

## SUPPLEMENTAL MATERIALS

### **Comprehensive mapping of the 5' and 3' untranslated regions of *Aspergillus fumigatus* reveals diverse mechanisms of mRNA processing including premature transcription termination.**

Lukas Schrettenbrunner<sup>1</sup>, Corinne Maufrais<sup>2,3</sup>, Guilhem Janbon<sup>2</sup>, Edward W. J. Wallace<sup>4</sup>,  
Matthew G. Blango<sup>1\*</sup>

<sup>1</sup> Junior Research Group RNA Biology of Fungal Infections, Leibniz Institute for Natural Product Research and Infection Biology—Hans Knöll Institute (Leibniz-HKI), 07745 Jena, Germany.

<sup>2</sup> Institut Pasteur, Université Paris Cité, Unité Biologie des ARN des Pathogènes Fongiques, Département de Mycologie, F-75015, Paris, France

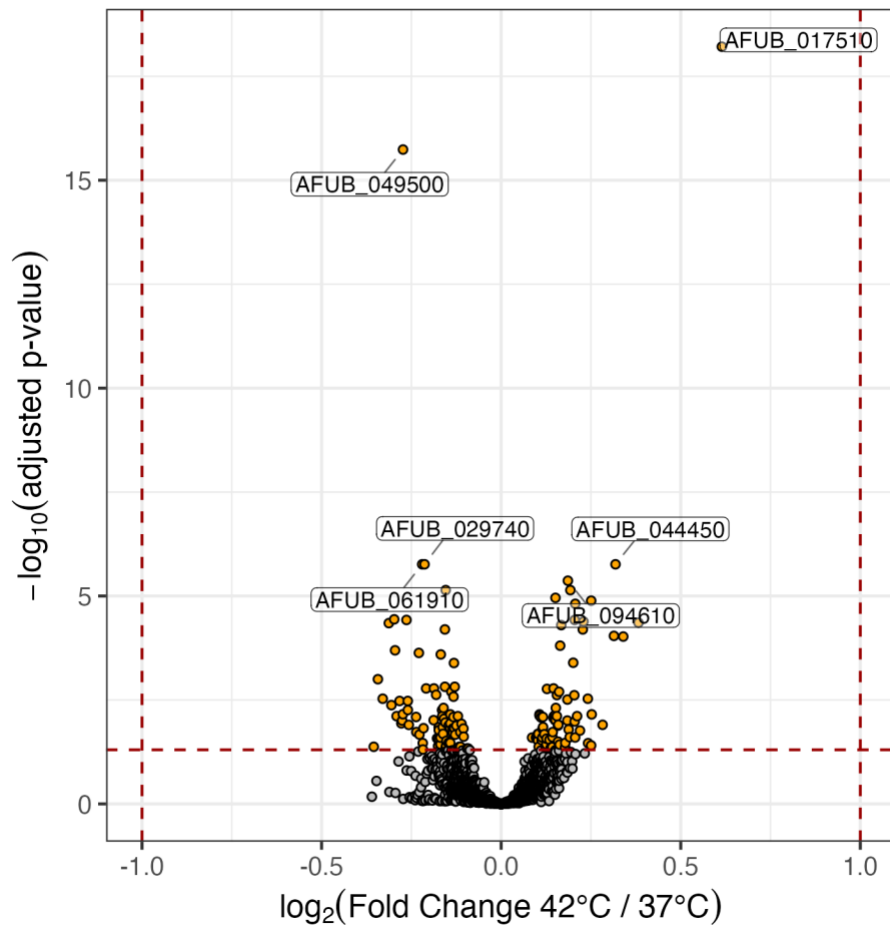
<sup>3</sup> Institut Pasteur, Université Paris Cité, HUB Bioinformatique et Biostatistique, C3BI, USR 3756 IP CNRS, F-75015, Paris, France

<sup>4</sup> University of Edinburgh, Institute for Cell Biology and Centre for Engineering Biology, School of Biological Sciences, University of Edinburgh, C.H. Waddington Building, Max Born Crescent, Edinburgh, Scotland, EH9 3BF, UK

