

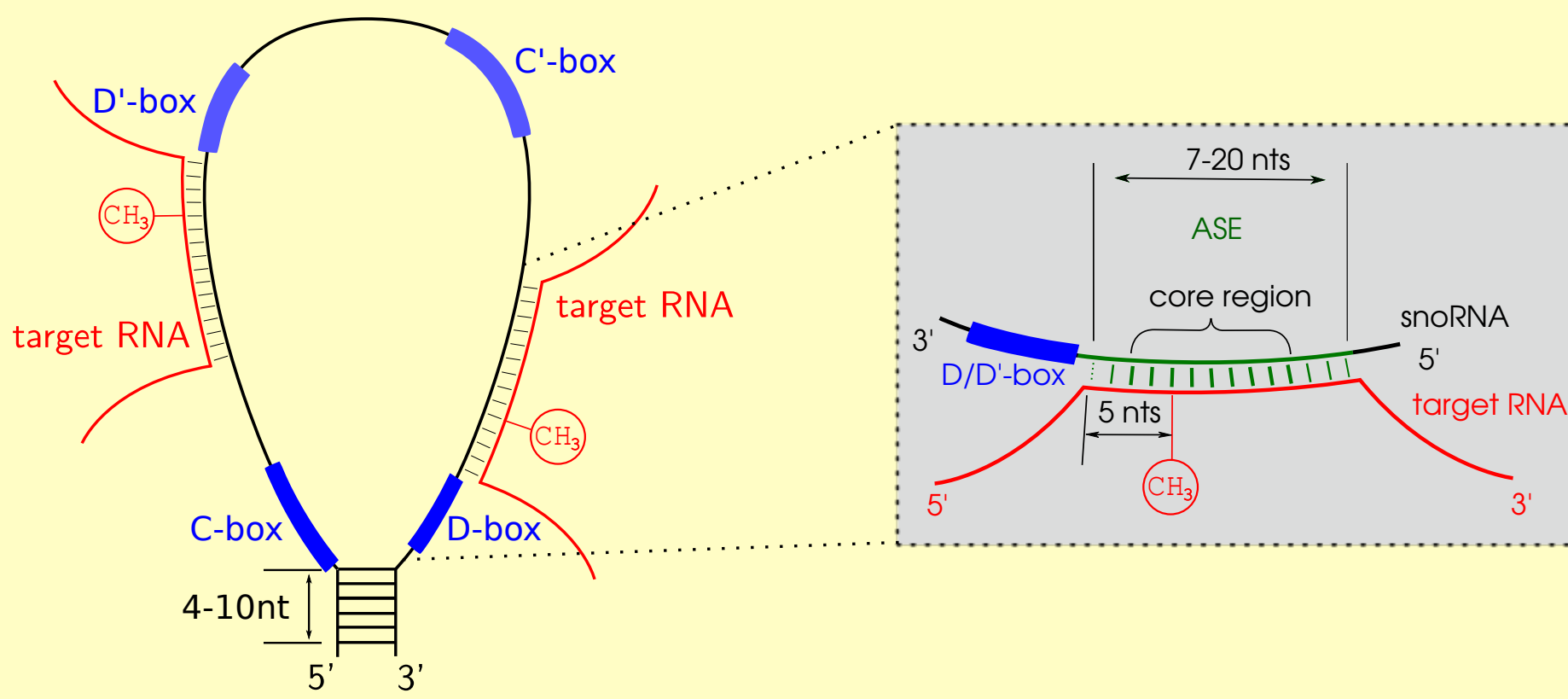
Matching Of Soulmates: SnoRNA Families and Their Targets

Stephanie Kehr*, Sebastian Bartschat, Hakim Tafer, Peter Stadler, Jana Hertel

Bioinformatics Group, Department of Computer Science and Interdisciplinary Center for Bioinformatics, University of Leipzig

