

Homology Based Micro RNA Prediction



H. Alexander Ebhardt¹, Yifeng Liu² and Duane Szafron²

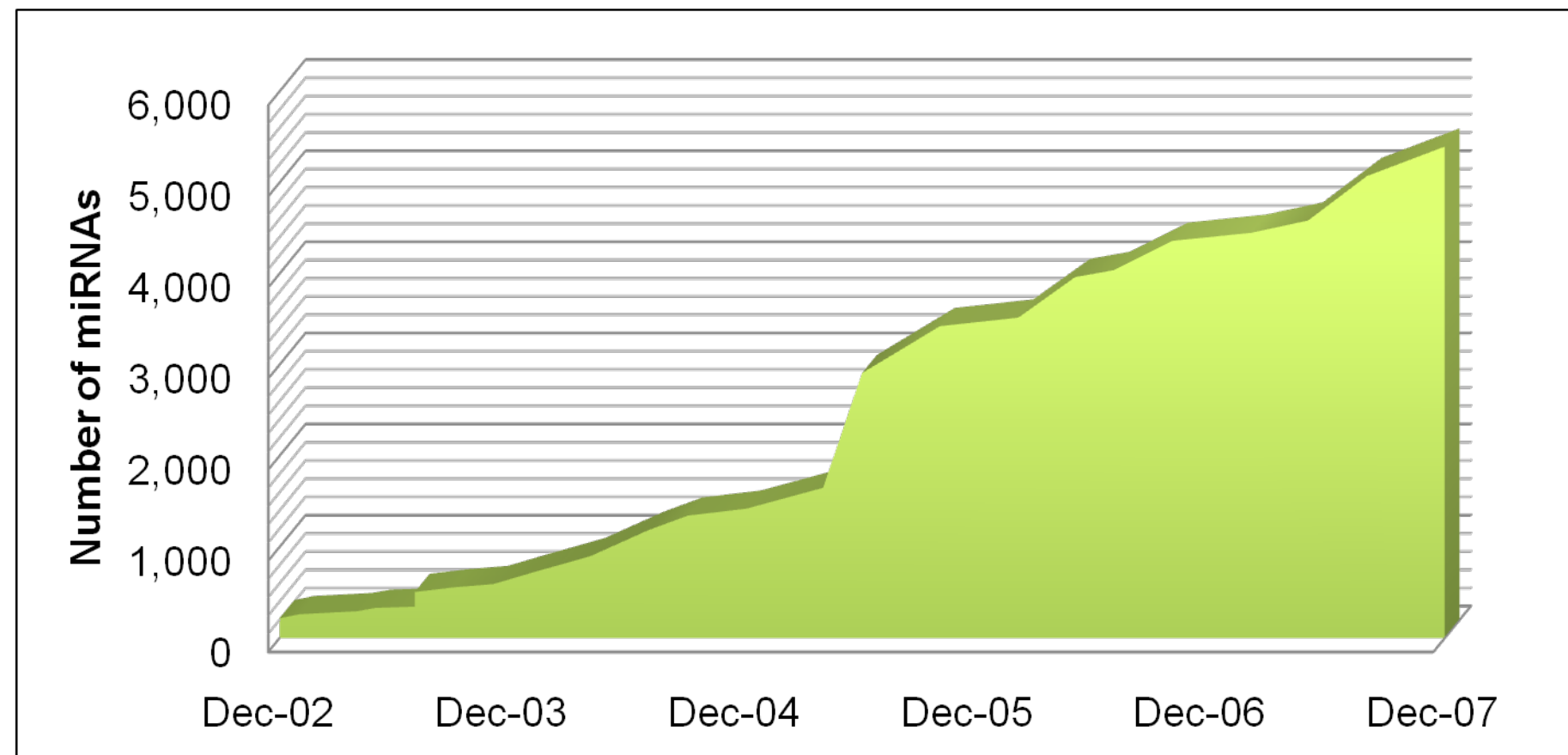
¹Department of Biochemistry, ²Department of Computing Science