**Running Title:** Defining *A. fumigatus* poly-(A)-enriched RNA ends.

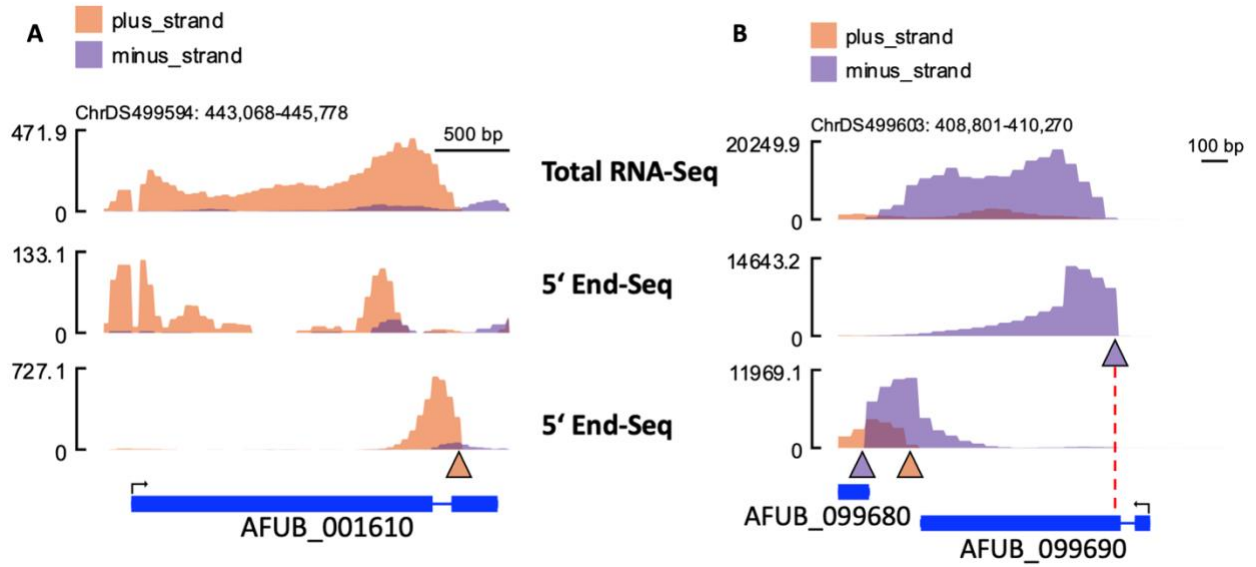
\*Correspondence: E-mail [matthew.blango@leibniz-hki.de](mailto:matthew.blango@leibniz-hki.de); Tel. +49 (0)3641-532 1072.

