DE-NOVO CANONICAL MIRNAS

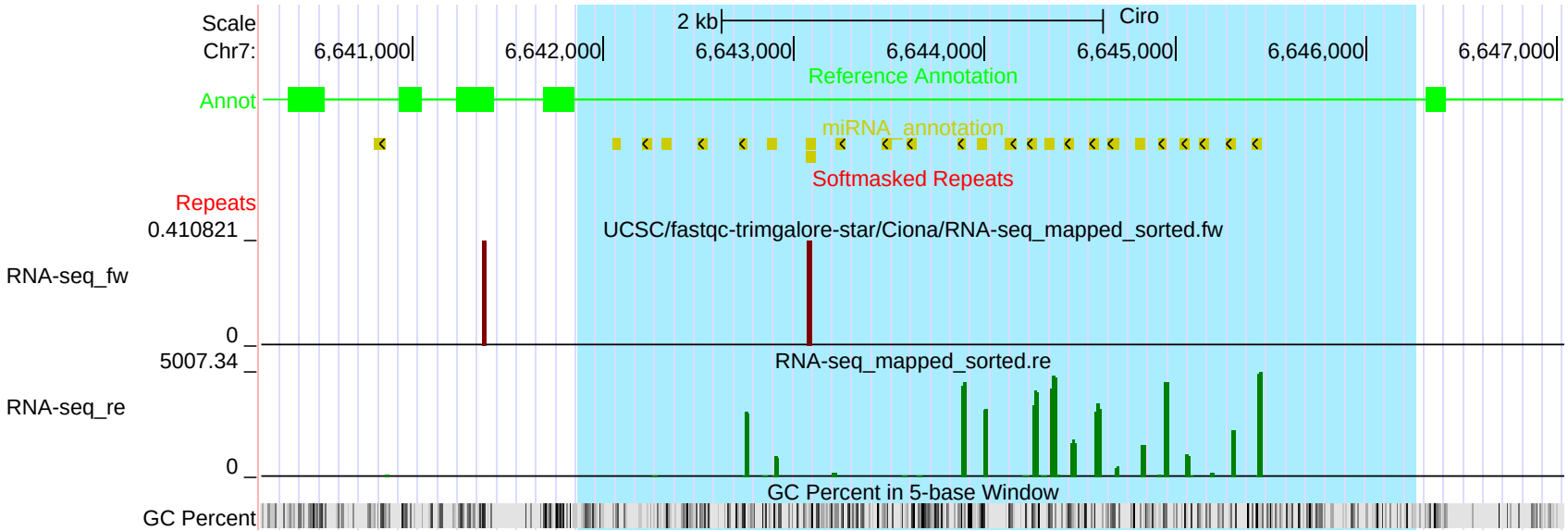
A PILOT STUDY INCLUDING ITERATIVE MODELS

