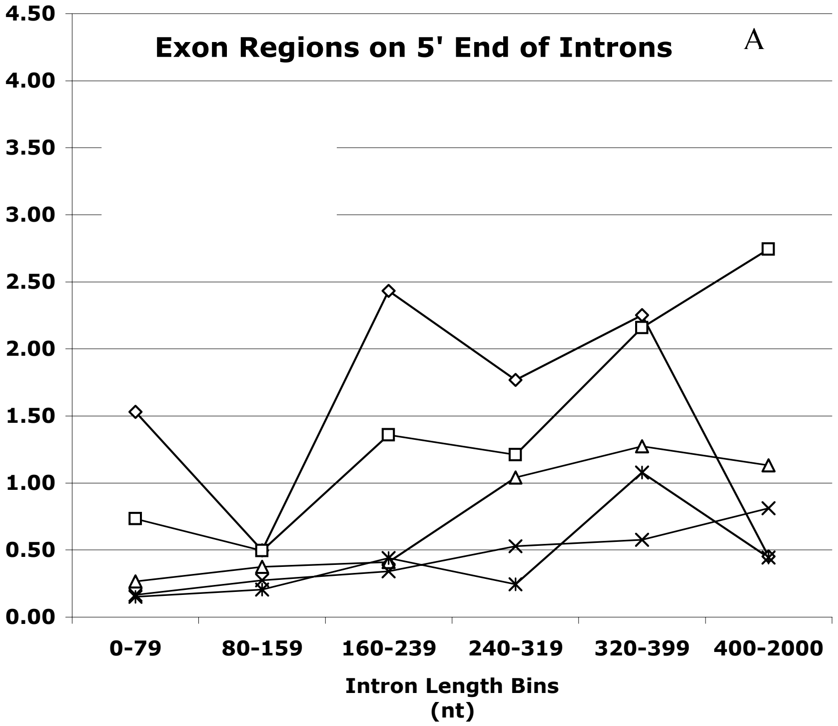
FIG. S1. Information content (bits) in 5 nt of the exon regions adjacent to the intron in various length ranges of introns.



