

ncRNAs in yeast introns

Sam Giffiths-Jones, Daniela Delneri

Katarzyna Hooks

University of Bordeaux Twitter: @KBHooks



Introns and evolution



24 July 2018

Reasons for intron retention



- Evolutionary remains
- Functional:
 - Splicing
 - Sequence of intron
 - Structure of intron
 - Non-coding RNA genes

Intron evolution in Saccharomycotina



• WGD

- 250 orthologous introns in 20 species
- Perfect intron losses

 (gene replaced with its mRNA by homologous recombination)
- Some evidence for micro-homology mediated loss

Hooks et al., Genome biology and evolution (2014)

Intronic ncRNA (*in trans*)



24 July 2018



Intron structure (in cis)

Hooks et al., Genome biology and evolution (2014)



Approach



Computational screen





- 14 intersecting predictions
- + 2 paralogs
- + 3 other high-scoring

• conserved RNA structures up to *Candida* sp.

 only two from 19 RNA structures predicted in introns have not been shown to have function or a phenotype when deleted



mRNA intron ncRNA spliced -++

snoRNAs putative ncRNA

Hooks et al., Genetics (2016)





Hooks et al., Genetics (2016)



RNA structure and introns

- 1. *HAC1 / XBP1*
- 2. GLC7
- 3. RPL28

HAC1



Filipiowicz, EMBO J (2014)

HAC1



Hooks & Griffiths-Jones, RNA Biol (2011)

