

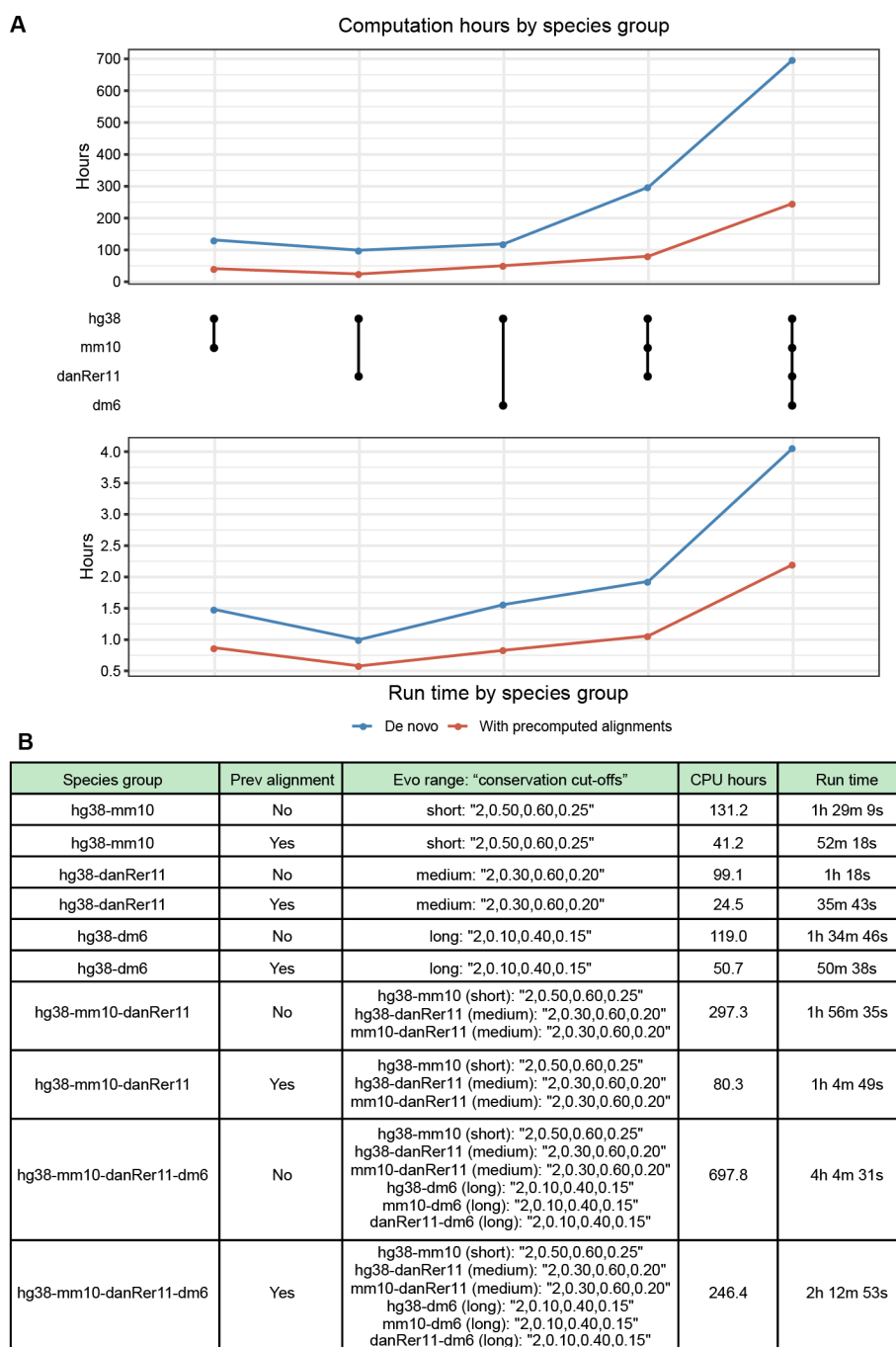
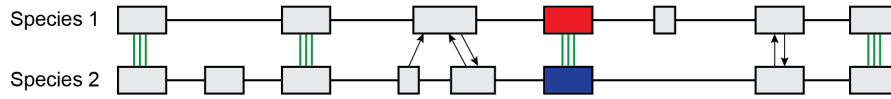


