Evolution of Yeast Noncoding RNAs Reveals an Alternative Mechanism for Widespread Intron Loss

Quinn M. Mitrovich *et al.*

Science **330**, 838 (2010);
DOI: 10.1126/science.1194554

difference in fluorescence signal between the activated and kirromycin-stalled states (*34*).

This structure of 70*S*–tRNA–EF-Tu–GDPCP provides an atomic-resolution model of a translational GTPase in its activated state. Codon recognition leads to a series of conformational changes in the 30*S* subunit, tRNA, and EF-Tu that position EF-Tu for GTP hydrolysis. GTPase activation does not require a large opening of the "hydrophobic gate," but instead requires the positioning of the catalytic histidine into the active site by the SRL residue A2662. The high level of sequence and structural conservation between all translational GTPases suggests that although each factor recognizes a distinct ribosomal state, each must bind in such a way that the SRL interacts with its catalytic histidine. Therefore, GTPase activation by the SRL is the universal mechanism for triggering GTP hydrolysis on the ribosome across all kingdoms of life.

References and Notes

1. T. Margus, M. Remm, T. Tenson, *BMC Genomics* **8**, 15 (2007).
2. T. E. Dever, M. J. Glynias, W. C. Merrick, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1814 (1987).
3. D. Moazed, J. M. Robertson, H. F. Noller, *Nature* **334**, 362 (1988).
4. L. Holmberg, O. Nygård, *Biochemistry* **33**, 15159 (1994).
5. J. F. Eccleston, M. R. Webb, *J. Biol. Chem.* **257**, 5046 (1982).
6. H. Berchtold *et al.*, *Nature* **365**, 126 (1993).
7. M. Kjeldgaard, P. Nissen, S. Thirup, J. Nyborg, *Structure* **1**, 35 (1993).
8. E. Jacquet, A. Parmeggiani, *EMBO J.* **7**, 2861 (1988).
9. T. Pape, W. Wintermeyer, M. V. Rodnina, *EMBO J.* **17**, 7490 (1998).
10. J. R. Mesters, A. P. Potapov, J. M. de Graaf, B. Kraal, *J. Mol. Biol.* **242**, 644 (1994).
11. T. P. Hausner, J. Atmadja, K. H. Nierhaus, *Biochimie* **69**, 911 (1987).
12. H. R. Bourne, D. A. Sanders, F. McCormick, *Nature* **349**, 117 (1991).
13. D. Mohr, W. Wintermeyer, M. V. Rodnina, *Biochemistry* **41**, 12520 (2002).
14. L. Vogeley, G. J. Palm, J. R. Mesters, R. Hilgenfeld, *J. Biol. Chem.* **276**, 17149 (2001).
15. T. M. Schmeing *et al.*, *Science* **326**, 688 (2009).
16. J. Sengupta *et al.*, *J. Mol. Biol.* **382**, 179 (2008).
17. E. Villa *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1063 (2009).
18. J. C. Schuette *et al.*, *EMBO J.* **28**, 755 (2009).
19. P. Nissen *et al.*, *Science* **270**, 1464 (1995).
20. M. Valle *et al.*, *EMBO J.* **21**, 3557 (2002).
21. D. J. Taylor *et al.*, *EMBO J.* **26**, 2421 (2007).
22. C. Knudsen, H. J. Wieden, M. V. Rodnina, *J. Biol. Chem.* **276**, 22183 (2001).
23. D. M. Blow, J. J. Birktoft, B. S. Hartley, *Nature* **221**, 337 (1969).
24. M. Selmer *et al.*, *Science* **313**, 1935 (2006).
25. N. Bilgin, M. Ehrenberg, *J. Mol. Biol.* **235**, 813 (1994).
26. F. Rambelli *et al.*, *Biochem. J.* **259**, 307 (1989).
27. J. M. Ogle *et al.*, *Science* **292**, 897 (2001).
28. R. Mittal, M. R. Ahmadian, R. S. Goody, A. Wittinghofer, *Science* **273**, 115 (1996).
29. K. Rittinger, P. A. Walker, J. F. Eccleston, S. J. Smerdon, S. J. Gamblin, *Nature* **389**, 758 (1997).
30. J. J. Tesmer, D. M. Berman, A. G. Gilman, S. R. Sprang, *Cell* **89**, 251 (1997).
31. M. Diaconu *et al.*, *Cell* **121**, 991 (2005).
32. H. J. Wieden, W. Wintermeyer, M. V. Rodnina, *J. Mol. Evol.* **52**, 129 (2001).
33. M. Remacha *et al.*, *Mol. Cell. Biol.* **15**, 4754 (1995).
34. M. V. Rodnina, R. Fricke, W. Wintermeyer, *Biochemistry* **33**, 12267 (1994).