CRISTIAN A. VELANDIA H
TBI
UNIVERSITY OF VIENNA

cavelandiah@tbi.univie.ac.at

microRNAs on *Ciona robusta*

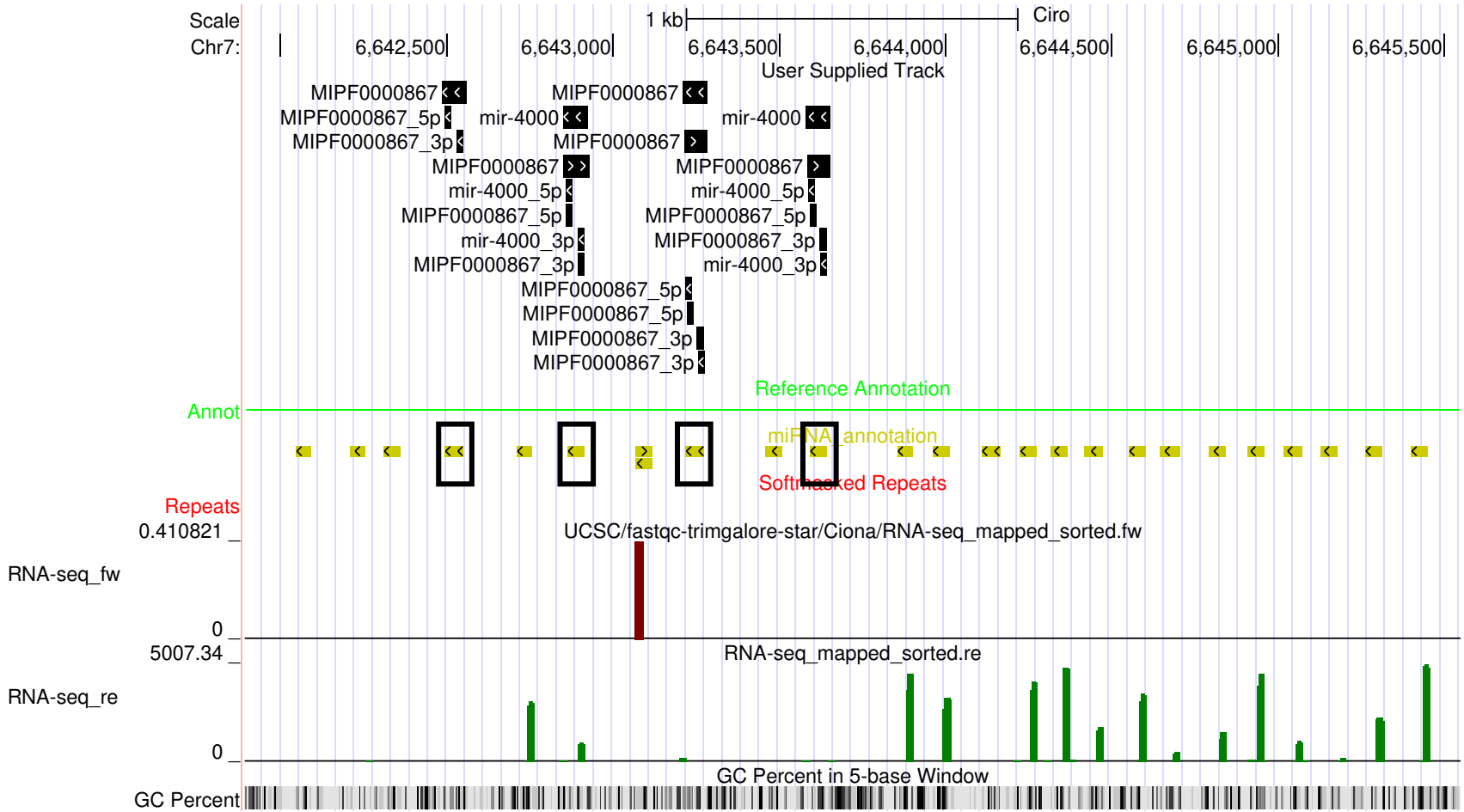
Chromosome 7 cluster (28 miRNA loci over 3713 nt)



`miRBase` annotations without 'miR-family' classification.

microRNAs on *Ciona robusta*

Homology annotation using `miRNAture` + `miRBase`: low coverage.



Methods: *De novo* miRNA annotation in *Ciona robusta*

Table: Available SRA small-RNA/miRNA-seq for *C. robusta*

Run	**Experiment**
SRP002173	small-RNA for two developmental stages
SRP079886	Expression Oral Siphon Regeneration
SRP116990	small-RNA transgenic ascidians