## SUPPLEMENTAL FIGURES

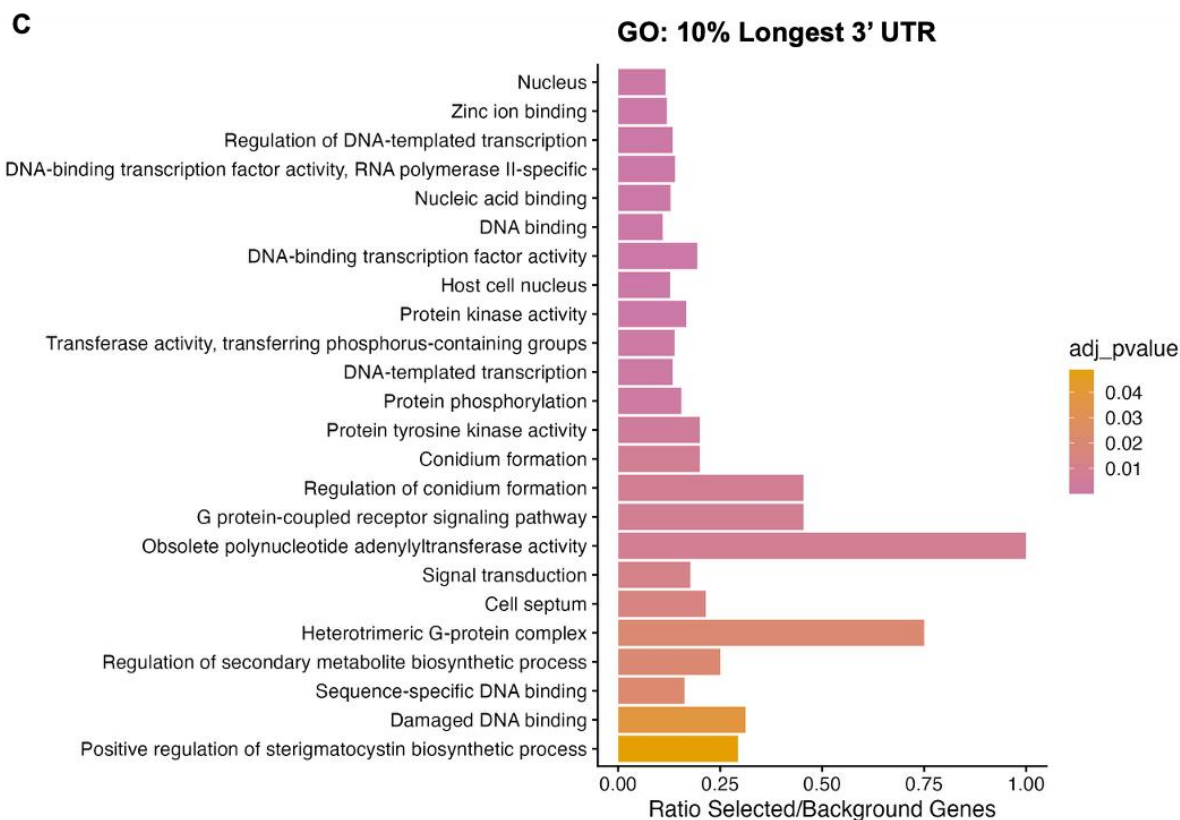
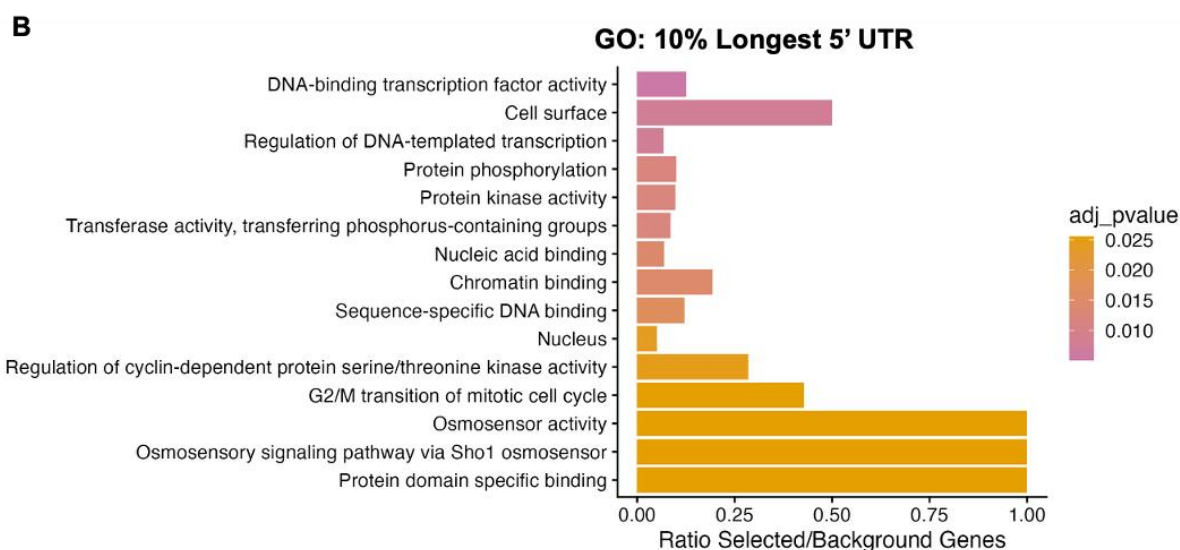
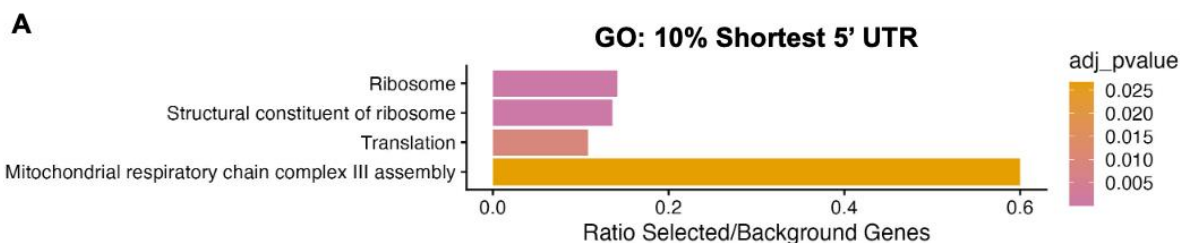