- 3' splice site

H3

7 bp Arthropoda

50

0.0

Conservation

1.0

60

HAC1 conservation

| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 | 130 |
|----------------|--|------------------------------------|-------------------|---|--|------------------------------|---------------|--|----------------------------------|------------------------------------|----------------------------------|---------------------|-------------|
| B_floridae | G A U C | C A G G U C - | C U | C U G A A U G | C | C – A G – – – – | A C A A U - G | – – U G – C A C C U <mark>C</mark> U G (| A G C A G G G G - | с <mark>а</mark> – – – – – – – – – | G A U C U | - G G A U C | |
| _rerio | – – – – – – – – G <mark>U</mark> G A C – – | C G G G U C - | U U | C U G A G U (| <mark>C C G C A G C -</mark> A C U | C – A G – – – – | G C U A C - G | – – U G – U G C C U <mark>C</mark> C <mark>G (</mark> | A G C A G G U G - | с <mark>а</mark> – – – – – – – – – | <mark>G G</mark> C C C | A G C A G U C C C - | |
| A carolinensis | U C C C C G G - C | | G G | C U G A G U G | C | U – G G – – – – | C C U A A - G | - | AGAAGGGA- | C A | GGCCC | AGCACCGGGG | G – – – – – |
| H sapiens | – – – – – – – – G <mark>U</mark> G G C – – | C G G G U C - | U G | C U G A G U G | C | C – A G – – – – | A C U A C - G | – – U G – C A C C U <mark>C</mark> U G (| A G C A G G U G - | C A | GGCCC | A G U U G U C A C - | |
| C_teleta | – – – – – – G A G <mark>U</mark> G U – – – | G U G U C - | | – – – A <mark>G A</mark> G <mark>U (</mark> | <mark>C U G C A G</mark> U – A C U | C - U | C A A G U A | - | C | | A A G G C A C | ACACU' | C |
| H_robusta | G C A C C G - C A | ACGCACA- | C C | C – A C <mark>G</mark> G G U (| <mark>C U G C A G C</mark> – – C C | C – G U G – – – | U U G | – – A A – C A <mark>C C C C C G (</mark> | C A G G G G G G A G - 1 | u u | – – – G G U <mark>G</mark> U G U | U G G U G | C |
| L_gigantea | U G G A G | A C G A A A - | U C C C A | C A G A U (| <mark>C U G C A G C</mark> – A U C | U - G | C | - | C | с <mark>а</mark> – – – – – – – – – | G A U U U C G | A C U C C | A |
| A_californica | – – – – – – – G U <mark>U</mark> U – – – – | G U G U G U - | U U G U U C | – – – C A G G <mark>U (</mark> | <mark>C C G C A G C</mark> – A C C | U - G | C | - | AGAAGGAG- | с <mark>а</mark> – – – – – – – – – | G G U U C A | G A A U C | |
| C_elegans | G A U C G C | C G U G C C - | U U | – – – U G A A U (| C A G C A G C – A U U | C – A – – – – – | U U A A | – – <mark>U G</mark> A G C <mark>C U – C</mark> A <mark>G (</mark> | AGU-GGGAA | C A | – – – – – <mark>G G</mark> C – C | CGAU | C |
| P_pacificus | U U G U C A | CACCC- | C U | – – – <mark>UG</mark> GAU | C A G C C G C – A U C | C – A – – – – – | U U U G | – – <mark>U G</mark> A A C U C – <mark>C</mark> A <mark>G (</mark> | AGA-GAGCU | C A | – – – – – <mark>G G</mark> G U G | U G G C | U - A |
| B_malayi | C A U C A | | C C U | U G G A U G | C U G C C G A – A U C | C – A – – – – – | U U C G | – – C G G U C C C – <mark>C</mark> A G (| C A G A – G G G A A | с <mark>а</mark> – – – – – – – – – | – – – – – <mark>G G</mark> C G G | UGAU | G |
| C briggsae | – – – – – – G G A <mark>U</mark> C – – – A | ACGAGCC- | U | U G A A U (| C A G C A G A – A U U | C – A – – – – – | U U A A | – – U G C A C C C – <mark>C</mark> A G (| AGA-GGG-A | с <mark>а</mark> – – – – – – – – – | – – – – A <mark>G G</mark> C – C | AACGGAUU | C |
| D_melanogaster | – – – – G G C U G <mark>U</mark> G – – – – | C G U C C A - | C C A A C | – C U <mark>U G</mark> G A U G | <mark>C U G C A G C</mark> – A U C | C – A A A G – – | C | – – U G – A C C C U <mark>C</mark> U <mark>G</mark> (| C | u <mark>a</mark> – – – – – – C a | A C A G G U <mark>G</mark> G A C | A C A C A | G - U C |
| C_pipens | – – – – G G U U G <mark>U</mark> G – – – – | U U U C C A - | - C C C U A C U A | – – – <mark>– G</mark> G A <mark>U</mark> (| <mark>C A G C A G C – A</mark> U C | C – – – – – – – – | G C C C A | – – A A – C C <mark>C C U C U G (</mark> | C | U U – – – – A C U U | UCAA – U <mark>G</mark> GAA | G C A C A | A – A C – – |
| I_castaneum | – – – – G A A U G <mark>U</mark> G – – – – | A – – – – G A G U G – | C A A | – – G <mark>U G</mark> G A C <mark>(</mark> | C A G C A G U – G U C | <mark>C</mark> – – C – – – – | | – – <u>– –</u> – U U <mark>C U U C</mark> G <mark>G (</mark> | C | | – – – – G C <mark>G</mark> C U C | A C A C A | U – U C – – |
| D_pulex | – – – – G G A U G <mark>U</mark> – – – – – | A U C C | U | G – U <mark>U</mark> A A G <mark>U (</mark> | C A G C C G C – A C U | U – A U U A – – | A | – – U G – G A C C U <mark>C</mark> U <mark>G</mark> (| C | C A | – – – – – – <mark>G</mark> GGU | U A C A | U - C C |
| B_cinerea | – – – – – – – G U <mark>U</mark> C U U – – | – – – – U C G <mark>G</mark> U – – | - A U G A C A | C – A A C A U <mark>U</mark> (| <mark>C U G C U G C – A</mark> A U | G – U U G U G – | C | – – – – – – – G A C <mark>C U G (</mark> | A G U G U C – – – | A | <u>A</u> U C G | – A – – A A G U A C | |
| A_nidulans | U G A G U C C | C C C G A U U - | U G A C A | C – A A C A U C 🤇 | <mark>C U G C A G C</mark> – G A U | G – U U G U G – | C | – – – – – – – G A C <mark>C U G (</mark> | A G U G U C – – – | | – – – – – A <mark>G</mark> U C G | - G - C G G G C U C | G |
| C_albicans | - U C C A A C U A A U | C A U U C U - | A U A G | – – C <mark>U G</mark> A U <mark>U</mark> U | J A <mark>G C A G C</mark> – A A U | C – A G U C U – | - A | – – <u>– –</u> – – – – G C <mark>C</mark> A <mark>G (</mark> | A G A G C | C A A U - | – – – – – A <mark>G</mark> A A U | C A U U U C A · | G – U U G G |
| 5_cerevisiae | C A A U U C A A U U G - A | U C – – – U U G A – – | C A A U U G G | C – G <mark>U</mark> A A U C <mark>(</mark> | C A <mark>G C</mark> C <mark>G</mark> U – G A U | U – A C G – – – | A U U G G | C U <mark>U G</mark> U A C U G <mark>U C</mark> C <mark>G</mark> 2 | A <mark>A G</mark> C G C A G U – | C A G G U – – – – – | U U G A | A – UUCAUUUG | A – A U U G |
| | | | | | | • | | | | | | | • |