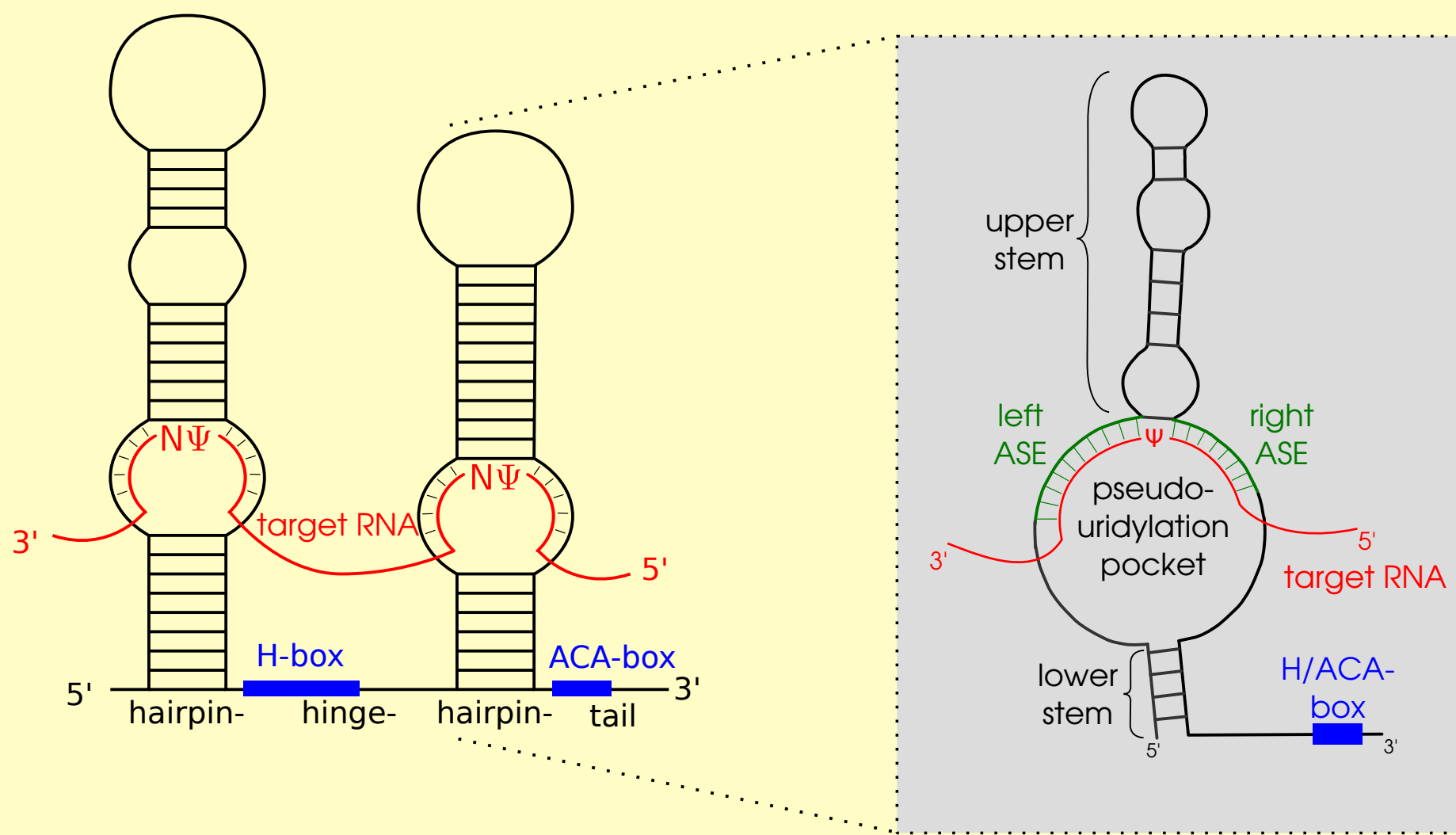
SnoRNA Function

- single nt modifications in ribosomal and small spliceosomal RNAs
- enriched at active centers
- pattern retained during evolution



Box C/D snoRNAs

- 2'-O-ribose-methylation
- mediated by methyltransferase fibrillarin
- upstream of D/D'-box small duplexes with target RNA