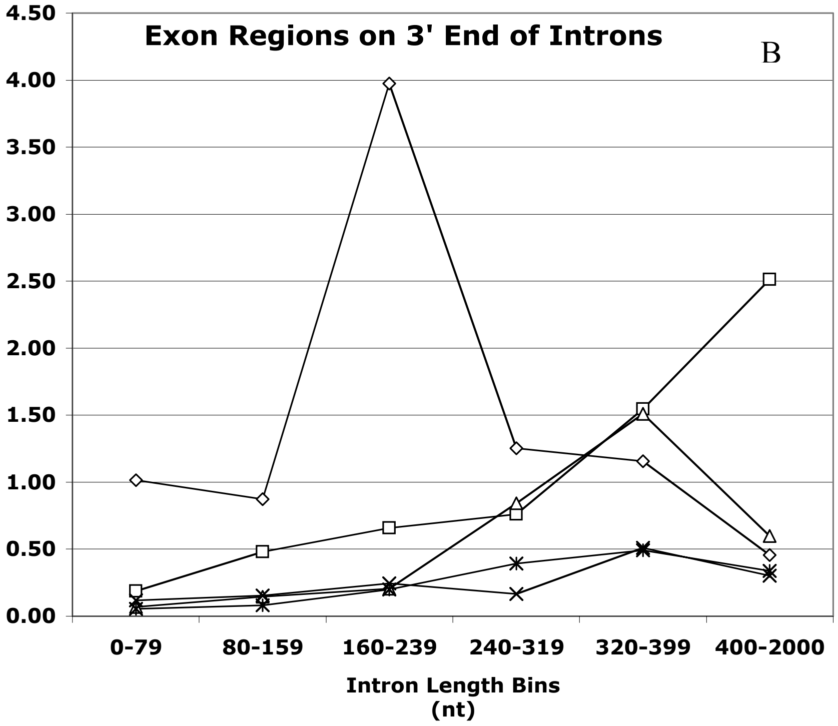
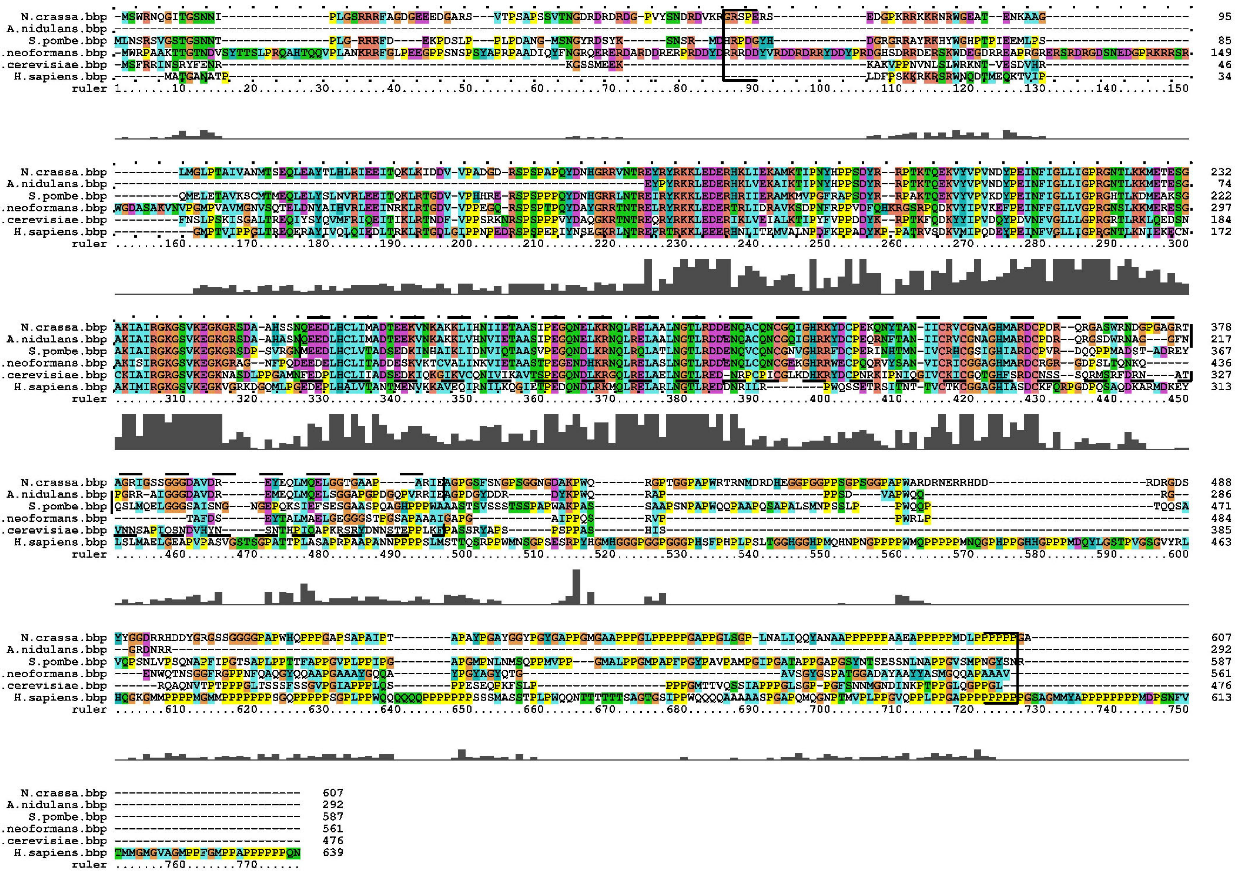