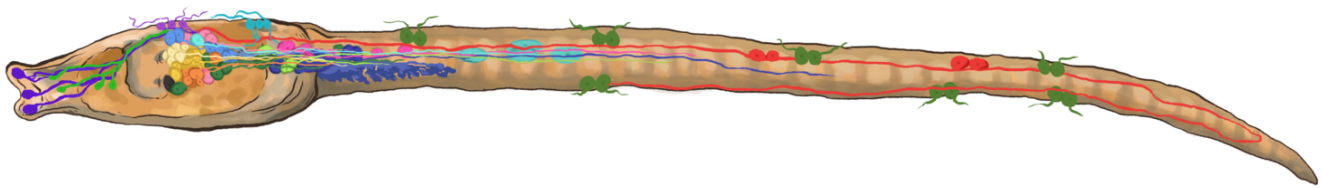
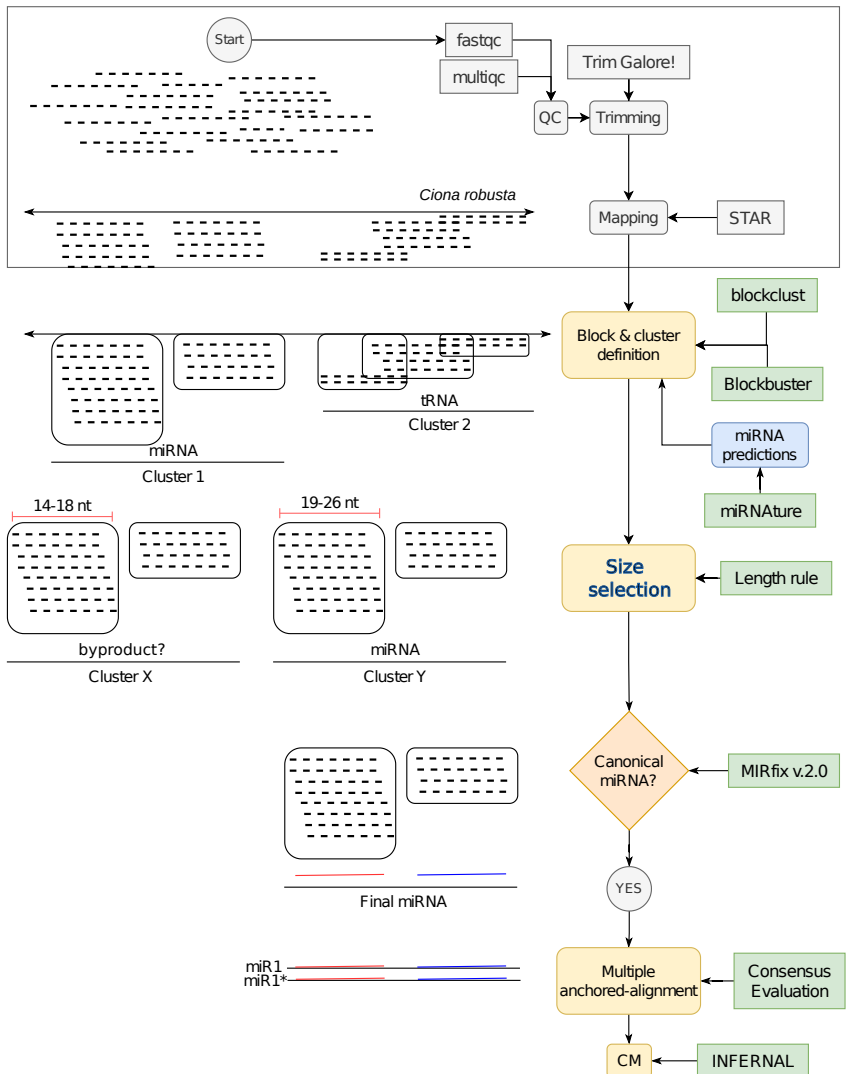