Box H/ACA snoRNAs

- conversion of uridine to pseudouridine
- mediated by dyskerin
- target RNA bound inside of interior loop
- two small duplex adjacent to upper hairpin

Interaction Conservation Index

- single sequence target predictions
 - Plexy¹ → box C/D snoRNAs
 - RNAsnoop² → box H/ACA snoRNAs
- ICI: conservation of interaction
 - quality of single interactions
 - degree of conservation

$$S(t, k) = \{s | X(t, s, k) \neq \emptyset\}$$

snoRNAs that target t in species k

$$T(s, k) = \{t | X(t, s, k) \neq \emptyset\}$$

targets of family s in species k

$$O(t, s) = \{k | X(t, s, k) \neq \emptyset\}$$

organisms where family s targets t

$$X(t, s, k) = \begin{cases} \text{set of snoRNAs from} \\ \text{family } s \text{ predicted to} \\ \text{guide modification of} \\ \text{target } t \text{ in species } k \end{cases}$$

$$ICI_{mod}(t, s) = \sum_{k \in O(t, s)} \frac{\varepsilon(t, s, k)}{\bar{\varepsilon}(t, k)}$$

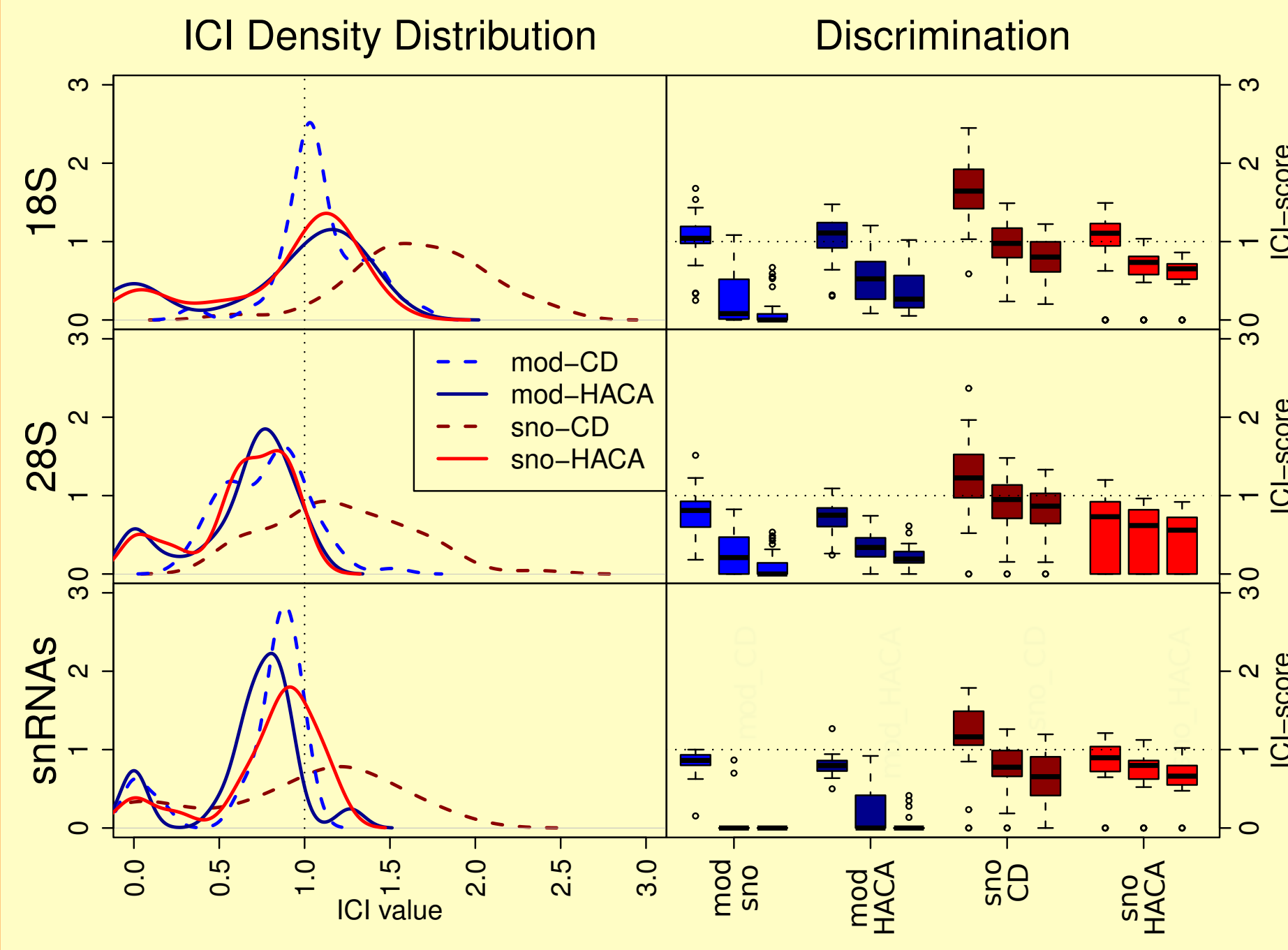
$$ICI_{sno}(t, s) = \sum_{k \in O(t, s)} \frac{\varepsilon(t, s, k)}{\hat{\varepsilon}(s, k)}$$

