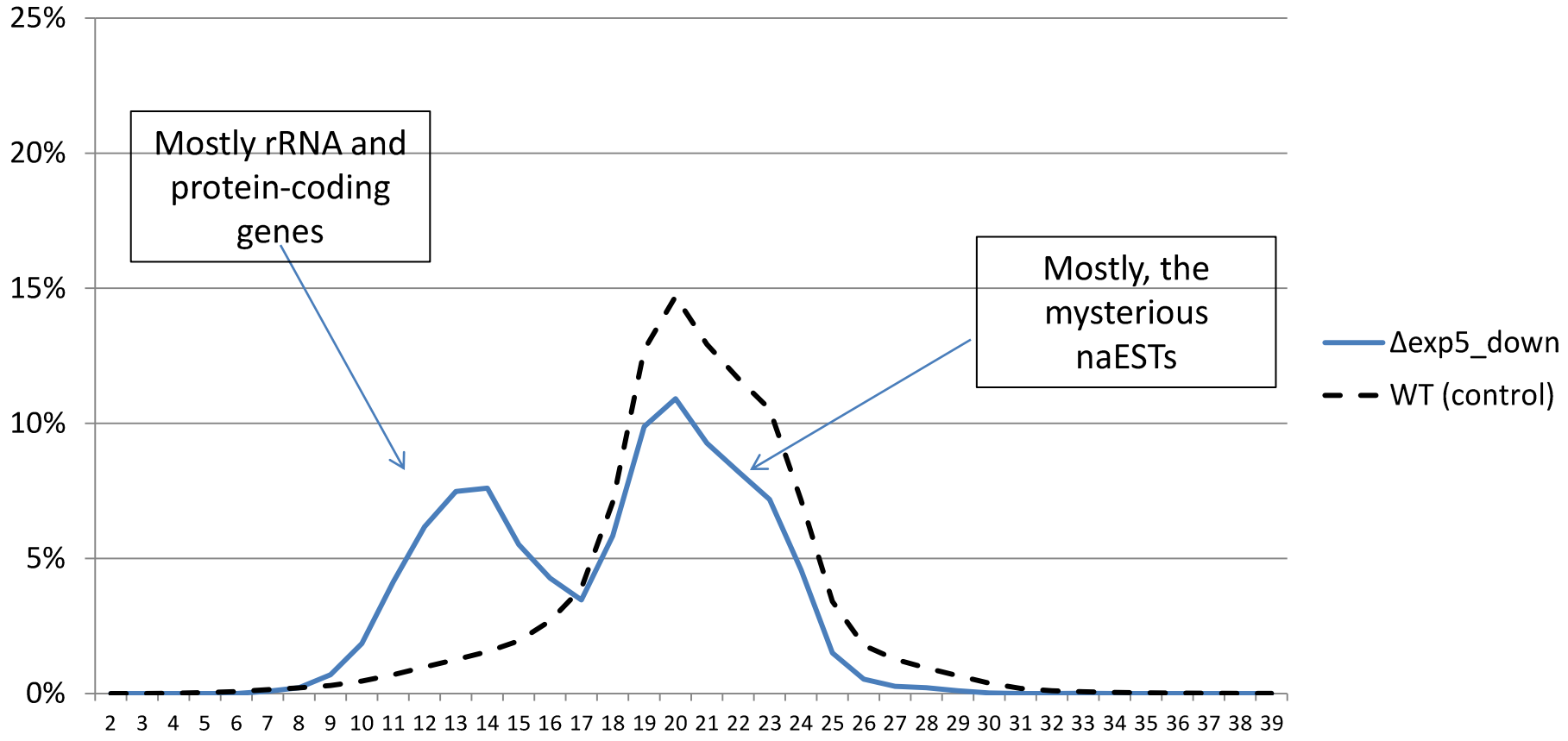
Δrbp35 reads look kind of shorter, and *Δexp5* reads are missing a peak at 14 bp…