35. We thank R. Green for reagents and D. de Sanctis at the European Synchrotron Radiation Facility for facilitating data collection. This work was supported by the Medical Research Council UK, the Wellcome Trust, the Agouron Institute, and the Louis-Jeantet Foundation. R.M.V. is supported by a Gates-Cambridge scholarship, and T.M.S. by the Human Frontier Science Program and Emmanuel College. V.R. is on the Scientific Advisory Board and holds stock options in Rib-X pharmaceuticals. Transfer of *Thermus thermophilis* ribosome strain T.th. HB8-MRC-MSAW1 requires a Materials Transfer Agreement with the MRC. Coordinates and structure factors have been deposited at the Protein Data Bank with accession codes 2xqd and 2xqe.

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6005/835/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References

29 June 2010; accepted 3 September 2010
10.1126/science.1194460

Evolution of Yeast Noncoding RNAs Reveals an Alternative Mechanism for Widespread Intron Loss

Quinn M. Mitrovich,[1,2]* Brian B. Tuch,[1,3]*† Francisco M. De La Vega,[3] Christine Guthrie,[2]‡ Alexander D. Johnson[1,2]‡

The evolutionary forces responsible for intron loss are unresolved. Whereas research has focused on protein-coding genes, here we analyze noncoding small nucleolar RNA (snoRNA) genes in which introns, rather than exons, are typically the functional elements. Within the yeast lineage exemplified by the human pathogen *Candida albicans*, we find—through deep RNA sequencing and genome-wide annotation of splice junctions—extreme compaction and loss of associated exons, but retention of snoRNAs within introns. In the *Saccharomyces* yeast lineage, however, we find it is the introns that have been lost through widespread degeneration of splicing signals. This intron loss, perhaps facilitated by innovations in snoRNA processing, is distinct from that observed in protein-coding genes with respect to both mechanism and evolutionary timing.

In eukaryotes, protein-coding genes are frequently interrupted by introns, which must be precisely removed from RNA transcripts by the nuclear spliceosome (*1*). Over evolutionary time scales, the presence of introns is dynamic, with intron gain and loss rates varying substantially across eukaryotic lineages (*2*–*4*). The mechanisms of intron gain and loss speak to questions about both the origins of introns and the markedly different intron-exon patterns observed across eukaryotes (*5*)—for example, whether spliceosomal introns arose within eukaryotes ("introns late"), within an ancestor of both prokaryotes and eukaryotes ("introns early"), or even before the emergence of protein-coding genes ("introns first") (*6*). The last two hypotheses depend on the feasibility of comprehensive intron loss within both the prokaryotic and archaeal lineages, whose modern representatives lack spliceosomal introns. Within eukaryotes, the hemiascomycetous yeasts show substantial intron loss, with modern species like *Saccharomyces cerevisiae* and *Candida albicans* devoid of introns in >90% of their genes (*7*). A postulated mechanism for this loss is reverse transcription of spliced RNA, followed by homologous DNA recombination that replaces the intron-containing genomic sequence with the intronless copy (*8*). Previous studies of intron loss have focused on protein-coding genes, however, and are therefore likely to be biased toward mecha-

[1]Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94143–2200, USA. [2]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143–2200, USA. [3]Genetic Systems Division, Research and Development, Life Technologies, Foster City, CA 94404, USA.

*These authors contributed equally to this work.
†Present address: Genome Analysis Unit, Amgen, South San Francisco, CA 94080, USA.
‡To whom correspondence should be addressed. E-mail: christineguthrie@gmail.com (C.G.); ajohnson@cgl.ucsf.edu (A.D.J.)