$$\varepsilon(t, s, k) = \frac{\sum_{t \in T(s, k)} \varepsilon(t, s, k)}{|T(s, k)|}$$

lowest minimum free energy of snoRNA $x \in X(t, s, k)$ interacting with region around target t

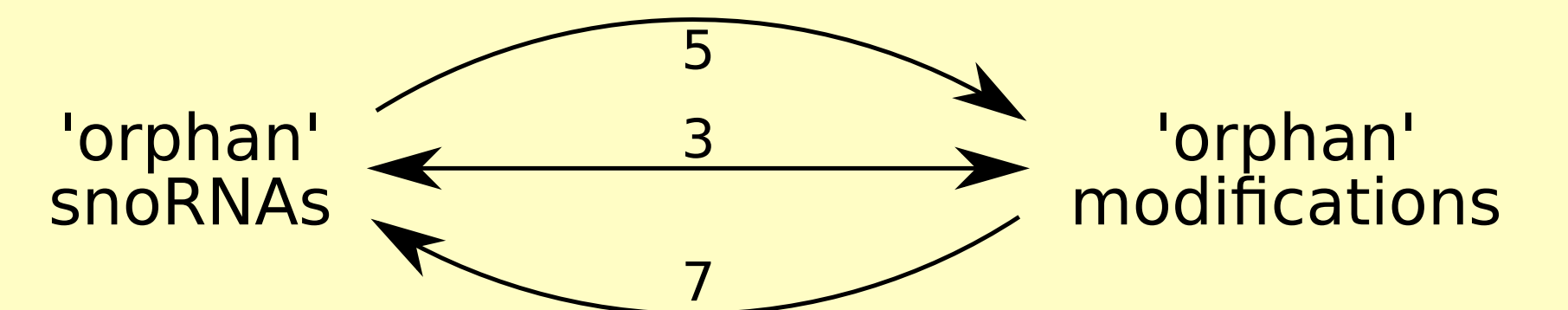
Known Interactions

- 18S: average ICI > 1 for known interactions
- 28S: low alignment quality due to incomplete sequences
- score highly discriminating