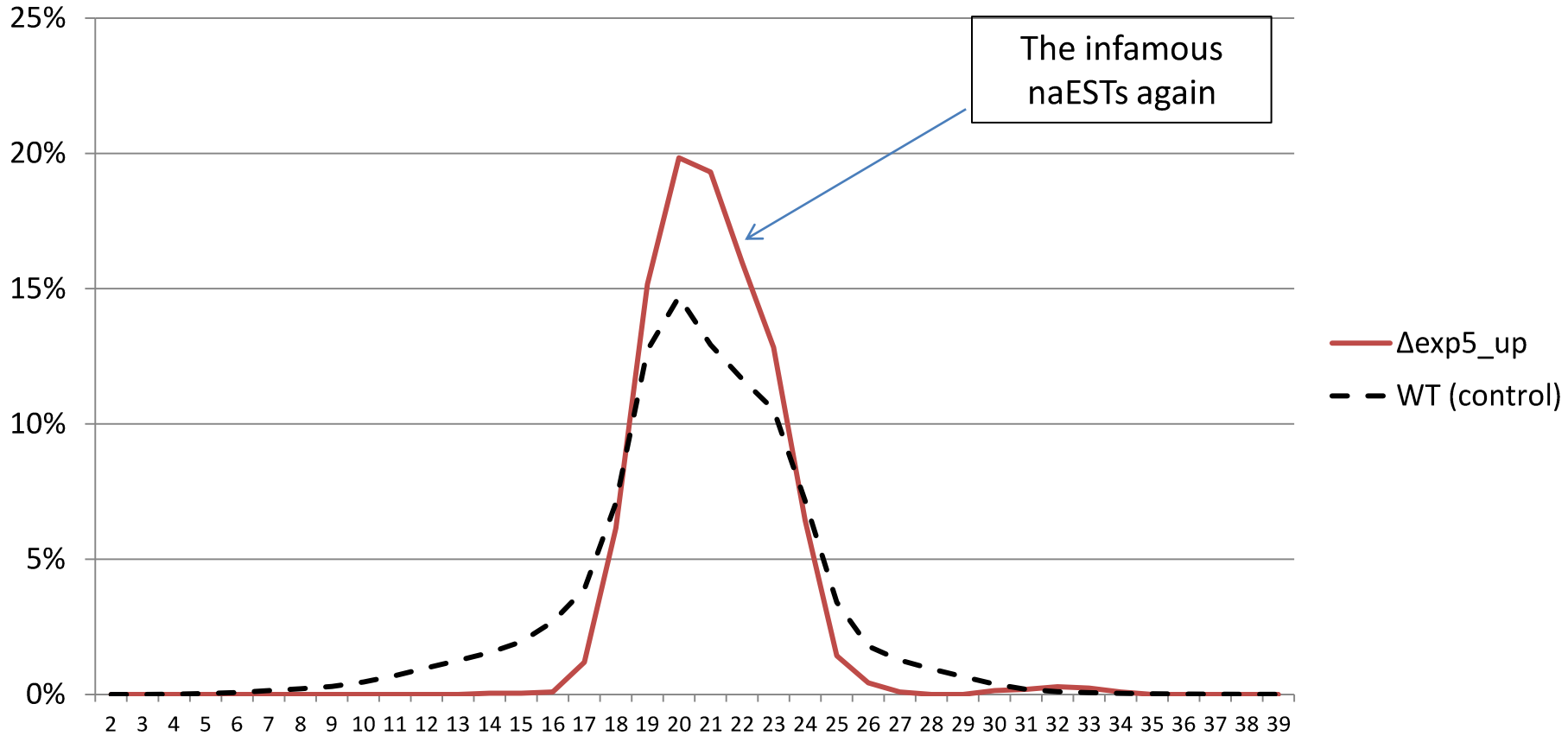




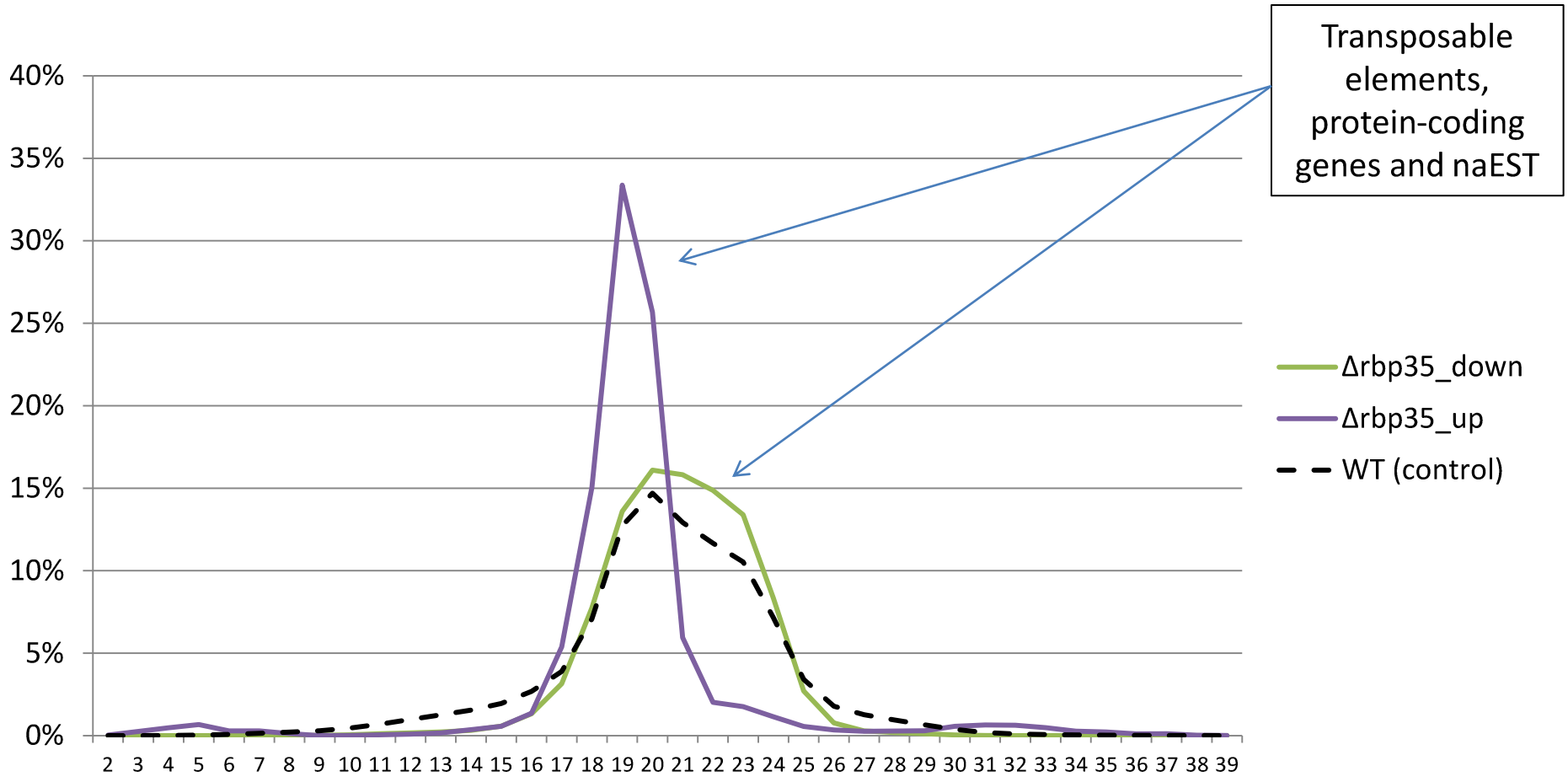
In Δexp5 we observe the down-regulation of very short fragments