FIG. S2. Branchpoint binding protein (BBP) homologs. (A) Sequence alignment of the BBP homologs in *S. cerevisiae* (AC# NP_013217.1), *S. pombe* (AC# AF073779 1), *A. nidulans* (1.84 (scaffold 5) http://www.broad.mit.edu/annotation/fungi/aspergillus/geneindex.html and AC# AA784699, region homologous to *S. pombe* BBP)*, N. crassa* (AC# CAD36971.1), *C. neoformans*, B3501 serotype D (locus163.m02728 http://www.tigr.org/tdb/e2k1/cna1/), and *H. sapiens* SF1, (AC# CAA70018.1). The SF1/BBP (KOG0119) motif is in brackets, the MSL5 domain is in a dotted box, and the AIR domain is in a dashed box. (B) Phylogram for BBP homologs.

A



Bracket KOG0119
Dotted MSL5
Dashed AIR

B

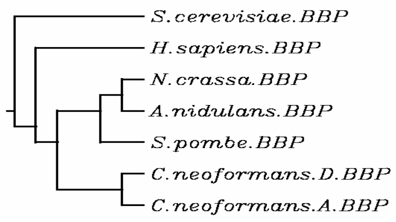
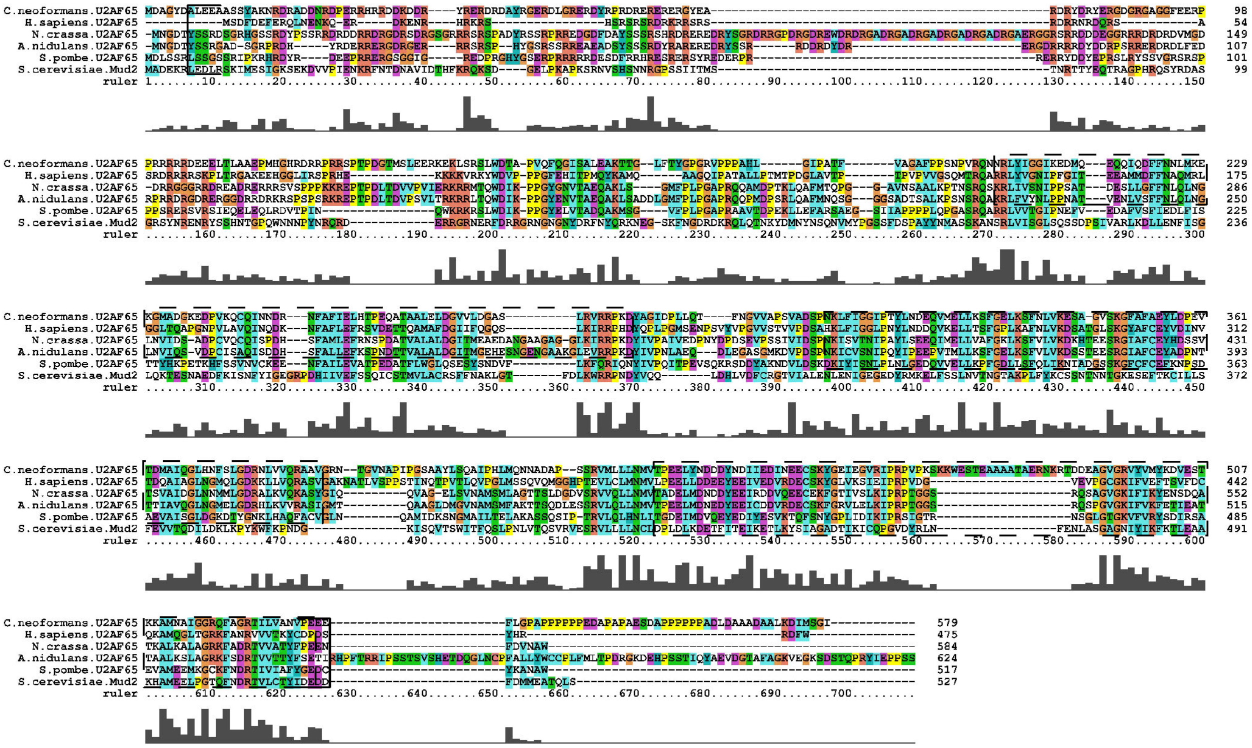
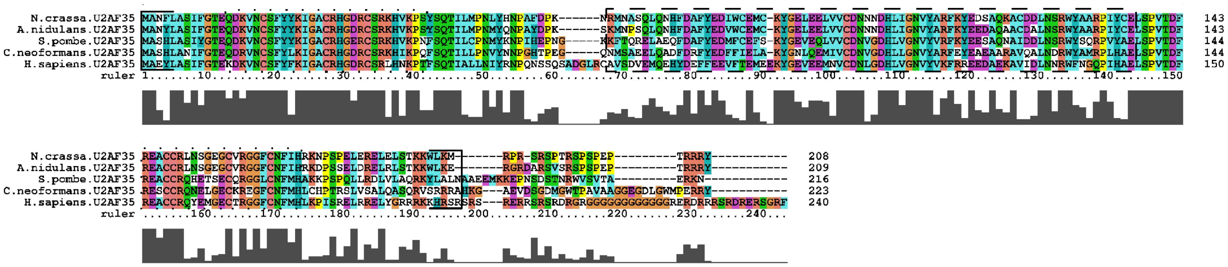