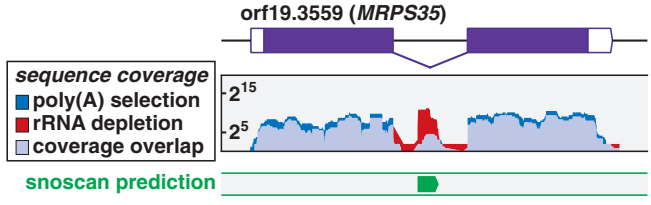
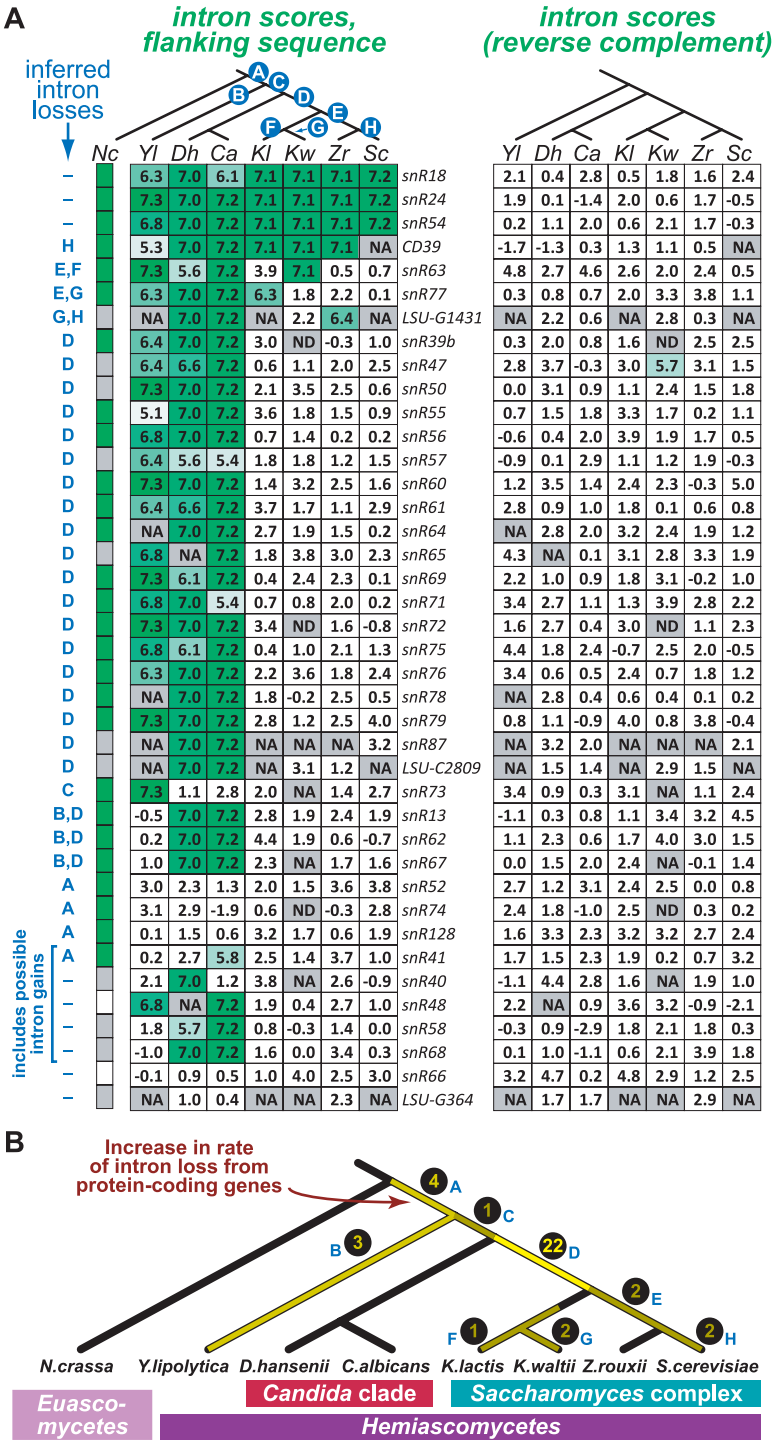


Fig. 1. Sequence library comparisons reveal noncoding RNAs. RNA sequence data are shown for *MRPS35*, a representative protein-coding gene that hosts a snoRNA within its intron. Nonadenylated RNAs, such as mature snoRNAs, are enriched in the rRNA-depleted library relative to the poly(A)-selected library. Sequence depth is represented on a $\log_2$ axis. (Bottom track) One of 1706 lower-confidence snoRNA predictions generated using the snoscan algorithm (*22*).