**Figure S1: Volcano plot of the differentially expressed genes between 37°C and 42°C.**

Graphical representation of Log<sub>2</sub> Fold Change data from Table S1 as a volcano plot. Orange dots represent genes with a  $p_{adj} < 0.05$ . Genes with missing adjusted p-values from DESeq2 were excluded.

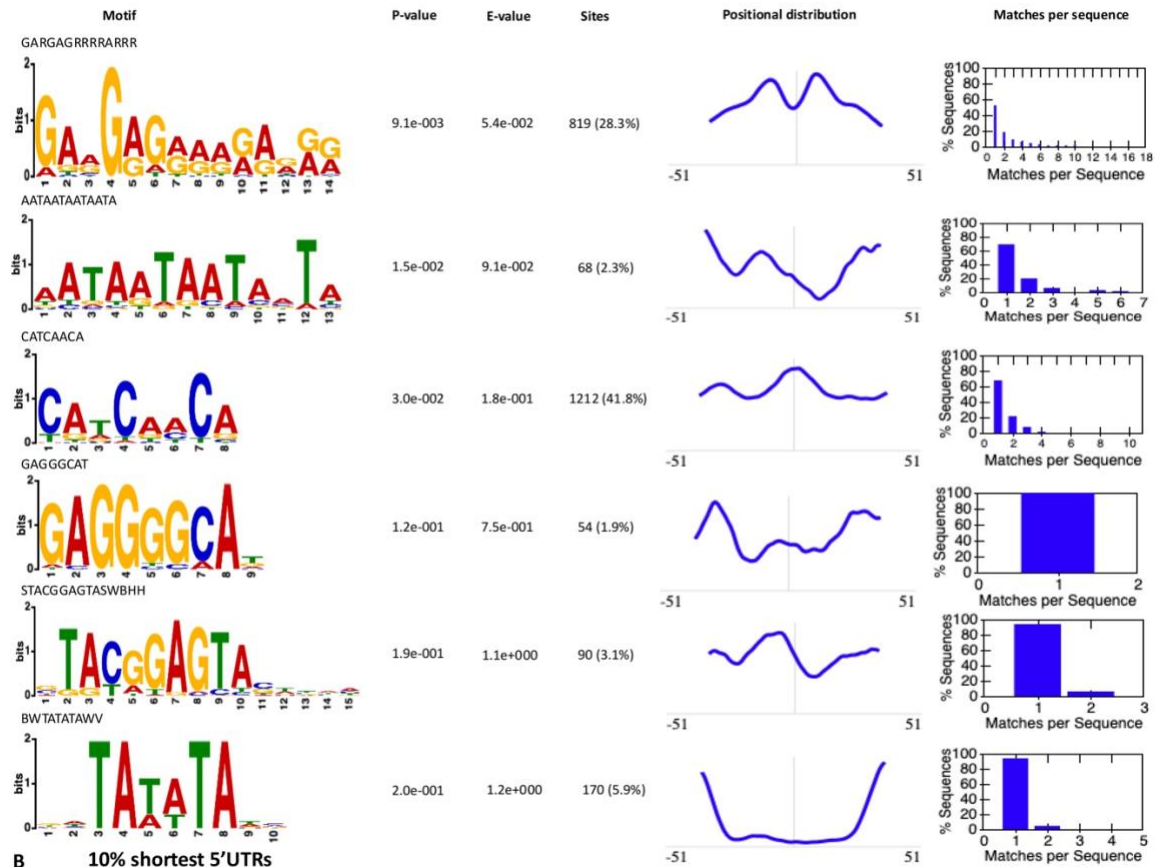


**Figure S2: UTR length bias originating from genes with predicted annotation.** (A) Read alignment of *AFUB\_099680* encoding a putative transmembrane transporter and *AFUB\_001610* encoding a protein of unknown function. (B) Read alignment of *AFUB\_099690* encoding a protein of unknown function. In both (A) and (B), the y-axis indicates total read-counts. Arrow heads indicate the positions of detected high confidence 5'/3' end sites. Read tracks were smoothed to improve viewability, this can also lead to some intron being depicted as unspliced, although they are actually spliced.