In Δexp5 we observe the down-regulation of very short fragments

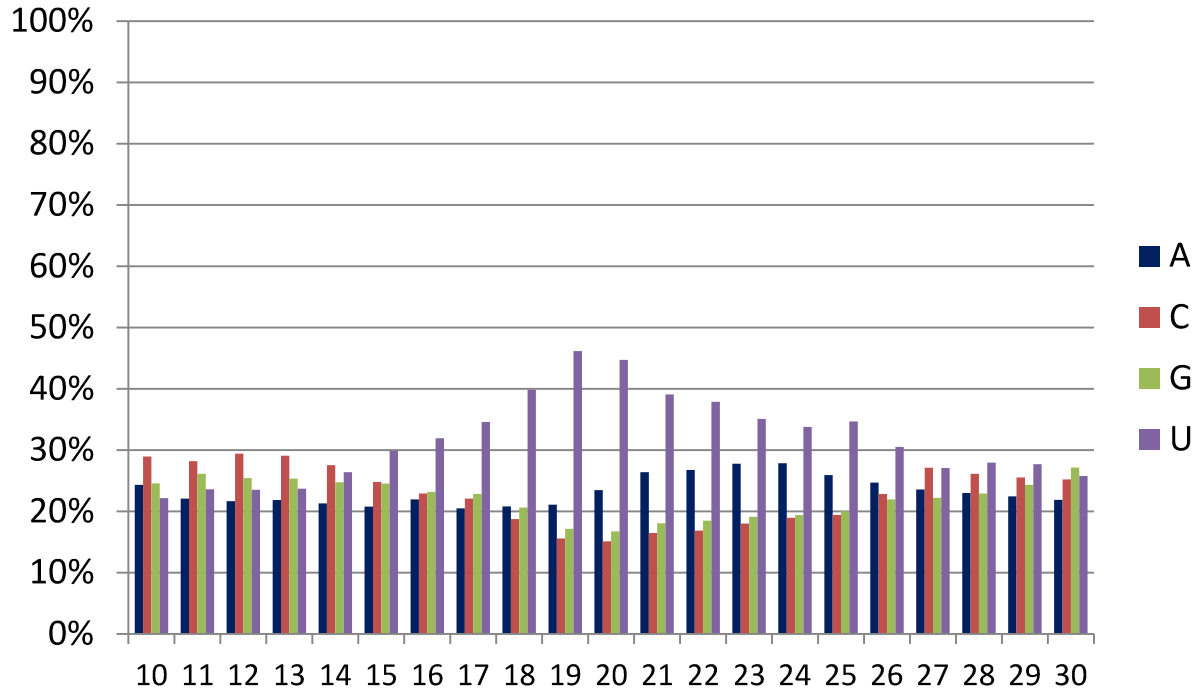


In Δrbp35 we observe a peak of up-regulated 19 nt sequences



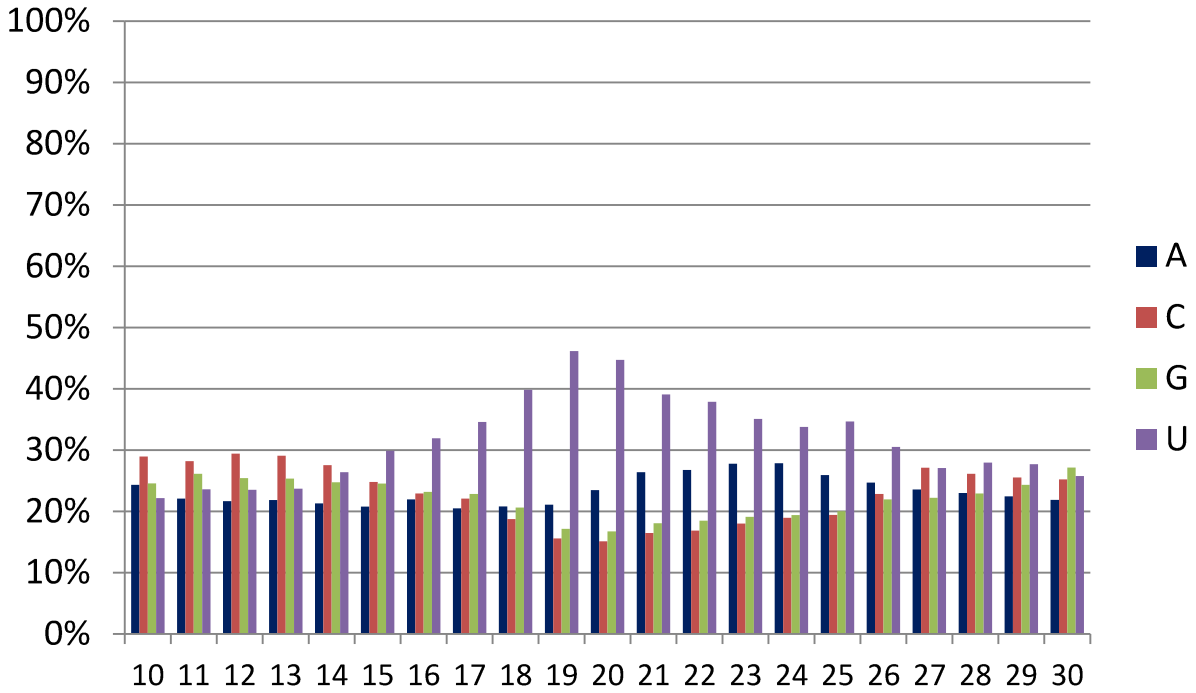
There is an overall U preference as 1st nucleotide, in reads with length 16-26 nt