Figure S1 – Computational performance of *ExOrthist*. (a) CPU hours and run time required by each of the pairwise and multi-species runs used to evaluate *ExOrthist* computational performance (see Methods). The species considered in each run are indicated by black dots in the middle section of the plot. The blue line charts refer to *de novo ExOrthist* runs, while the orange line charts refer to runs performed with all pre-computed pairwise alignments (--prevaln option). (b) Complete information for the *ExOrthist main* runs depicted in (a): required CPU hours, run time as well as other essential information (species, evolutionary distance ranges, conservation cut-offs, and whether or not the run included

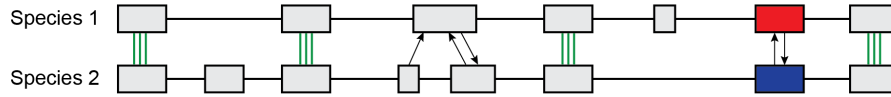
previously generated alignments). The cut-offs are in the format required by *ExOrthist main*: “int_num,ex_seq,ex_len,prot_sim”. Additional information can be found in the *Nextflow* report for each of the runs (Additional file 2: Files S1-S10).

1) Conservation scenarios:

* **Conserved:** both compared exons (red and blue) belong to the same exon cluster.

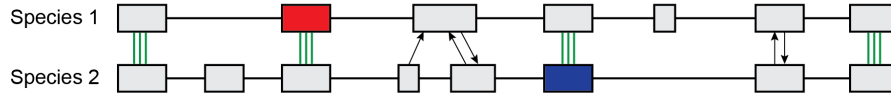


* **Best hit:** the compared exons are reciprocal or single best hits (if main.nf output is provided).

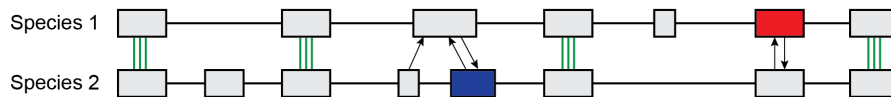


2) NON-conservation scenarios:

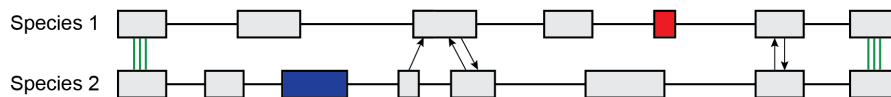
* **Non-Conserved (A):** each compared exon belongs to a different exon cluster.



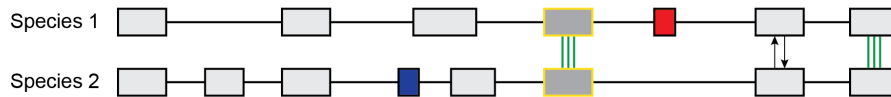
* **Non-Conserved (B):** the compared exons have different best hits (if main.nf output is provided).



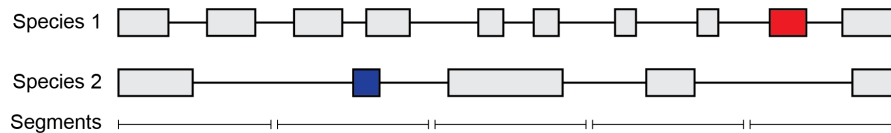
* **Non-Conserved (C):** difference in exon length is larger than a given cut-off



* **Non-Conserved (D):** the exons are at opposite sides of a conserved exon (anchor; yellow rim).



* **Non-Conserved (E):** no anchors available, both genes have ≥ 5 exons AND the exons are in different "exon segments" of the gene (≥ 2 segments away).



3) Unclear scenarios: at the same side of all anchors or in the same/nearby segment if no anchor.

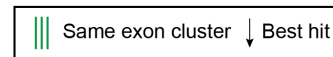
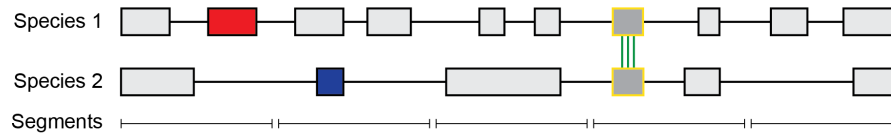


Figure S2 - Scenarios of pairwise exon comparisons in orthologous genes by *compare_exon_sets*. Each non-conservation scenario is evaluated in a consecutive manner (i.e. if the conditions for A are met, the rest of the scenarios are not assessed).