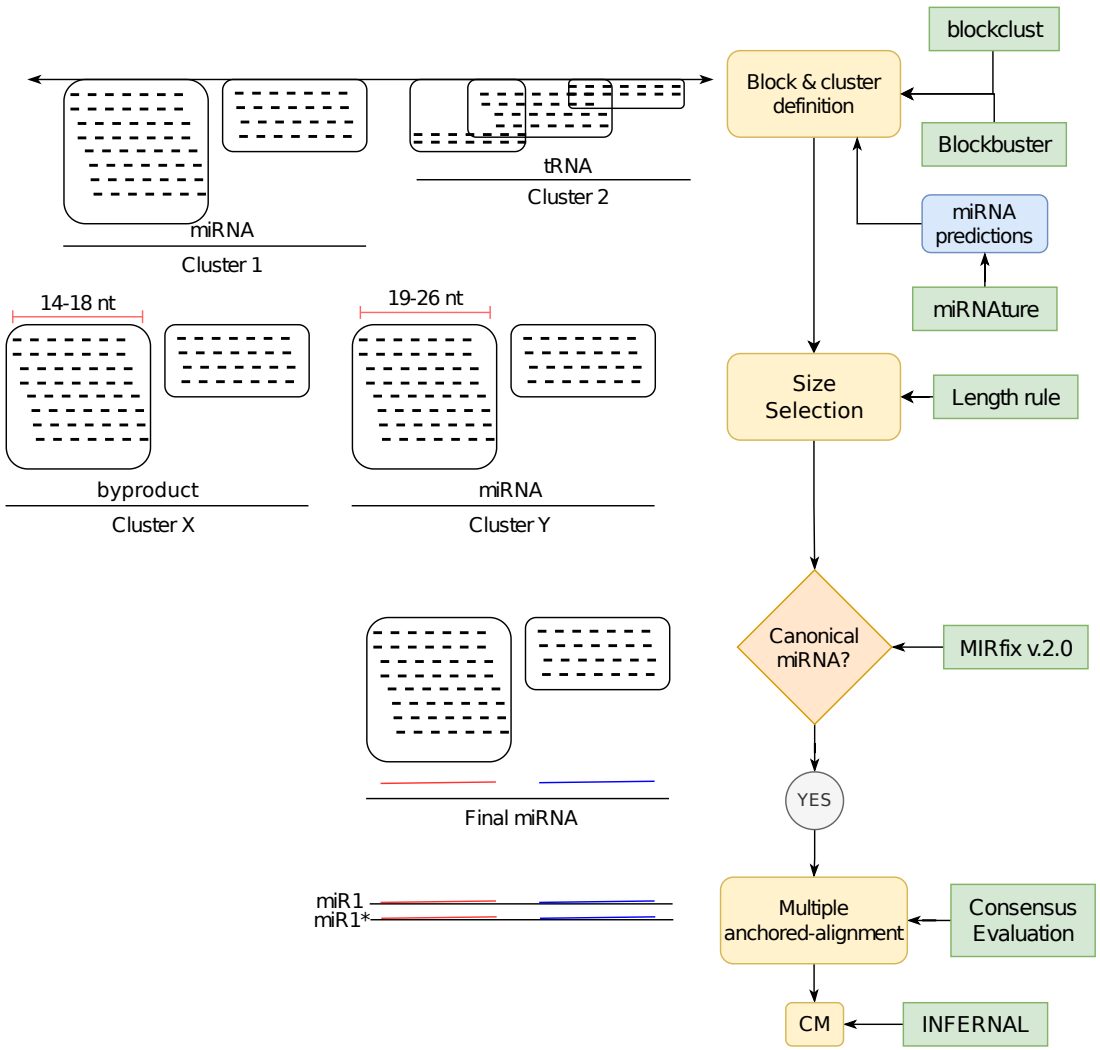