*ebhardt@ualberta.ca



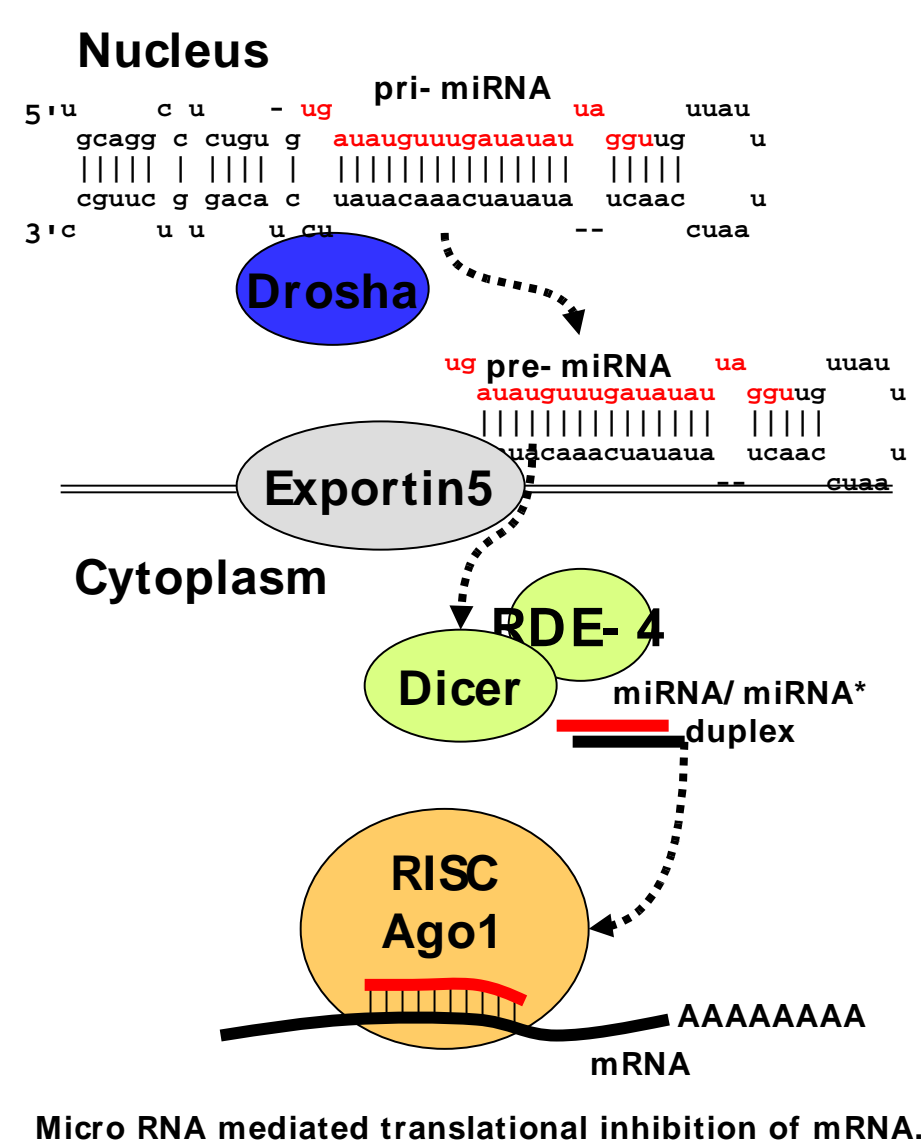
Abstract

- Growing number of confirmed micro RNAs (miRNAs) reported in miRBase (see below)



- Goal: automated prediction of miRNAs based on homology with high accuracy.**
- Presented here is our algorithm together with independent validation.