		Yl	Dh	Ca	Kl	Kw	Zr	Sc		Yl	Dh	Ca	Kl	Kw	Zr	Sc
–		6.3	7.0	6.1	7.1	7.1	7.1	7.2	*snR18*	2.1	0.4	2.8	0.5	1.8	1.6	2.4
–		7.3	7.0	7.2	7.1	7.1	7.1	7.2	*snR24*	1.9	0.1	-1.4	2.0	0.6	1.7	-0.5
–		6.8	7.0	7.2	7.1	7.1	7.1	7.2	*snR54*	0.2	1.1	2.0	0.6	2.1	1.7	-0.3
H		5.3	7.0	7.2	7.1	7.1	7.1	NA	*CD39*	-1.7	-1.3	0.3	1.3	1.1	0.5	NA
E,F		7.3	5.6	7.2	3.9	7.1	0.5	0.7	*snR63*	4.8	2.7	4.6	2.6	2.0	2.4	0.5
E,G		6.3	7.0	7.2	6.3	1.8	2.2	0.1	*snR77*	0.3	0.8	0.7	2.0	3.3	3.8	1.1
G,H		NA	7.0	7.2	NA	2.2	6.4	NA	*LSU-G1431*	NA	2.2	0.6	NA	2.8	0.3	NA
D		6.4	7.0	7.2	3.0	ND	-0.3	1.0	*snR39b*	0.3	2.0	0.8	1.6	ND	2.5	2.5
D		6.4	6.6	7.2	0.6	1.1	2.0	2.5	*snR47*	2.8	3.7	-0.3	3.0	5.7	3.1	1.5
D		7.3	7.0	7.2	2.1	3.5	2.5	0.6	*snR50*	0.0	3.1	0.9	1.1	2.4	1.5	1.8
D		5.1	7.0	7.2	3.6	1.8	1.5	0.9	*snR55*	0.7	1.5	1.8	3.3	1.7	0.2	1.1
D		6.8	7.0	7.2	0.7	1.4	0.2	0.2	*snR56*	-0.6	0.4	2.0	3.9	1.9	1.7	0.5
D		6.4	5.6	5.4	1.8	1.8	1.2	1.5	*snR57*	-0.9	0.1	2.9	1.1	1.2	1.9	-0.3
D		7.3	7.0	7.2	1.4	3.2	2.5	1.6	*snR60*	1.2	3.5	1.4	2.4	2.3	-0.3	5.0
D		6.4	6.6	7.2	3.7	1.7	1.1	2.9	*snR61*	2.8	0.9	1.0	1.8	0.1	0.6	0.8
D		NA	7.0	7.2	2.7	1.9	1.5	0.2	*snR64*	NA	2.8	2.0	3.2	2.4	1.9	1.2
D		6.8	NA	7.2	1.8	3.8	3.0	2.3	*snR65*	4.3	NA	0.1	3.1	2.8	3.3	1.9
D		7.3	6.1	7.2	0.4	2.4	2.3	0.1	*snR69*	2.2	1.0	0.9	1.8	3.1	-0.2	1.0
D		6.8	7.0	5.4	0.7	0.8	2.0	0.2	*snR71*	3.4	2.7	1.1	1.3	3.9	2.8	2.2
D		7.3	7.0	7.2	3.4	ND	1.6	-0.8	*snR72*	1.6	2.7	0.4	3.0	ND	1.1	2.3
D		6.8	6.1	7.2	0.4	1.0	2.1	1.3	*snR75*	4.4	1.8	2.4	-0.7	2.5	2.0	-0.5
D		6.3	7.0	7.2	2.2	3.6	1.8	2.4	*snR76*	3.4	0.6	0.5	2.4	0.7	1.8	1.2
D		NA	7.0	7.2	1.8	-0.2	2.5	0.5	*snR78*	NA	2.8	0.4	0.6	0.4	0.1	0.2
D		7.3	7.0	7.2	2.8	1.2	2.5	4.0	*snR79*	0.8	1.1	-0.9	4.0	0.8	3.8	-0.4
D		NA	7.0	7.2	NA	NA	NA	3.2	*snR87*	NA	3.2	2.0	NA	NA	NA	2.1
D		NA	7.0	7.2	NA	3.1	1.2	NA	*LSU-C2809*	NA	1.5	1.4	NA	2.9	1.5	NA
C		7.3	1.1	2.8	2.0	NA	1.4	2.7	*snR73*	3.4	0.9	0.3	3.1	NA	1.1	2.4
B,D		-0.5	7.0	7.2	2.8	1.9	2.4	1.9	*snR13*	-1.1	0.3	0.8	1.1	3.4	3.2	4.5
B,D		0.2	7.0	7.2	4.4	1.9	0.6	-0.7	*snR62*	1.1	2.3	0.6	1.7	4.0	3.0	1.5
B,D		1.0	7.0	7.2	2.3	NA	1.7	1.6	*snR67*	0.0	1.5	2.0	2.4	NA	-0.1	1.4
A		3.0	2.3	1.3	2.0	1.5	3.6	3.8	*snR52*	2.7	1.2	3.1	2.4	2.5	0.0	0.8
A		3.1	2.9	-1.9	0.6	ND	-0.3	2.8	*snR74*	2.4	1.8	-1.0	2.5	ND	0.3	0.2
A		0.1	1.5	0.6	3.2	1.7	0.6	1.9	*snR128*	1.6	3.3	2.3	3.2	3.2	2.7	2.4
A		0.2	2.7	5.8	2.5	1.4	3.7	1.0	*snR41*	1.7	1.5	2.3	1.9	0.2	0.7	3.2
–		2.1	7.0	1.2	3.8	NA	2.6	-0.9	*snR40*	-1.1	4.4	2.8	1.6	NA	1.9	1.0
–		6.8	NA	7.2	1.9	0.4	2.7	1.0	*snR48*	2.2	NA	0.9	3.6	3.2	-0.9	-2.1
–		1.8	5.7	7.2	0.8	-0.3	1.4	0.0	*snR58*	-0.3	0.9	-2.9	1.8	2.1	1.8	0.3
–		-1.0	7.0	7.2	1.6	0.0	3.4	0.3	*snR68*	0.1	1.0	-1.1	0.6	2.1	3.9	1.8
–		-0.1	0.9	0.5	1.0	4.0	2.5	3.0	*snR66*	3.2	4.7	0.2	4.8	2.9	1.2	2.5
–		NA	1.0	0.4	NA	NA	2.3	NA	*LSU-G364*	NA	1.7	1.7	NA	NA	2.9	NA

Fig. 2. snoRNA-associated introns were lost in the *Saccharomyces* lineage. (**A**) Intron prediction scores [combined 5′ splice site and branch site species-specific PWM scores (*15*)] for 40 C/D box snoRNA flanking sequences (± 200 nt) in seven representative hemiascomycetes. Each snoRNA (horizontal row) is labeled according to the nomenclature of *S. cerevisiae* (where applicable) or *N. crassa* (*CD39*) or by the predicted *C. albicans* rRNA modification site. Intron scores greater than 5.0 (false-positive rate <0.4%) are shaded green. Inferred intron loss events for each snoRNA, based on parsimony, are indicated on the left and correspond to labeled branches (not drawn to scale) in the phylogeny shown above and in (B). snoRNAs that were not identified or that lack sufficient flanking sequence to score are gray and labeled NA or ND, respectively. Locations of *N. crassa* orthologs (green for intronic, white for exonic) are derived from (*12*). Intron scores for reverse-complements of flanking sequences (right) are provided as a negative control. (**B**) Phylogenetic pattern of snoRNA-associated intron loss. The number of loss events assigned to each branch is indicated in yellow. Species: *N. crassa*, *Y. lipolytica*, *Debaryomyces hansenii*, *C. albicans*, *Kluyveromyces lactis*, *Kluyveromyces waltii*, *Zygosaccharomyces rouxii*, and *S. cerevisiae*.

nisms that lead to precise intron removal. Here, we examine instead the evolution of splicing patterns in yeast noncoding genes.