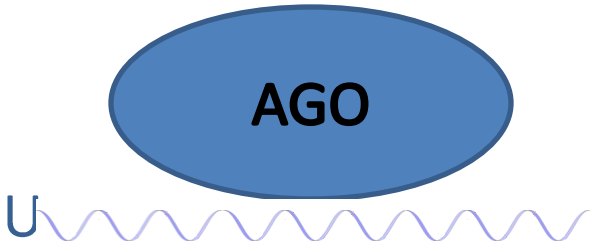
Mainly in protein-coding genes and transposable elements



There is an overall U preference as 1st nucleotide, in reads with length 16-26 nt

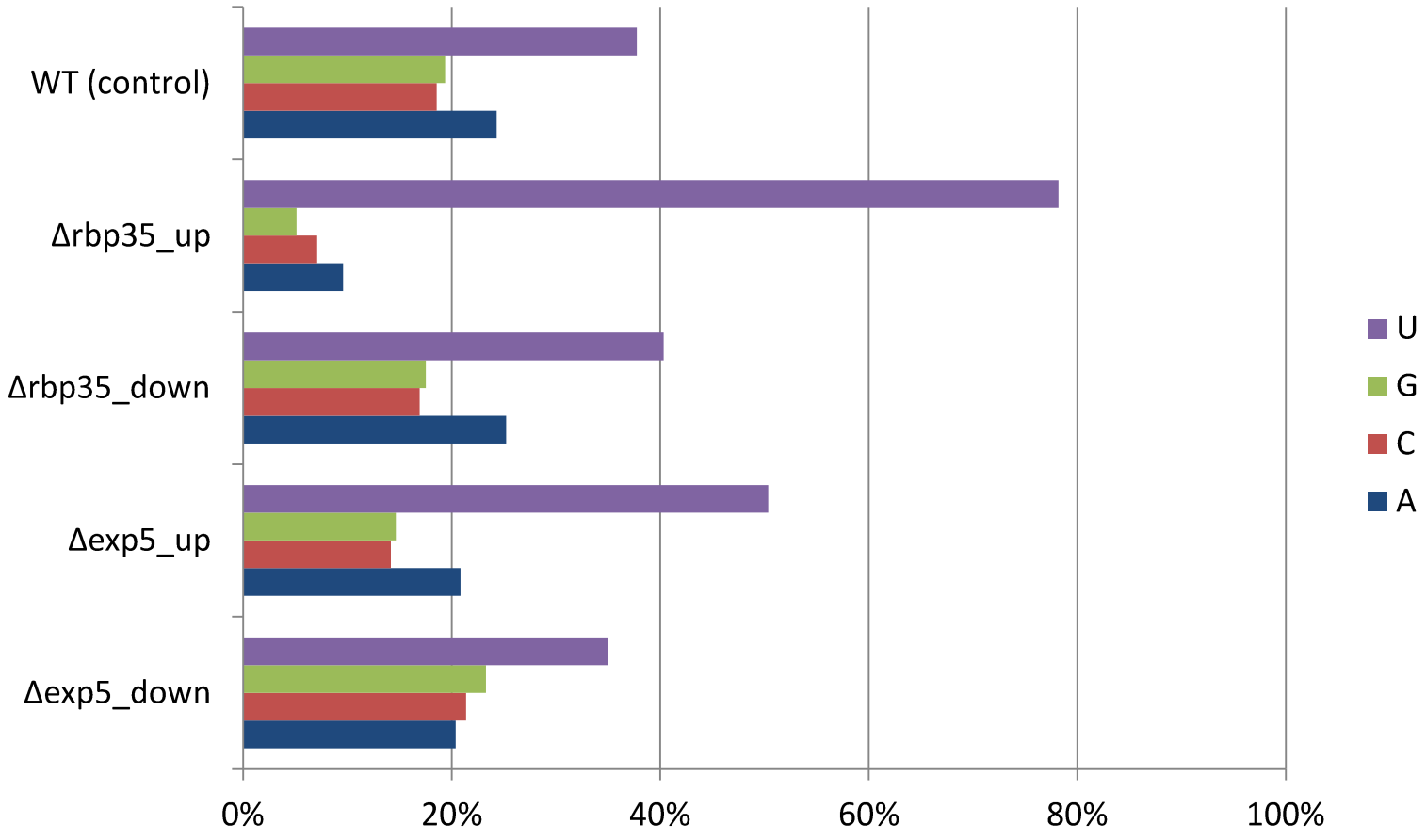
Mainly in protein-coding genes and transposable elements