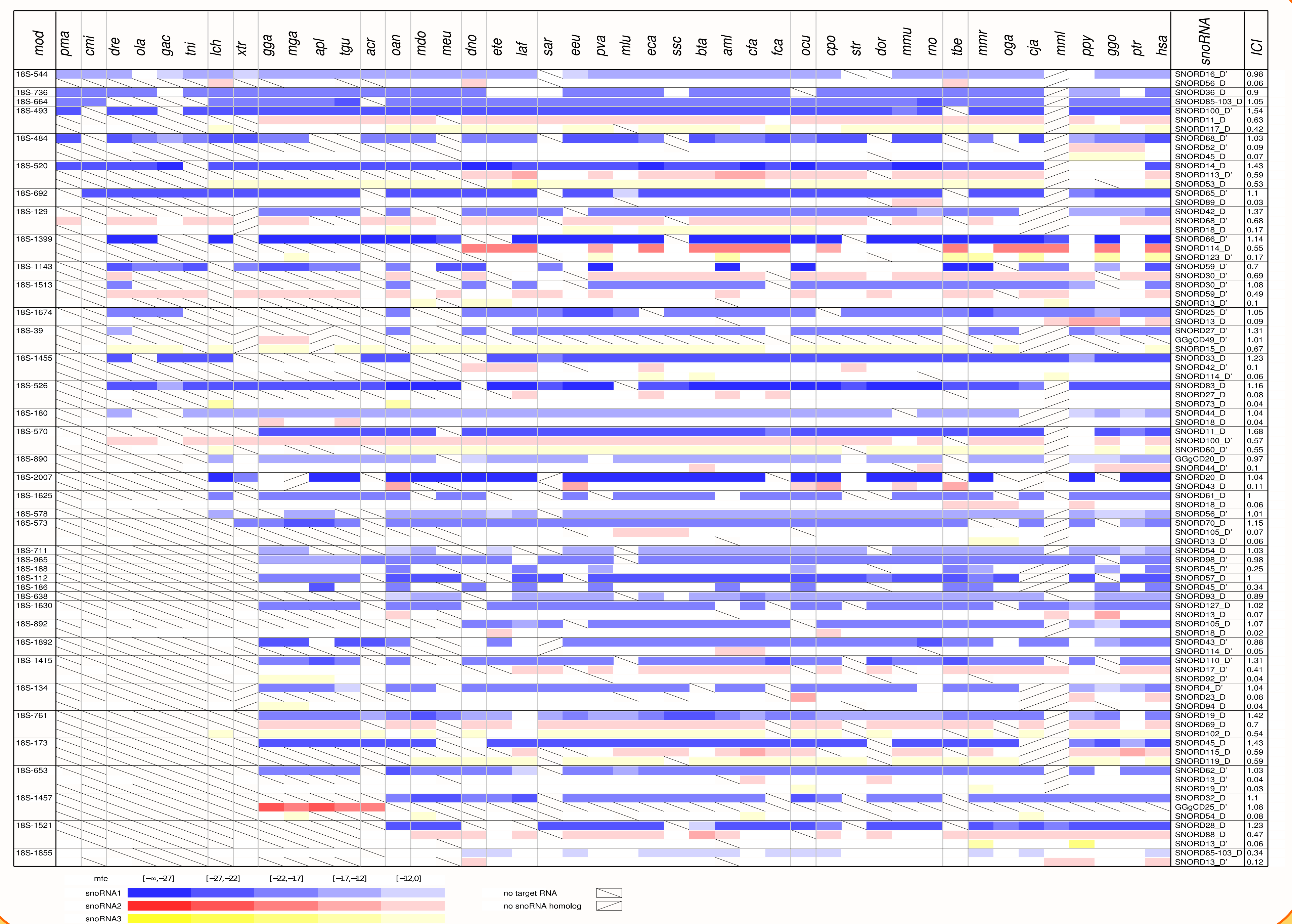