We performed massively parallel ligation-based sequencing of RNA libraries (RNA-seq) and mapped the resulting ~160 million 50-nucleotide (nt) strand-specific sequence reads to the *C. albicans* genome (*9*). Our data confirm 89% of 421 previously annotated spliceosomal introns (*7*, *10*) (table S1), while correcting or rejecting seven of these annotations (table S2). We also find 68 previously unannotated splice junctions, identifying 15 new introns in protein-coding genes (table S3), 30 examples of alternative splicing [table S4 and supporting online (SOM) text], and 23 new introns in previously unannotated transcripts that lack substantial protein-coding potential (table S5). Analysis of 11 of these spliced, noncoding RNAs revealed that their exons have no apparent function, but that their introns contain C/D box snoRNAs—noncoding RNAs that target modifications to ribosomal RNA (rRNA) (*11*). In the nonhemiascomycetous fungus *Neurospora crassa*, snoRNAs are also generally processed from the introns of non–protein-coding precursors (*12*). This is different, however, from the more closely related hemiascomycete *S. cerevisiae*, where nearly all snoRNAs arise from unspliced primary transcripts and, therefore, require a splicing-independent processing pathway (*13*). This difference between *C. albicans* and *S. cerevisiae* suggests that the transition of snoRNAs from intron sequences to unspliced, dedicated transcripts occurred within the *Saccharomyces* lineage, well after the onset of rapid intron loss from protein-coding genes in the hemiascomycete ancestor.

To trace the evolution of snoRNAs throughout the hemiascomycetes, we first generated 40 high-confidence C/D box snoRNA predictions for *C. albicans* (e.g., Fig. 1). Among these are the 11 identified in our intron analysis (above), as well as the previously identified *snR52* (*14*). We searched for splicing signals in sequences adjacent to snoRNAs, and predict that 33 of the 40 identified *C. albicans* snoRNAs are intronic. This confirms the difference between *C. albicans* and *S. cerevisiae*, as C/D box snoRNAs in the latter species are rarely intronic [6 out of 47 (*13*)].

We next identified orthologous snoRNAs from other sequenced yeasts (Fig. 2), beginning with computational predictions (or, for *S. cerevisiae*, existing annotations), identifying candidate orthologs of our *C. albicans* set based on their predicted rRNA target sites, and finally confirming and refining our predictions by searching for limited primary sequence identity among predicted snoRNAs (*15*). This final refinement identified snoRNAs whose rRNA target sites had changed between species, demonstrating evolutionary plasticity in yeast rRNA methylation sites (fig. S1). We predict that 105 of the 255 analyzed snoRNAs are located within introns (Fig. 2A). Individual species vary considerably: Those in the *Saccharomyces* complex have few intronic snoRNAs (three to five), whereas others have substantially more (23 to 33). The most par-

simonious explanation for these data is massive loss of snoRNA-associated introns, most of which took place in the common ancestor of the *Saccharomyces* complex (Fig. 2B).

The intron loss mechanism proposed for protein-coding genes—retrotransposition of spliced mRNAs—cannot explain the pattern we observe, as it would eliminate the snoRNAs along with the introns. Rather, the introns appear to have been lost through degeneration of their splicing signals, effectively converting them into unspliced exon sequence. In *N. crassa* (*12*) and hemiascomycetes outside the *Saccharomyces* complex, snoRNAs in the *snR72-78* polycistronic cluster are mostly contained within individual introns (Fig. 3A). In *S. cerevisiae*, the genes are arranged identically, but are cotranscribed as a single unspliced precursor, then processed into individual snoRNAs by the RNase III enzyme Rnt1 (*16*). The conservation of genomic synteny among these species strongly suggests intron loss through splice site degeneration ("de-intronization") with *Yarrowia lipolytica* and the *Candida* clade representing an intermediate state of partial intron loss and *S. cerevisiae* representing complete intron loss. The *snR57* cluster similarly supports this idea (below), as do the gene structures of *snR47* and *snR79* (SOM text).