| | 10 | | 20 | 30 | | 40 | 50 | | 60 | 70 | | 80 | 9 | 0 | 100 | 1 | 110 | 1 | 120 | 130 |
|----------------|---------------------------|-------------|-------------|-------------------|-----------|-------------|-----------------|-----------|-------------|-------------------|-------------|-------------|---------|-------------|-------------|---------------|------------------------------|-------------------------|-----------------|--------------------|
| B_floridae | G A U C | | CAGGUC | C U | 0 | CUGAAU | CCGCAGC- | AUUC | – A G – – – | – A C A A U – G – | U G | - C A C C U | CUGCAGC | AGGGG | - C A | | G A U C U | I - G - | - G A U C | |
| D_rerio | G U G A C | | CGGGUC | U U | 0 | CUGAGU | CCGCAGC- | ACUC | – A G – – – | – G C U A C – G – | U G | – U G C C U | CCGCAGC | AGGUG | - C A | | – – – <mark>G G C C C</mark> | AGC | A G U C C C | |
| A carolinensis | U C C C G G | – C – – – – | UGGGCC | – – – G G – – – – | 0 | CUGAGU | CCGCAGC- | ACUU | – G G – – – | – C C U A A – G – | U G | - u u c c u | CUGCAGA | AGGGA | - C A | | – – – <mark>G G C C C</mark> | AGC | ACCGGGGG | |
| H_sapiens | G U G G C | | C G G G U C | – – – U G – – – – | 0 | CUGAGU | CCGCAGC- | ACUC | – A G – – – | – A C U A C – G – | U G | - C A C C U | CUGCAGC | AGGUG | - C A | | – – – <mark>GGCCC</mark> | AGU | U G U C A C | |
| C_teleta | G A G U G U - | | – G U G U C | | | - AGAGU | CUGCAGU- | ACUC | - U | – C A A G U A – – | | - c u c c u | CUGCCGA | AGGAG | | | – A A <mark>G G C A C</mark> | : - · | A C A C U C | |
| H_robusta | G C A C C G - | C A A C | - G C A C A | – C C – – – – – – | C - I | ACGGGU | CUGCAGC- | - C C C - | - G U G | - U U G | A A | - C A C C C | CCGCAGG | GGGAG | - U U | | – G G <mark>U G U G U</mark> | / - · | – – U G G U G C | |
| L_gigantea | U G G A G | – – A – – – | – CGAAA | – U C C C A – – – | | - C A G A U | CUGCAGC- | AUCU | - G | - C | U G | - u u c c u | CUGCCGA | AGGAG | - C A | | – G A <mark>U U U C G</mark> | A – – • | C U C C A | |
| A_californica | G U U U | – – G – – – | – U G U G U | – U U G U U C – – | | - C A G G U | CCGCAGC- | ACCU | - G | – C – – – – – – – | U G | - บ บ С С บ | CUGCAGA | AGGAG | - C A | | G G U U C A | G | A A U C - | |
| C_elegans | G A U C G | – C C – – – | – G U G C C | – U U – – – – – – | | - U G A A U | CAGCAGC - | AUUC | – A – – – – | – U U A A – – – – | U G | AGCCU- | CAGCAGU | - G G G A | A C A | | – – – <mark>G G C – C</mark> | : | C G A U C | |
| P_pacificus | U U G U C A | | - C A C C C | - C U | | - U G G A U | CAGCCGC- | AUCC | – A – – – – | – U U U G – – – – | U G | AACUC - | CAGCAGA | – GAGC | U C A – – – | | – – – <mark>G G G U G</mark> | ; . | – – – U G G C U | - A |
| B malayi | C A U C A | | – U C G U C | – C C U – – – – – | | - U G G A U | CUGCCGA- | AUCC | – A – – – – | – U U C G – – – – | – – – – C G | G U C C C – | CAGCAGA | - G G G A | A C A | | – – – <mark>G G C G G</mark> | ; . | U G A U G | |
| C_briggsae | G G A U C | – A A C – – | - GAGCC | – U – – – – – – – | | - U G A A U | CAGCAGA - | AUUC | – A – – – – | – U U A A – – – – | U G | CACCC- | CAGCAGA | – G G G – 2 | A C A | | A G G C - C | AAC | G – – G A U U C | |
| D_melanogaster | – – – – G G C U G U G – – | – – C – – – | – GUCCA | – C C A A – – – – | – C – C t | JUGGAU | CUGCAGC- | AUCC | – A A A G – | | C – – – U G | – A C C C U | СИGССGС | AGGGU | A U A – – – | – – – C A A C | AGG <mark>UGGAC</mark> | A | C A C A G | - U C |
| C_pipens | – – – – G G U U G U G – – | U | - U U C C A | – C C C U A C U A | | G G A U | CAGCAGC - | AUCC | | – G C C C A – – – | A A | - c c c c u | СИGССGС | A G G G - 2 | A U U | – A C U U U C | AA - UGGAA | G | C A C A A | - A C |
| T_castaneum | G A A U G U G | – – A – – – | – GAGUG | - C A A | 0 | GUGGAC | CAGCAGU - | GUCC | – – C – – – | | | - | СGGCCGC | AAGGA | | | – – G <mark>C G C U C</mark> | A | C A C A U | - U C |
| D_pulex | G G A U G U | | - A U C C - | | – U G – U | JUAAGU | CAGCCGC- | ACUU | – A U U A – | A | U G | - G A C C U | СИGССGС | AGGGA | - C A | | <mark>- G G G U</mark> | U U | A C A U | - C C |
| B_cinerea | G U U C U U | | UCGGU- | – – A U G A – – – | C A C - 1 | AACAUU | C U G C U G C - | AAUG | – U U G U G | C | | – – – G A C | CUGCAGU | G U C – – | | – A – – – – – | A U C G | - A - | – AAGUAC – | |
| A_nidulans | U G A G U C C | – – C – – – | CCGAUU | – – – U G A – – – | C A C - 1 | AACAUC | CUGCAGC- | GAUG | – U U G U G | C | | – – – G A C | CUGCAGU | G U C – – | | | A G U C G | ; – G – <i>i</i> | C | |
| C_albicans | - U C C A A C U A A U | | CAUUCU | – – – A U – – – – | AG 0 | CUGAUU | UAGCAGC - | AAUC | – A G U C U | A | | – – – G C | CAGCAGA | G C – – – | | C A A U | A G A A U | 1 C | – AUUUCAG | – U U G G <i>I</i> |
| S_cerevisiae | – – C A A U U C A A U U G | – A U C – – | - U U G A - | – C A A U U G G – | C - C | GUAAUC | CAGCCGU- | GAUU | – A C G – – | – A U U – – – – – | - G C U U G | UACUGU | CCGAAGC | GCAGU | CAGGU | | | י ט – א | UCAUUUGA | - A U U G - |
| | | | | | | | | | | | | | | | | | | | | |