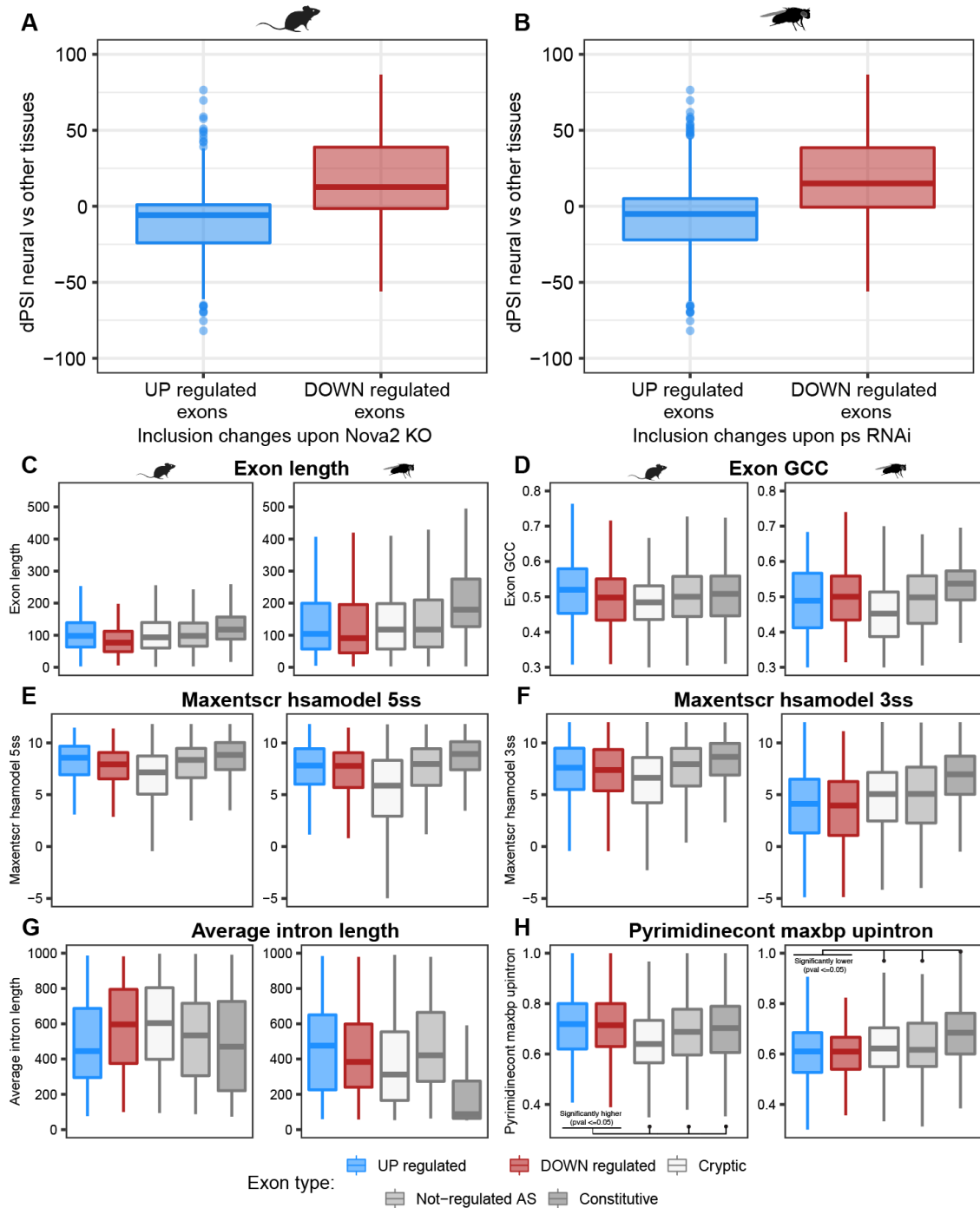


Figure S3 – Features of *Nova*-regulated exons in mouse and fruitfly. (a,b) Distributions of delta PSIs (i.e. differences in inclusion levels) between neural and non-neural samples for exons that are up-regulated (blue) or down-regulated (red) upon *Nova2* depletion in mouse (a) or *ps* knock-down in fruitfly (b). (c-h) Comparison of selected exon features as assessed by *matt cmpr_exons* (see Methods). Boxplots in each panel show the distributions of the feature across *Nova*-regulated exons (blue and red; respectively up- and down-regulated upon *Nova* perturbation) and three sets of control exons (cryptic exons, alternative exons not

regulated by *Nova* and constitutive exons) for the mouse (left) or fruitfly (right). Significant differences (Mann-Whitney U tests) between regulated exons and the control sets were only reported for features where the regulated exons presented different trends in the two species. The single panels refer to exon length (c), exon GC content (d), 5' splice site strength (e), 3' splice site strength (f), average intron length (between upstream and downstream introns) (g), and length of the pyrimidine tract in the upstream intron (h).