Intron loss through splice-site degeneration would not be expected to occur within protein-coding genes, as this would disrupt the encoded protein. Consistent with this prediction, of the five snoRNAs located in introns within protein-coding regions in *C. albicans*, four remain associated with introns in nearly all the *Saccharomyces* complex species [*snR18*, *snR24*, *snR54*, and *CD39*; (Fig. 2A)]. The fifth, *snR39b*, was likely displaced from its associated coding sequence through a genomic duplication event (fig. S2B). The snoRNA *CD39* is located within the ribosomal protein gene *MRPS35* intron in all but one of the species we analyzed (Fig. 2A). The exception is *S. cerevisiae*, where *MRPS35* has lost its intron, presumably through retrotransposition rather than deintronization. This appears to have eliminated *CD39* from the genome as well: The predicted rRNA methylation site for this snoRNA is unmodified in *S. cerevisiae* (*13*).

In *C. albicans*, one unusual case of splicing may reflect the particular processing requirements imposed by intron-hosted snoRNAs. The snoRNAs *snR57*, *snR55*, and *snR61* are processed from three introns of a single precursor (Fig. 3B). We find that two of the introns (introns 2 and 3) lie entirely within the sequence of an enveloping intron (intron 1). The splicing signals for intron 1 are not present in the primary transcript, but are created upon splicing of introns 2 and 3. Thus, splicing of intron 1 can occur only after introns 2 and 3 have been removed. (See SOM text for similar scenarios in protein-coding genes of other eukaryotes.) Analysis of both the *snR57/55/61* and the *snR72-78* clusters in other species indicates a significant reduction in the sizes of internal exons within the hemiascomycetes, driven perhaps by the same pressures that streamlined their genomes (*17*) and leading ultimately to nested splicing within the *Candida* clade (fig. S3). In animals and fungi, intron-hosted C/D box snoRNAs obey a strict "one-snoRNA-per-intron" rule, a requirement imposed by their exonucleolytic maturation pathways (*18*). Nested splicing of snoRNA host transcripts fulfills this requirement by allowing sequential removal of individual introns despite the absence of intervening exons.

Selective pressures are proposed to have driven intron loss from hemiascomycete protein-coding genes (*19*), and these same pressures may have driven the loss we observe here. The dependence of snoRNAs on splicing for their proper maturation, however, would have imposed a constraint against loss of their associated introns (*18*). This constraint may have been overcome by the innovation of an alternative snoRNA processing pathway in the *Saccharomyces* lineage, where exonic C/D box snoRNAs first undergo endonucleolytic cleavage by the RNase III enzyme Rnt1 (*20*). Nonetheless, some capacity to process exonic snoRNAs must be more ancient, because other hemiascomycetes (Fig. 2) and more distantly related fungi (*12*) do have some exon-hosted snoRNAs. It is unknown, however, whether processing of such snoRNAs involves Rnt1 (SOM text).

Studies of protein-coding genes have revealed a dramatic increase in the rate of intron loss in the hemiascomycete ancestor (*21*). By focusing instead on noncoding RNAs, we describe unexpected patterns of both exon and intron loss. A drastic reduction in the sizes of internal exons has ultimately led to their complete loss in *Candida* species, resulting in a complex form of splicing that maintains the independent removal of now overlapping introns. The *Saccharomyces* complex, however, has experienced a second wave of intron loss, perhaps facilitated by innovations in snoRNA processing and mediated by a mechanism—splice site degeneration—distinct from that which has acted on protein-coding genes.