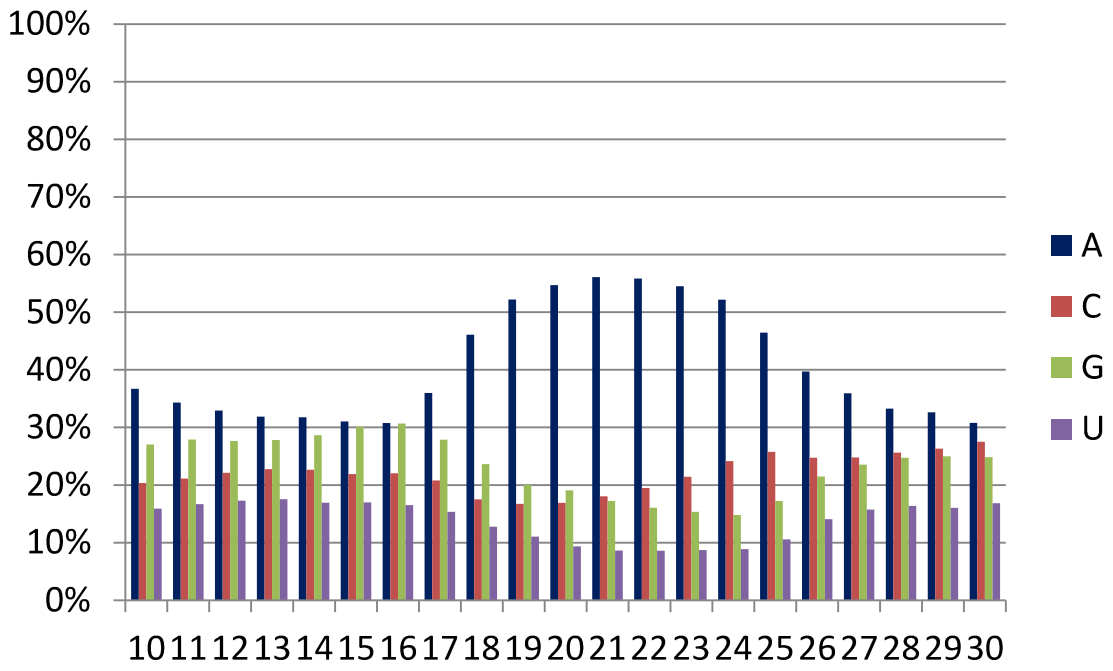
Certain proteins involved in RNA-interference like Argonaut have a preference for 5' uraciles

Δrbp35 small RNAs with 1st nucleotide U-preference are up-regulated



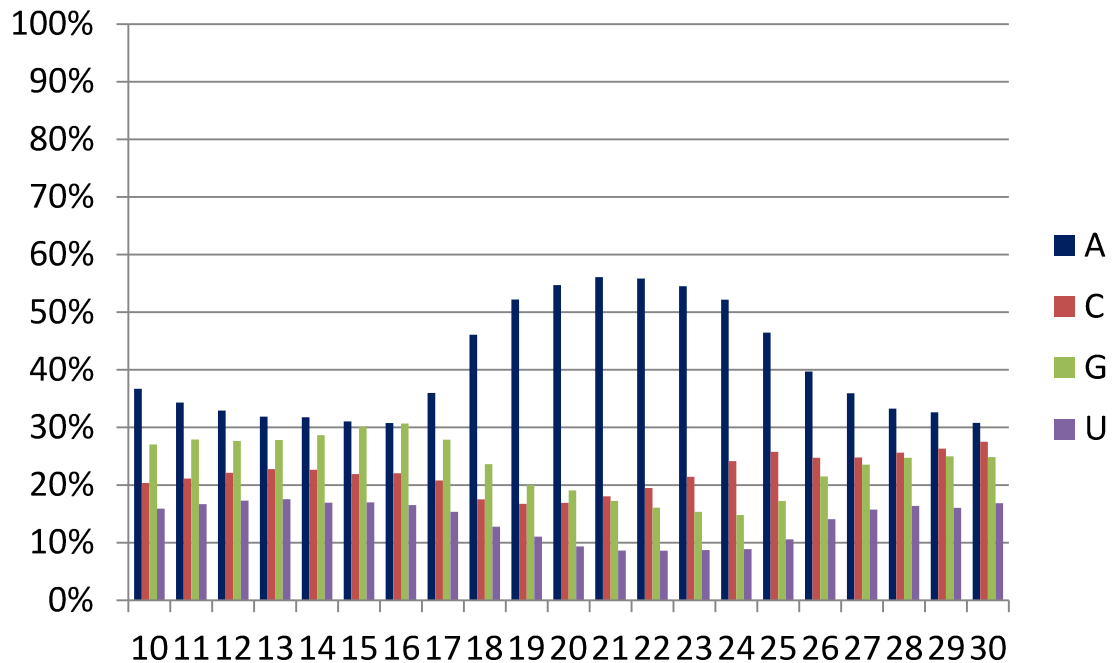
There is an overall A preference as last nucleotide, in reads with length 17-29