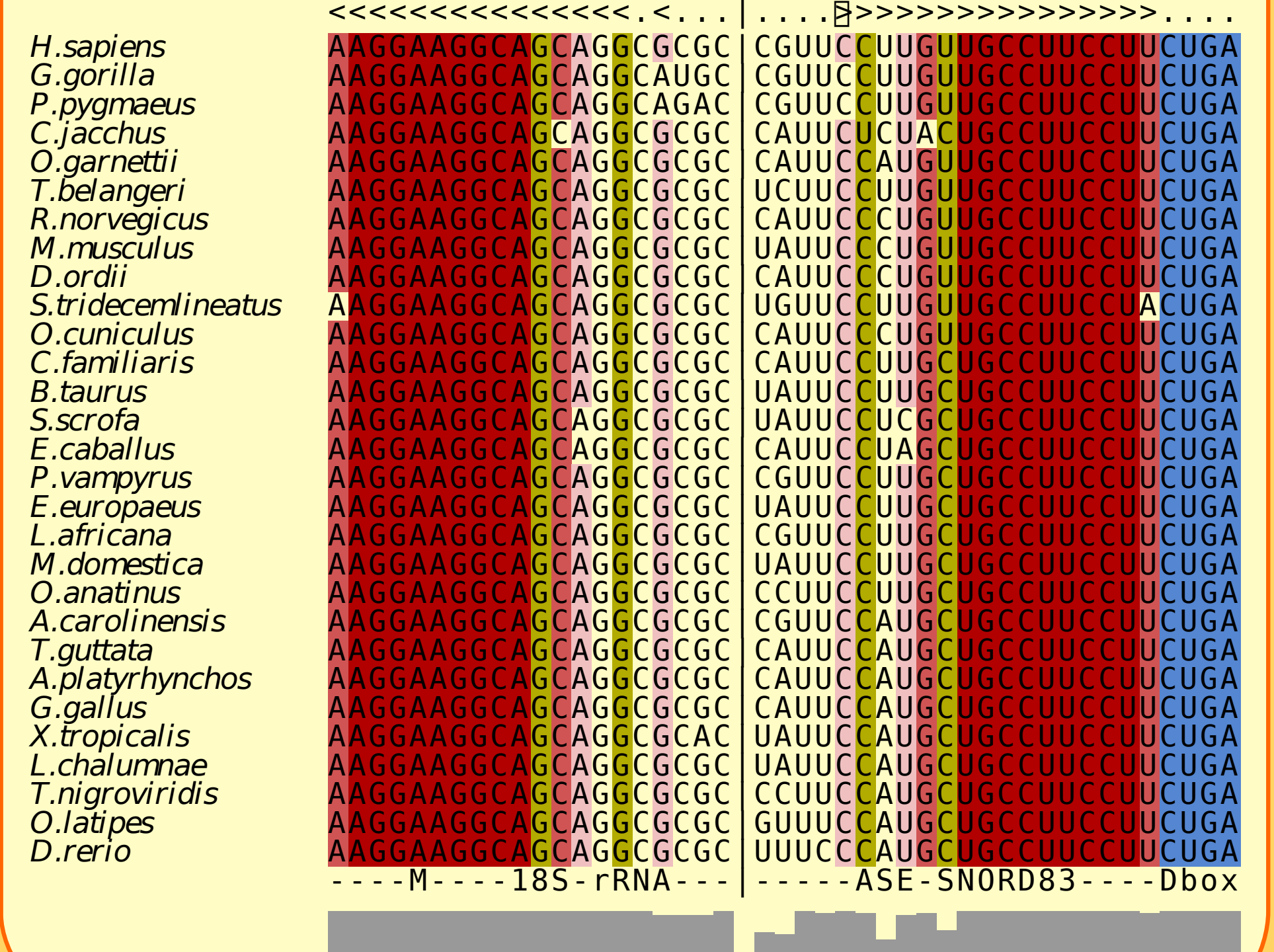