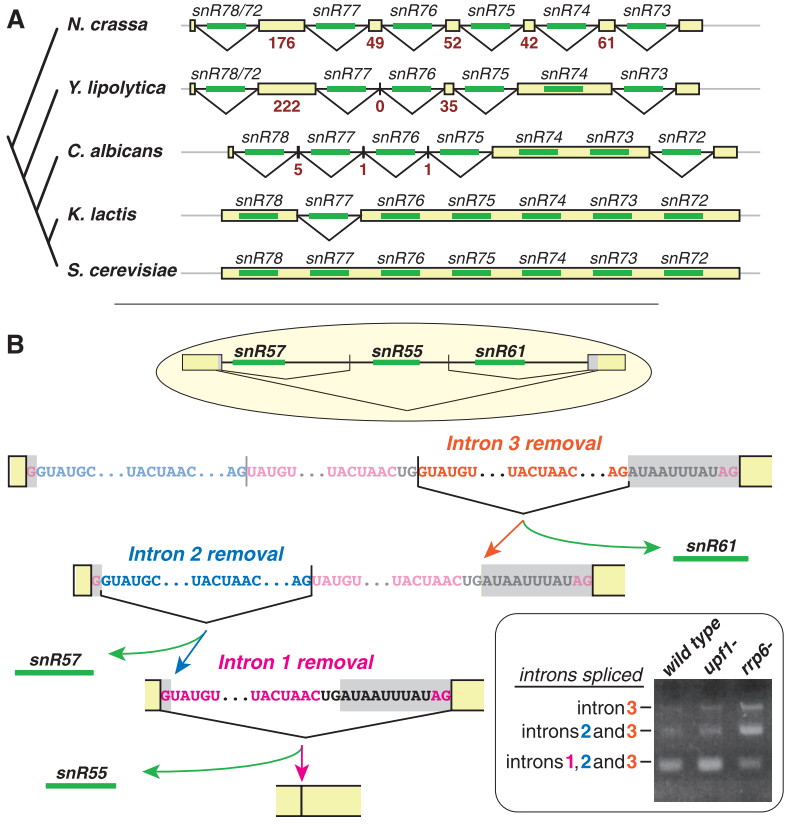


Fig. 3. Polycistronic snoRNA precursors exhibit unusual splicing patterns. (**A**) Splicing of the *snR72-78* cluster in various fungal species (phylogenetic relations on left). snoRNAs are shown in green, introns as lines, and exons as yellow boxes, with internal exons labeled by size. Cotranscription of entire clusters has been demonstrated only for *S. cerevisiae* (*16*) and *N. crassa* (*12*). (**B**) Nested splicing of the *snR57/55/61* snoRNA cluster in *C. albicans*. The 5′ splice site, branch site, and 3′ splice site sequences are shown for introns 1 (pink), 2 (blue), and 3 (orange). Gray "exons" represent sequences joined with intron 1, by splicing of introns 2 and 3, to create de novo intron 1 splice sites. (Inset) Reverse transcription polymerase chain reaction products of the snoRNA host transcript from wild-type cells and those deficient for nonsense-mediated mRNA decay (*upf1-Δ*) or nuclear exonucleolytic decay (*rrp6-Δ*). We infer the order of splicing events from the observable products: Intron 3 is nearly always removed, intron 2 is removed in a subset of transcripts (lower two bands), and intron 1 only when 2 and 3 have also been removed (lower band). Patterns in decay mutants are consistent with degradation of partially spliced host transcripts in the nucleus.

References and Notes

1. M. C. Wahl, C. L. Will, R. Lührmann, *Cell* **136**, 701 (2009).
2. W. Li, A. E. Tucker, W. Sung, W. K. Thomas, M. Lynch, *Science* **326**, 1260 (2009).
3. L. Carmel, Y. I. Wolf, I. B. Rogozin, E. V. Koonin, *Genome Res.* **17**, 1034 (2007).
4. S. W. Roy, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5773 (2005).
5. S. W. Roy, W. Gilbert, *Nat. Rev. Genet.* **7**, 211 (2006).
6. D. Penny, M. P. Hoeppner, A. M. Poole, D. C. Jeffares, *J. Mol. Evol.* **69**, 527 (2009).
7. Q. M. Mitrovich, B. B. Tuch, C. Guthrie, A. D. Johnson, *Genome Res.* **17**, 492 (2007).
8. G. R. Fink, *Cell* **49**, 5 (1987).
9. B. B. Tuch *et al.*, *PLoS Genet.* **6**, e1001070 (2010).
10. B. R. Braun *et al.*, *PLoS Genet.* **1**, 36 (2005).
11. L. B. Weinstein, J. A. Steitz, *Curr. Opin. Cell Biol.* **11**, 378 (1999).
12. N. Liu *et al.*, *BMC Genomics* **10**, 515 (2009).
13. D. Piekna-Przybylska, W. A. Decatur, M. J. Fournier, *RNA* **13**, 305 (2007).
14. C. Marck *et al.*, *Nucleic Acids Res.* **34**, 1816 (2006).
15. Materials and methods are available as supporting material on *Science* Online.
16. L. H. Qu *et al.*, *Mol. Cell. Biol.* **19**, 1144 (1999).
17. B. Dujon *et al.*, *Nature* **430**, 35 (2004).
18. J. W. Brown, D. F. Marshall, M. Echeverria, *Trends Plant Sci.* **13**, 335 (2008).
19. M. Lynch, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6118 (2002).
20. G. Chanfreau, P. Legrain, A. Jacquier, *J. Mol. Biol.* **284**, 975 (1998).
21. J. E. Stajich, F. S. Dietrich, S. W. Roy, *Genome Biol.* **8**, R223 (2007).
22. T. M. Lowe, S. R. Eddy, *Science* **283**, 1168 (1999).

23. We thank C. Barbacioru, O. Homann, S. Kuersten, S. Ledoux, I. Listerman, T. Lowe, C. Maeder, K. Pande, A. Price, S. Roy, J. Steitz, and members of the Guthrie and Johnson laboratories for helpful discussions and support; and M. Barker and C. Monighetti for providing sequence reads. This work was supported by NIH grants GM021119 (C.G.) and GM37049 (A.D.J.) and by Life Technologies Corp. C.G. is an American Cancer Society Research Professor of Molecular Genetics. Sequence data are available at the Gene Expression Omnibus (accession number GSE21291).

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6005/838/DC1
Materials and Methods
SOM Text
Figs. S1 to S4
Tables S1 to S6
References

1 July 2010; accepted 21 September 2010
10.1126/science.1194554

Fate Mapping Analysis Reveals That Adult Microglia Derive from Primitive Macrophages

Florent Ginhoux,[1,2]* Melanie Greter,[1] Marylene Leboeuf,[1] Sayan Nandi,[3] Peter See,[2] Solen Gokhan,[4] Mark F. Mehler,[4,5] Simon J. Conway,[6] Lai Guan Ng,[2] E. Richard Stanley,[3] Igor M. Samokhvalov,[7] Miriam Merad[1]*

Microglia are the resident macrophages of the central nervous system and are associated with the pathogenesis of many neurodegenerative and brain inflammatory diseases; however, the origin of adult microglia remains controversial. We show that postnatal hematopoietic progenitors do not significantly contribute to microglia homeostasis in the adult brain. In contrast to many macrophage populations, we show that microglia develop in mice that lack colony stimulating factor-1 (CSF-1) but are absent in CSF-1 receptor–deficient mice. In vivo lineage tracing studies established that adult microglia derive from primitive myeloid progenitors that arise before embryonic day 8. These results identify microglia as an ontogenically distinct population in the mononuclear phagocyte system and have implications for the use of embryonically derived microglial progenitors for the treatment of various brain disorders.