secondary struct

secondary struct

RNA structure and introns



- 2. GLC7
- 3. RPL28

GLC7



GLC7 ncRNA indicated by RT-PCR and the Northern

GLC7 codes for a crucial protein phosphatase catalytic subunit

GLC7 phenotype



ncRNA deletion mutant is sensitive to salt stress

ncRNA deletion decreases the *GLC7* expression

no rescued by ncRNA expression

Hooks et al., Genetics (2016)

RNA structure and introns



3. *RPL28*

RPL28 structure



• Similar hairpin structures in premRNAs of *RPS14B*, *RPL30* and *RPL28*

RPL28 structure



• Similar hairpin structures in premRNAs of *RPS14B*, *RPL30* and *RPL28*

 pre-mRNA structure is recognized by ribosomal proteins with lower affinity than their rRNA targets

Negative autoregulatory loop

RPL28 expression



100

 10^{1}

 10^{-1}

 10^{3}

 10^{2} S. cerevisiae intron reads [RPKM] 10^{4}

RPL28 intron function

Mol Gen Genet (1986) 203:300-304

Intron mutations that affect the splicing efficiency of the CYH2 gene of Saccharomyces cerevisiae

Ulrike Swida, Eduardo Thüroff, and Norbert F. Käufer*



*aq*667.

RPL28 intron function

Mol Gen Genet (1986) 203:300-304

Intron mutations that affect the splicing efficiency of the CYH2 gene of Saccharomyces cerevisiae



d9 661

RNA structure and introns



Conclusions



Photo by: Rainis Venta, BY-SA 3.0

- Intron positions are perfectly conserved, sequences are frequently conserved
- Introns are expressed and have a life after splicing
- Introns can perform their function by RNA structures
- Intronic RNAs regulate the transcription of their host genes

Acknowlegments



Sam Giffiths-Jones

Antonio Marco

Daniela Delneri

Samina Naseeb

Steven Parker



The University of Manchester



Maureen O'Malley

Macha Nikolski



Elena Rivas Eric Westhof