- conserved interactions
- redundant guides
- changeover of guide

Orphans and Modifications

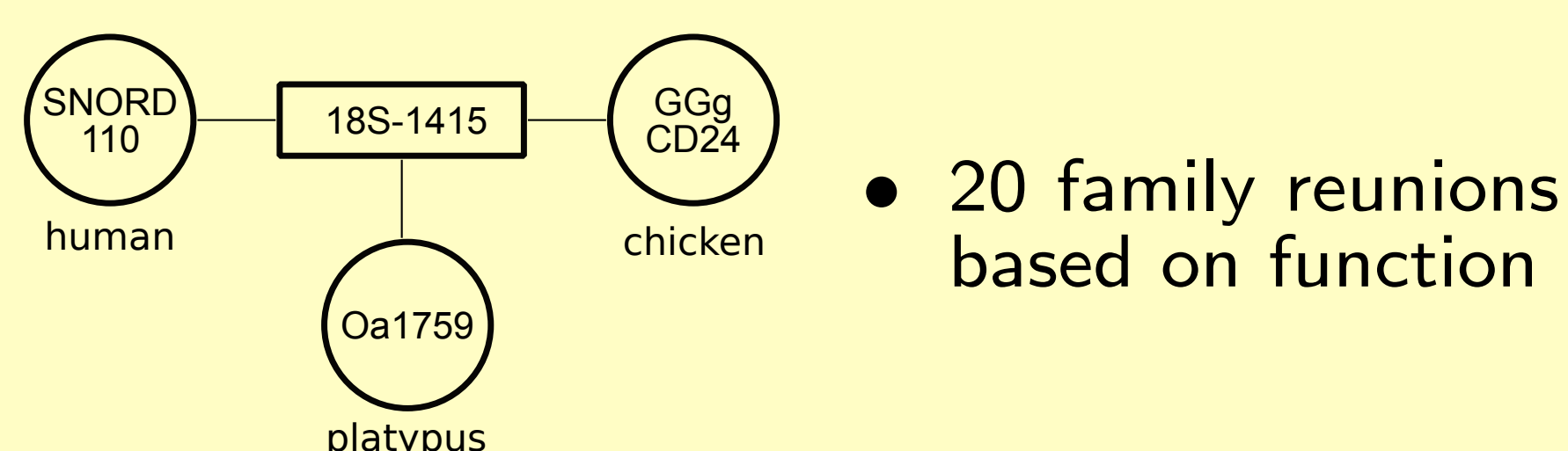


- high conservation at target binding sites
- compensatory mutations retain basepairing
- high entropy in rest of snoRNA sequence

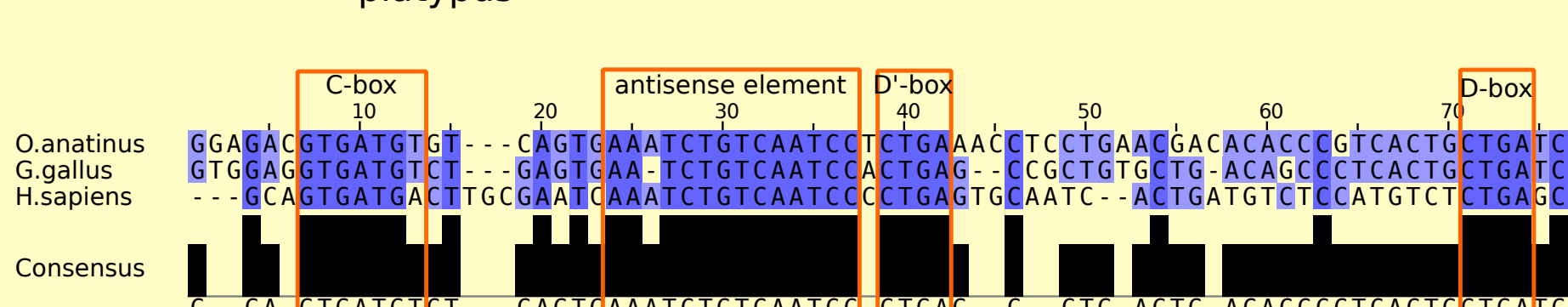


Distant Relatives

- sequence conservation to low for BLAST³
- identification of homologous function
- homology confirmed by alignment and conservation of antisense elements



- 20 family reunions based on function



References

1. Kehr S, Bartschat S, Stadler P F, and Tafer H. 2011 PLEXy: efficient target prediction for box C/D snoRNAs. *Bioinformatics*. 27:279-80
2. Tafer H, Kehr S, Hertel J, Hofacker I, and Stadler P F. 2010 RNAsnoop: efficient target prediction for box H/ACA snoRNAs. *Bioinformatics*. 26:610-6
3. Altschul S F, Gish W, Miller W and Myers E W, and Lipman D J. 1990 Basic local alignment search tool. *J Mol Biol*. 5:113

The work was funded by the European Union under the auspices of the FP-7 QUANTOMICS project.

*steffi@bioinf.uni-leipzig.de