Although microglial ontogeny is an extensive area of research, much controversy remains regarding the nature of microglial progenitors (*1*, *2*). The most consensual hypothesis to date is that embryonic and perinatal hematopoietic waves of microglial recruitment and differentiation occur in the central nervous system (CNS) (*1*, *2*). However, the exact contribution of embryonic and postnatal hematopoietic progenitors to the adult microglial pool in the steady state remains unclear. We examined the contribution of primitive and definitive hematopoiesis to the adult microglial population that populates the CNS during normal development. Our results provide direct evidence that adult microglia derive from primitive myeloid progenitors that arise before embryonic day 8 (age E8.0) and for the predominant contribution of primitive myeloid progenitors to an adult hematopoietic compartment.

To address the contribution of perinatal circulating hematopoietic precursors to microglial homeostasis, we reconstituted sublethally irradiated C57BL/6 CD45.2$^+$ newborns with hematopoietic cells isolated from CD45.1$^+$ congenic mice (*3*). Although more than 30% circulating leukocytes and tissue macrophages were of donor origin 3 months after transplant (fig. S1A), 95% of adult microglia remained of host origin at this time point (fig. S1, A and B). These results suggest that, in contrast to previous reports (*4*, *5*), perinatal circulating hematopoietic precursors, including monocytes, do not substantially contribute to adult microglial homeostasis. With use of adult congenic bone marrow chimera models, evidence in favor of (*6*–*8*) and against (*9*, *10*) the contribution of circulating hematopoietic cells to microglial homeostasis has been proposed. We found consistently that 10 to 20% of microglia in the brain parenchyma are of donor origin at 10, 15, and 21 months after transplant (fig. S1C). Parabiotic mice, which share the same blood circulation, provide a means to follow the turnover of adult circulating hematopoietic precursors without the need for exposure to radiation injuries. Although the mixing of the myeloid lineage is less efficient than the mixing of the lymphoid lineage (*11*), an average of 30% of monocytes and tissue macrophages were donor-derived at 1 month and 12 months after parabiosis (fig. S1D) (*12*). In contrast, less than 5% of microglia were donor-derived at these time points (fig. S1D), in agreement with a previous report on 5-month-old parabionts (*9*). Consistent with previous reports (*9*, *10*, *13*), these results suggest that the recruitment of bone marrow–derived cells to the brain of chimeric animals is dependent on radiation-induced brain injuries that followed the transplantation regimen. These results also suggest that postnatal microglia are maintained independently of circulating monocytes throughout life and are maintained by local radio-resistant precursors that colonize the brain before birth.

Next, we examined the origin of microglia during development. In mouse embryos, the first wave of hematopoietic progenitors appears in the extra-embryonic yolk sac and leads to the production of primitive hematopoiesis, which takes place between E7.0 and E9.0 (*14*, *15*). An independent wave of hematopoiesis termed "definitive hematopoiesis" is initiated within the embryo proper in the aorta, gonads, and mesonephros (AGM) region (*14*, *15*). Around E10.5, hematopoietic progenitors start to colonize the fetal liver, which serves as a major hematopoietic organ after E11.5, whereas later during develop-

[1]Department of Gene and Cell Medicine and the Immunology Institute, Mount Sinai School of Medicine, 1425 Madison Avenue, New York, NY 10029, USA. [2]Singapore Immunology Network (SIgN), 8A Biomedical Grove, IMMUNOS Building Nos. 3-4, BIOPOLIS, 138648, Singapore. [3]Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA. [4]Institute for Brain Disorders and Neural Regeneration, Rose F. Kennedy Center for Research on Intellectual and Developmental Disabilities, and Department of Neurology, Albert Einstein College of Medicine, 1410 Pelham Parkway South, Bronx, NY 10461, USA. [5]Departments of Neuroscience, Psychiatry, and Behavioral Sciences, Albert Einstein College of Medicine, 1410 Pelham Parkway South, Bronx, NY 10461, USA. [6]Herman B Wells Center for Pediatric Research, Indiana University School of Medicine, 1044 West Walnut Street, Indianapolis, IN 46202, USA. [7]Laboratory for Stem Cell Biology, Center for Developmental Biology (CDB), RIKEN Kobe, Kobe 6500047, Japan.

*To whom correspondence should be addressed. E-mail: Miriam.Merad@mssm.edu (M.M.); Florent_ginhoux@immunol.a-star.edu.sg (F.G.)