Background: micro RNAs



Premise

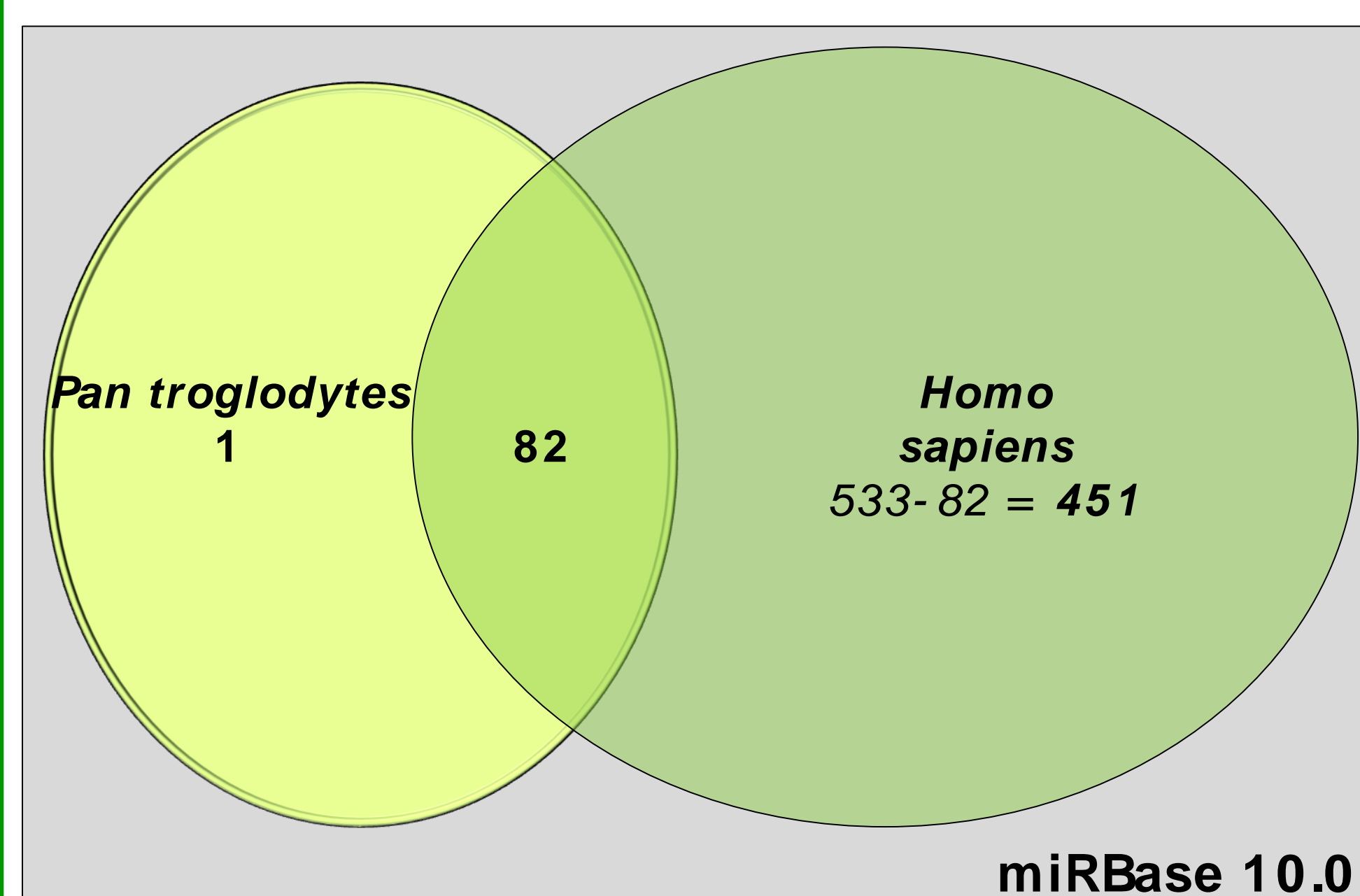


Figure 1: Flowchart

Flowchart of automated miRNA prediction algorithm.