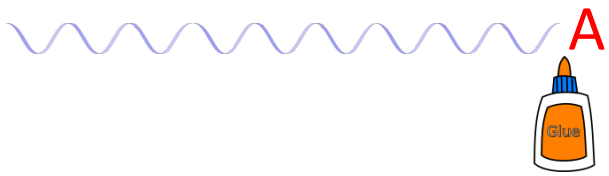
Again, in protein-coding genes and transposable elements



There is an overall A preference as last nucleotide, in reads with length 17-29

Again, in protein-coding genes and transposable elements



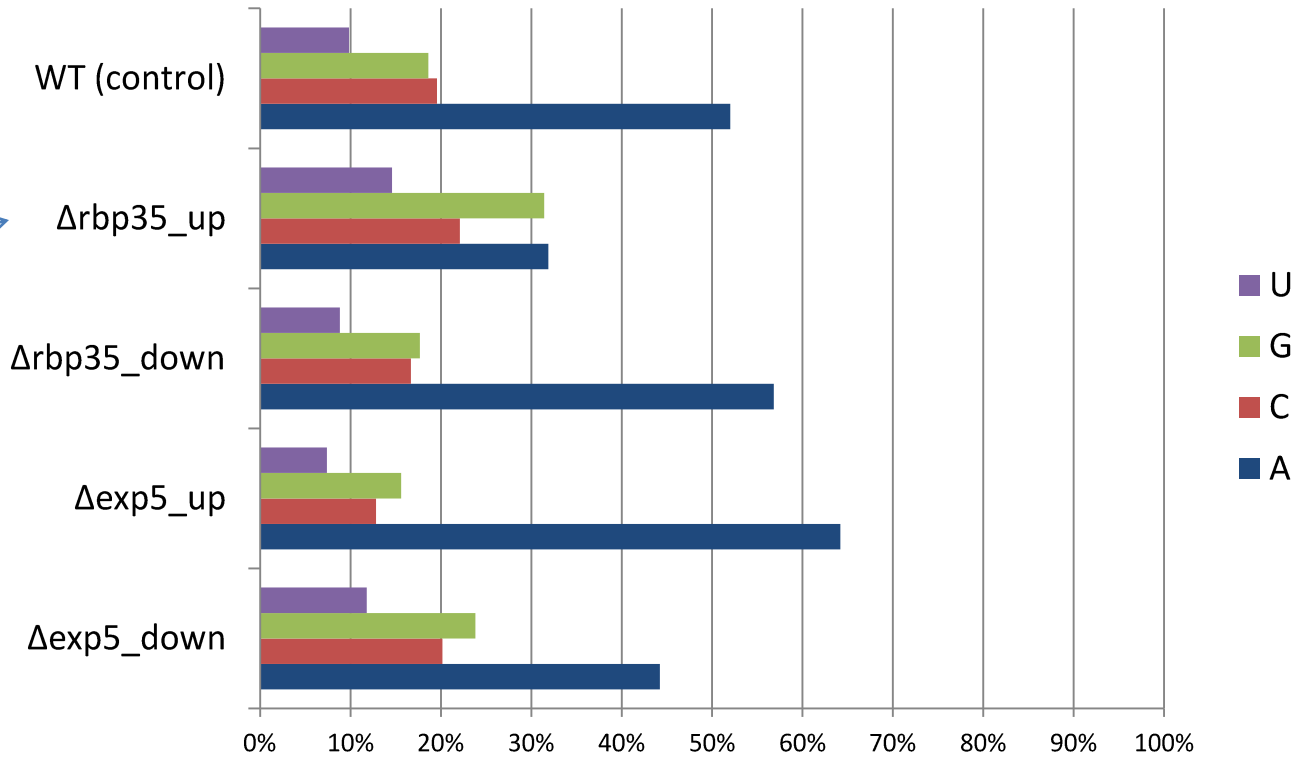


But it is a post-transcriptional modification!