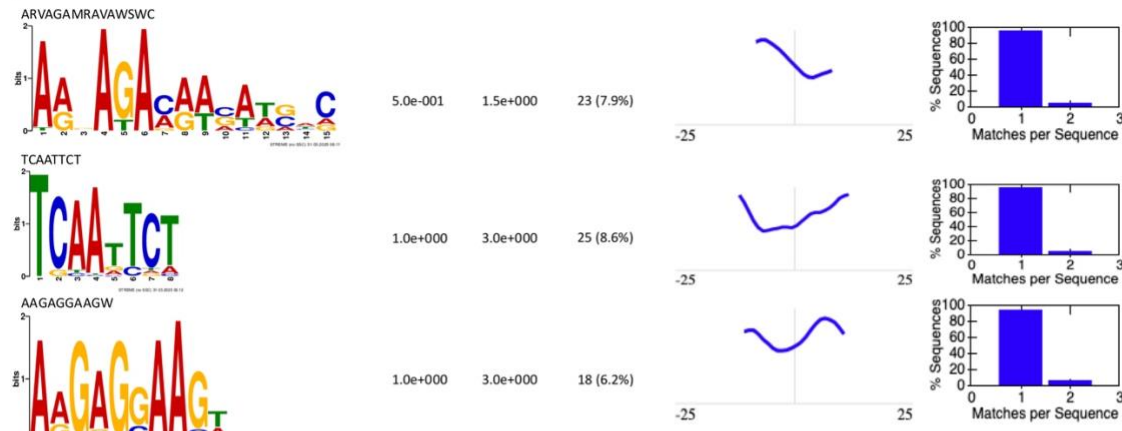


**Figure S3: GO-term analysis of the 10% shortest and longest UTRs.** Gene ontology terms associated with the genes with the (A) 10% shortest 5' UTRs, (B) 10% longest 5' UTRs, and (C) 10% longest 3' UTRs. No enrichment categories were detected for the 10% shortest 3' UTRs. GO terms were calculated with ShinyGO 0.82 (Ge et al. 2019).