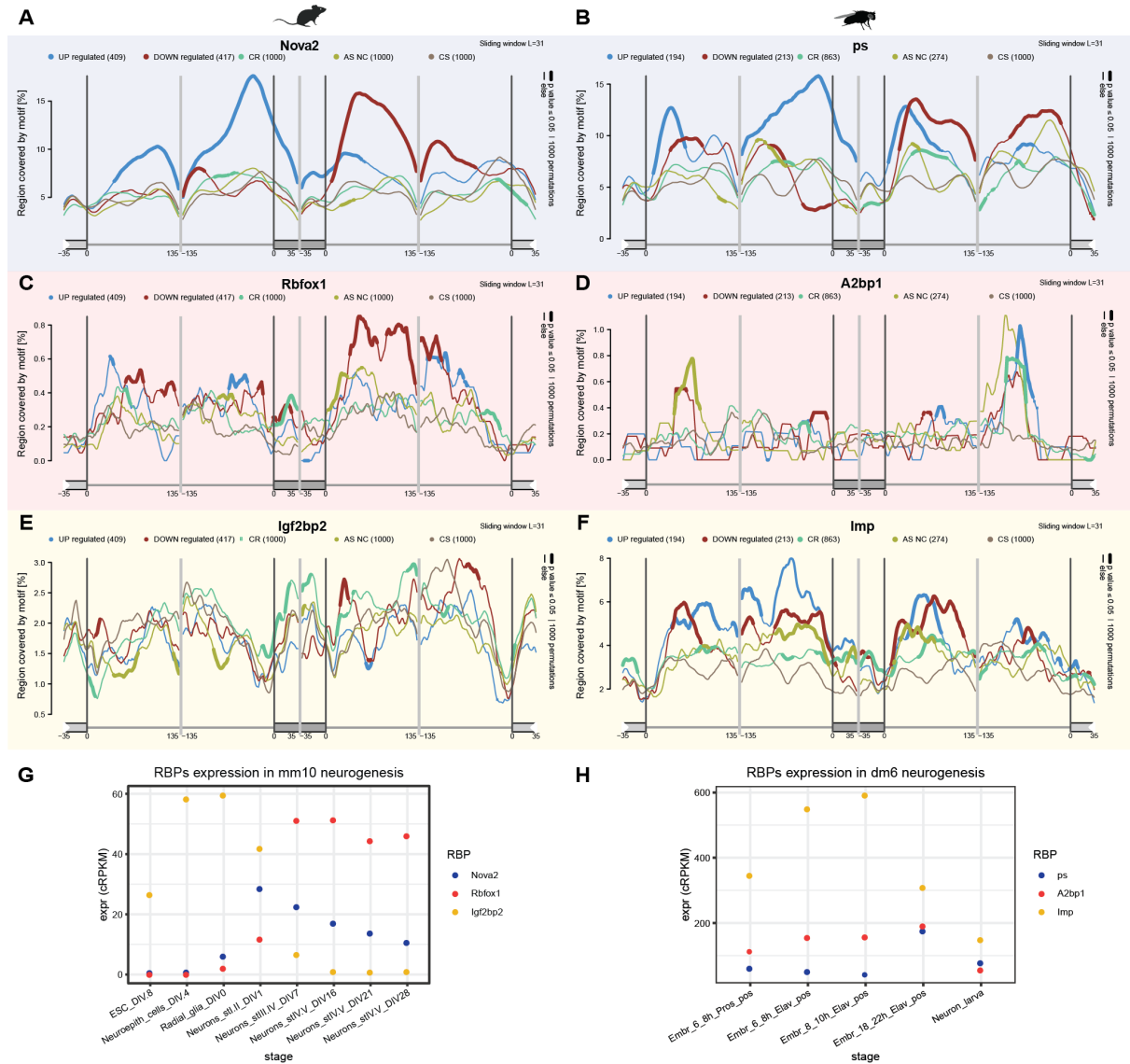


Figure S4 – Motif enrichments around *Nova*-regulated exons. (a-f) RNA maps testing motif enrichment around *Nova*-regulated exons for specific RNA binding proteins (RBPs) in mouse (a,c,e) and their respective orthologs in fruitfly (b,d,f). Thicker lines correspond to significantly enriched regions based on permutations tests done by *Matt*. **(g,h)** Expression patterns of the RBPs represented in (a-f) during mouse neuronal differentiation (g) and fruitfly neurogenesis (h). Expression data obtained from *VastDB*.