



Figure S4. The finding and definition of conserved dual coding regions from multi species alignments. The orthologous transcripts from four species were first aligned and then translated using the second reading frame. Hence, additional start and stop codons appeared in the translation. For each of the species, an uninterrupted segment of peptides were identified (the dotted line with arrow ends at both direction) and the first start codon was marked. The region between the closest start-stop codons was defined as ARF region. From the same set of transcripts, regions from the beginning to first stop codon in any one of the species and the last stop codon to the end of the transcript were defined as flanking ORF region.