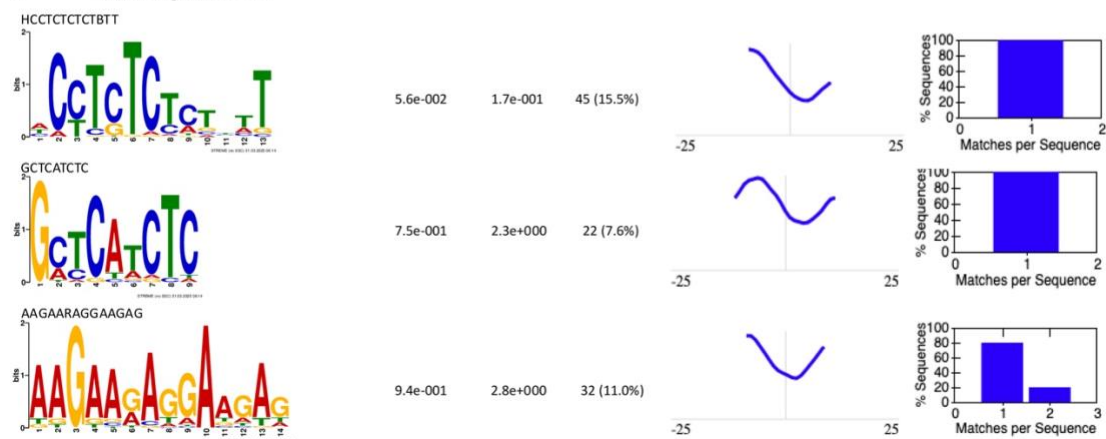
## A all TSS



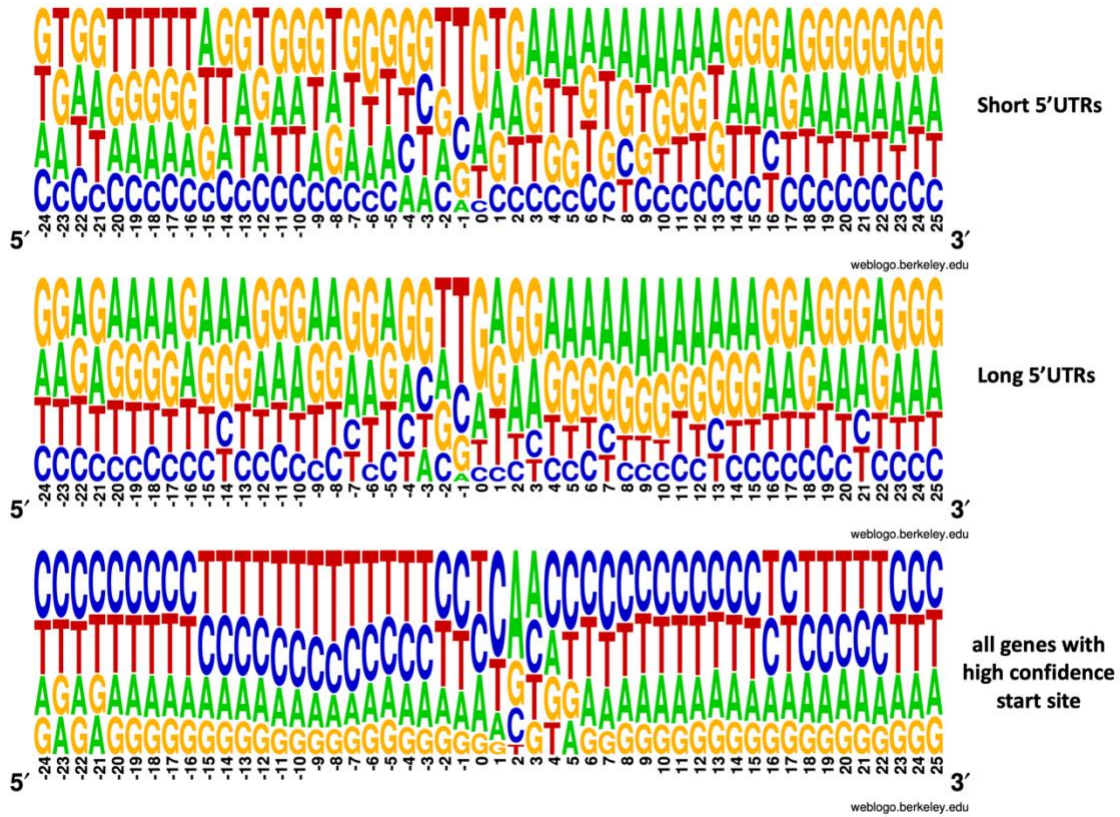
## B 10% shortest 5'UTRs



## C 10% longest 5'UTRs



**Figure S4: Motif analysis of transcriptional start sites.** Motif prediction using STREME with genomic sequences +/- 25 nt from the transcription start sites of the (A) all 5' UTRs, (B) 10% shortest 5' UTRs and (C) 10% longest 5' UTRs.



**Figure S5: Sequence context around transcriptional start sites.** WEB-logo of genomic sequences +/- 25 nt from transcription start sites for the denoted subsets or total high-confidence TES list. WEB-logo was created on <https://weblogo.threeplusone.com/>.

## **SUPPLEMENTAL TABLES**

**Table S1: DESeq2 analysis for differentially expressed genes in 42°C compared to 37°C.**

**Table S2: GO-terms of differentially expressed genes in 42°C compared to 37°C.**

**Table S3: Transcript start sites and transcript end sites for *A. fumigatus* mycelium.**

**Table S4: Number of detected uAUGs in high-confidence 5' UTRs.**

**Table S5: GO-term analysis of genes with the 10% shortest and longest 5' UTRs and 10% longest.**

**Table S6: Primers used for RT-qPCR.**

## **SUPPLEMENTAL REFERENCES**

Ge SX, Jung D, Yao R. 2019. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**: 2628-2629.