FIG. S3. U2AF^{65} protein homologs. Sequence alignment of the U2AF^{65} homologs in *S. pombe* (AC# CAB46760), *N. crassa* (AC#EAA34808), *A. nidulans* (locusAN6681.2 http://www.broad.mit.edu/ annotation/fungi/aspergillus/geneindex.html), *C. neoformans*, B3501 serotype D (locus162.m00297 http://www.tigr.org/tdb/e2k1/cna1/) and *H. sapiens* (AC#CAA5409.1) as well as MUD-2 (AAA64215) in *S. cerevisiae*. The RRM domains are enclosed by dashed boxes and the U2AF^{65} characteristic domain is in brackets.



Bracket KOG0120
Dashed RRM

FIG. S4. U2AF^{35} protein homologs. Sequence alignment of the U2AF^{35} homologs in *N. crassa* (contig 3.170 (scaffold 9), http://www-genome.wi.mit.edu/cgi-bin/annotation/Neurospora, region homologous with *S. pombe* U2AF^{35}), *A. nidulans* (1.172 (scaffold 16) http://www.broad.mit.edu/ annotation/fungi/aspergillus/geneindex.html), region homologous with *S. pombe* U2AF^{35}), *S. pombe* (AC#Q09176), *H. sapiens* (AC#Q01081), and *C. neoformans*, B3501 serotype D (locus162.m02737 http://www.tigr.or/tdb/e2k1/cna1). *S. cerevisiae* does not have a U2AF^{35} equivalent [Abovich, 1997 #83]. The RRM domain is in a dashed box, the Zn finger regions are enclosed in the dotted boxes, and the KOG2202 domain characteristic for the small U2AF subunit is in brackets.



Bracket KOG2202
Dashed RRM
Dotted Zn finger CCCH

FIG. S5. Nam8p, TIA-1 and TIAR homologs. (A) Sequence alignment of the TIA1/TIAR/NAM8 homologs in *S. pombe* (CSX1, AC# CAB16569.1), *N. crassa* (AC# NCU00768.1, XP_328580.1), *A. nidulans* (AC# EAA61923.1, EAA66209.1), *C. neoformans*, B3501 serotype D (locus 163.m06329, http://www.tigr.org/tdb/e2k1/cna1/), and *H. sapiens* (AC# NP_071505.1, A46174) as well as NAM8 (AC# Q00539), in *S. cerevisiae*. The RRM domains are enclosed by dashed boxes and KOG0148 is identified by brackets.

Bracket KOG0148
Dashed RRM