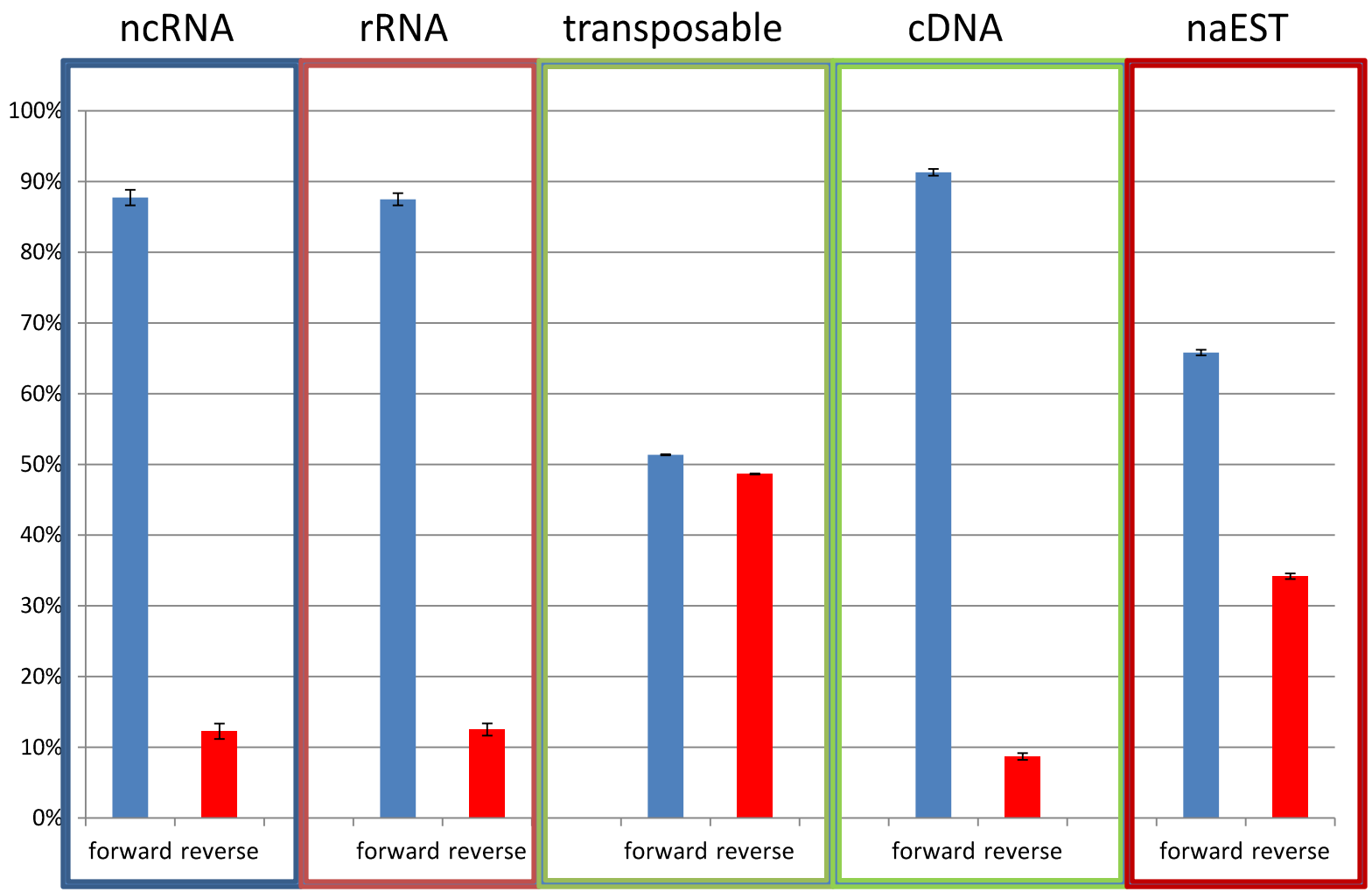
Also found in N.crassa:
Zhihong Xue, Haiyan Yuan, Jinhu Guo,* and Yi Liu† 2013

Last adenine preference is variable among mutants

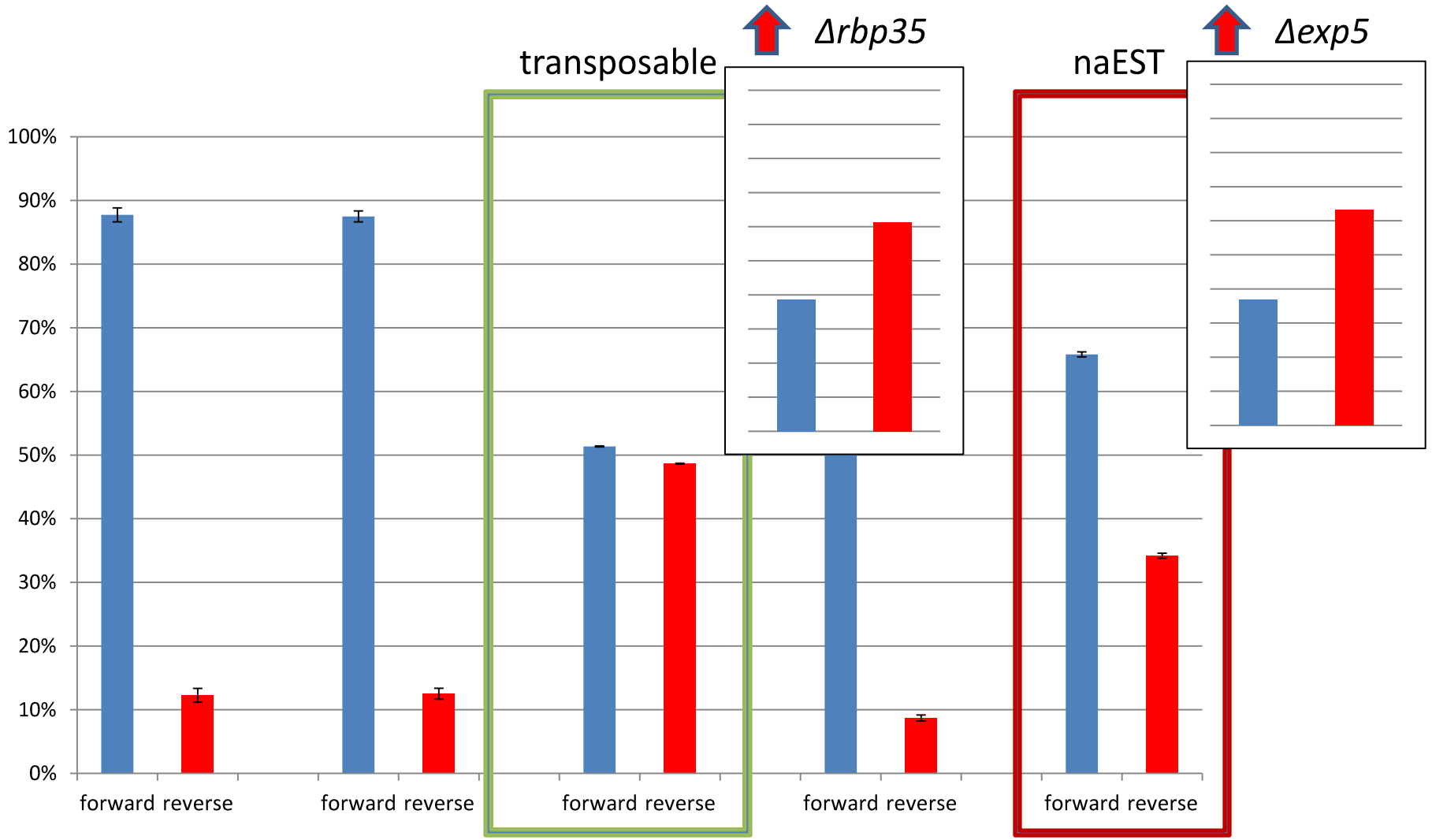
Maybe that's why it aligns better!