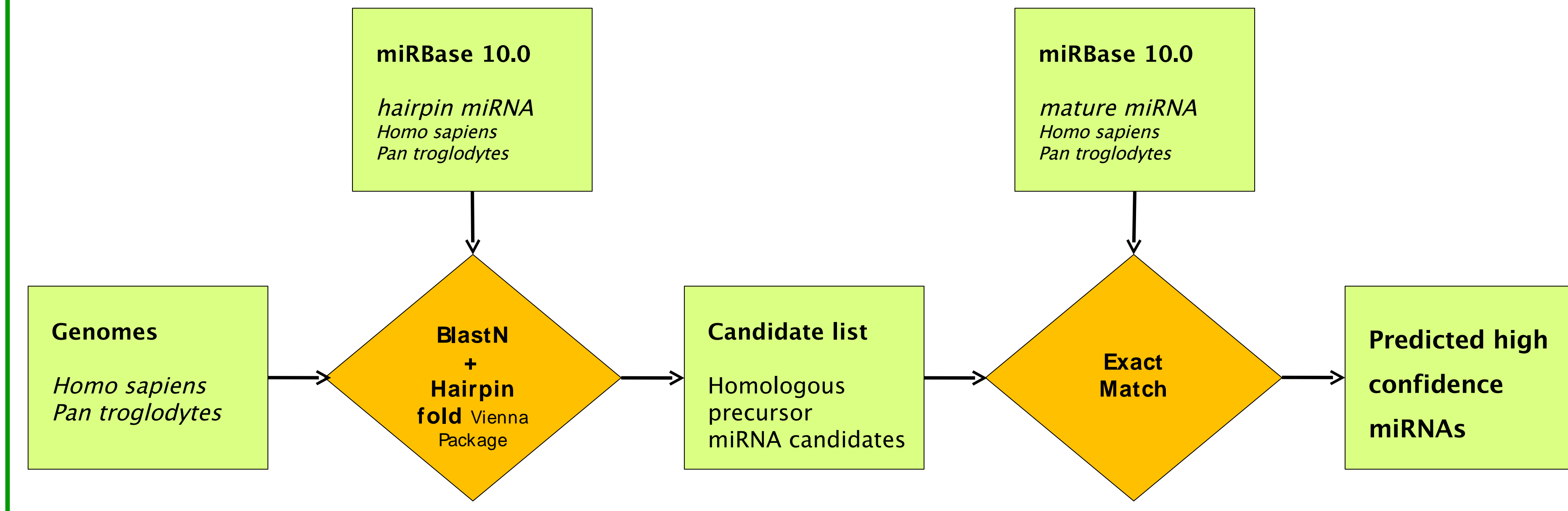
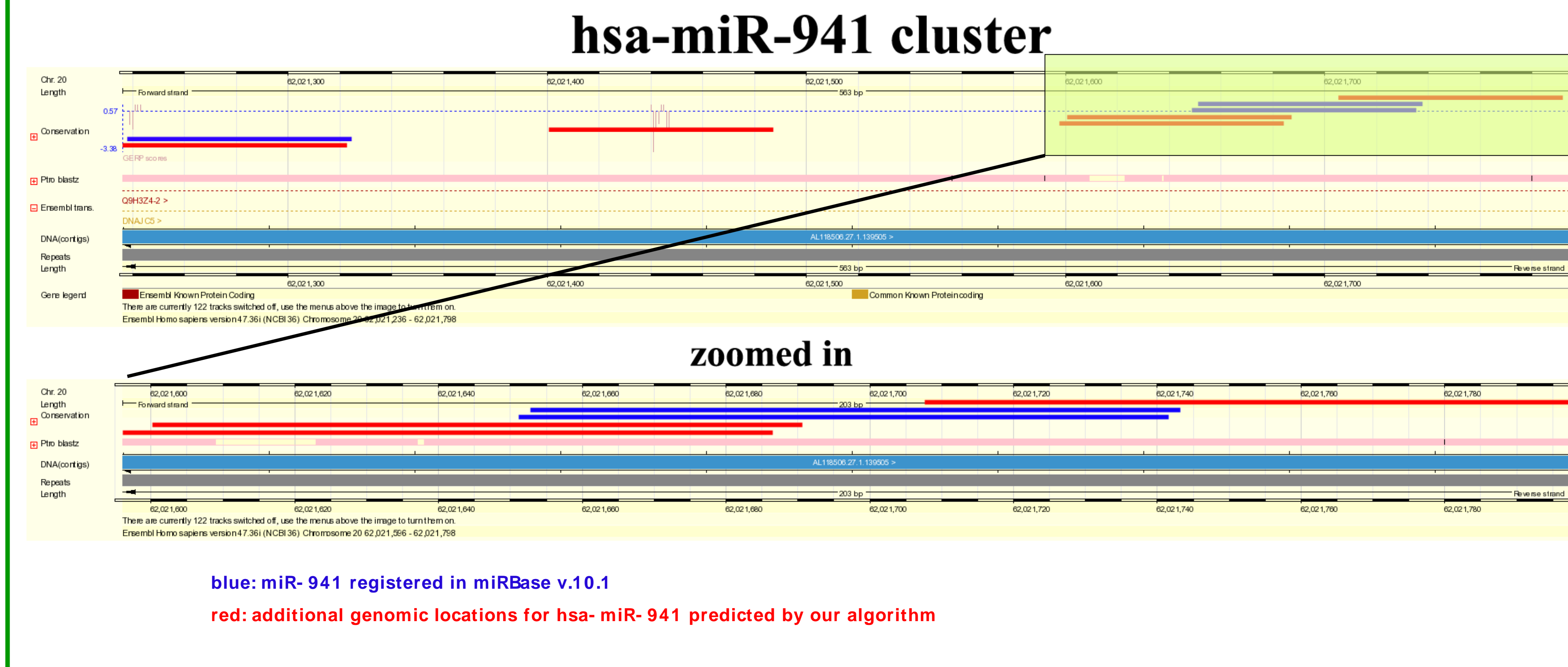


Figure 2: Results - hsa-miR-941

Our algorithm predicted overlapping genomic location of miR-941 in sub-telomeric region of chromosome 20 not recorded in current miRBase v.10.1.



Algorithm

Figure 1: Flowchart

- pri-miRNAs were collected from miRBase 10.0
- pri-miRNAs were compared to *hsa* and *ptr* genomes using BlastN and potential candidates folded into hairpins by RNAfold
- A list of candidate pre-miRNA was compiled and matched to a list of mature miRNAs (miRBase)
- Only pri-miRNA candidates that contain mature miRNAs made up final list of candidates

Advantage of procedure:

- fully automated
- low false positive predictions (conservative nature of prediction algorithm)

Results

- We predicted 483 unique miRNA sequences in *ptr* based on homology to validated miRNAs that were not recorded in miRBase 10.0.

Figure 2: Results - hsa-miR-941

- multiple locations of hsa-miR-941 precursor hairpin and mature miRNA not recorded in current miRBase registry
- hsa-miR-941 is located in sub-telomeric region of chromosome 20
- pre-miRNA overlap partially
- over a span of 600 nts 8 potential pre-miRNA hairpins

Validation of Results

Figure 3: Validation of results

- Our algorithm developed with miRBase 10.0
- miRBase 10.1 contained *Pan troglodytes* miRNAs that were predicted by our algorithm using miRBase 10.0
 - 11 *ptr-miR* predicted and confirmed
 - 2 *ptr-miR* not predicted due to stringency parameters in algorithm
 - 2 *ptr-miR* found pre-miRNA, but not exact match
 - 2 *ptr-miR* did not have homologues in 10.0
- Ongoing: conformation of prediction by Northern hybridization.

Figure 3: Independent validation of our results

Our algorithm was based on miRBase version 10.0. With the publication of miRBase version 10.1, a subset of our predicted miRNAs were validated by other researchers.