Transposable elements and naEST present a high degree of antisense transcription…



…incremented in the mutants



A new class of protein-coding related small RNA is found

A new class of protein-coding related small RNA is found

Unknown smallRNA!

At least 12 protein-coding genes

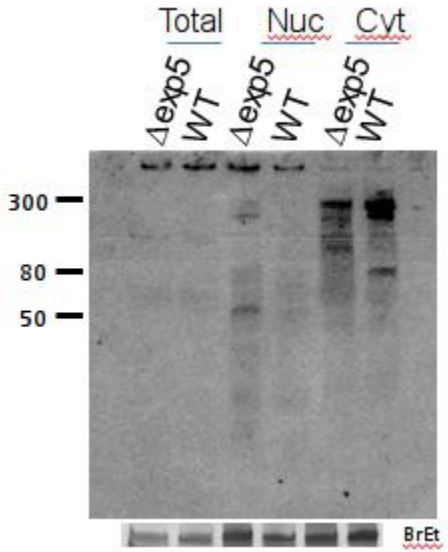
Same situation conserved in other filamentous fungi!
(but not yeasts)

A new class of protein-coding related small RNA is found

Unknown smallRNA!

At least 12 protein-coding genes!

Same situation conserved in other filamentous fungi! (but not yeasts)



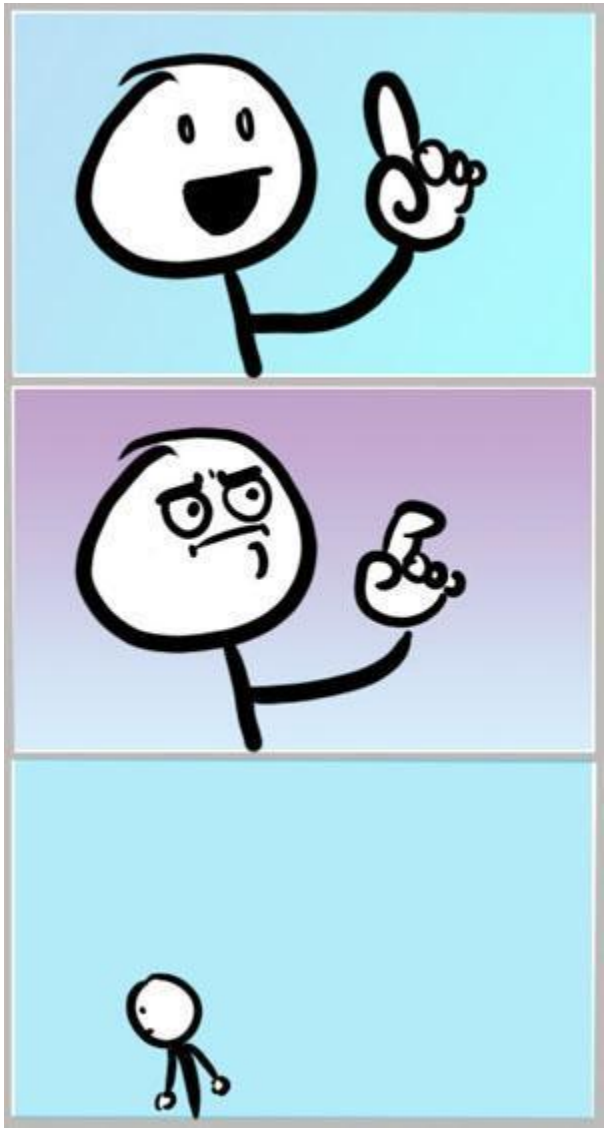
Part 2 - Conclusions

- The small RNAs world of *M. oryzae* is mostly unknown and heterogeneous
- The mutants affect different populations of small RNAs in different ways
- A (maybe) new class of small RNA has been discovered

Acknowledgements

CBGP
UPM-INIA

POLITÉCNICA
"Ingeniamos el futuro"

CAMPUS
DE EXCELENCIA
INTERNACIONAL

Questions?