Figure: *Ciona robusta*. Source: Lindsey Leigh, 2017

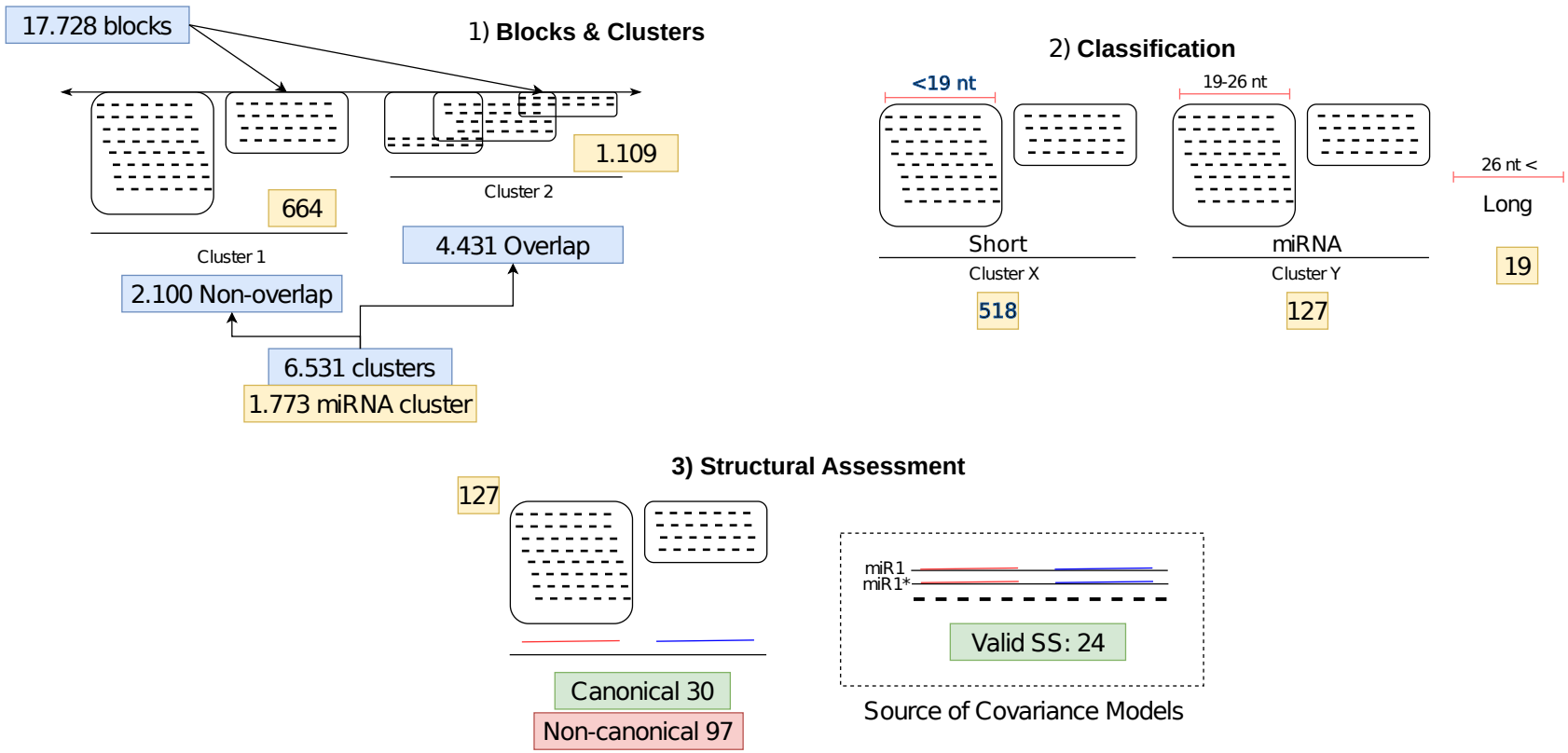
Methods: processing patterns → miRNAs



Methods: processing patterns → miRNAs

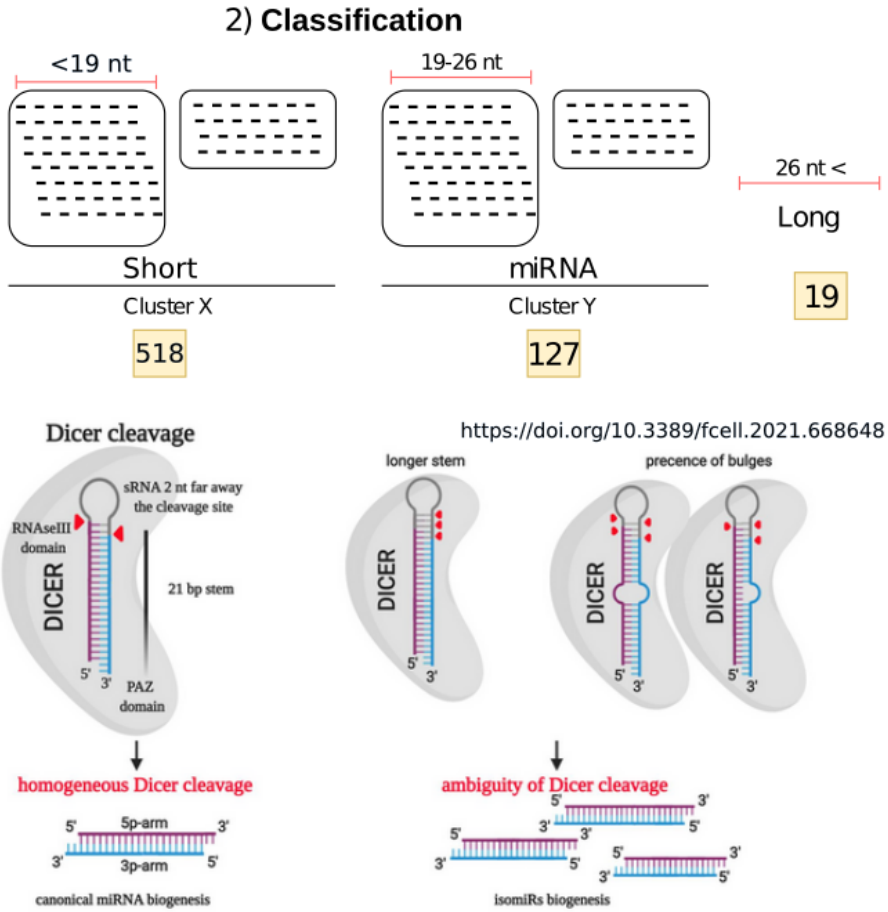


Results: Canonical refining



- Only 3.6% candidates detected as *canonical* miRNAs.

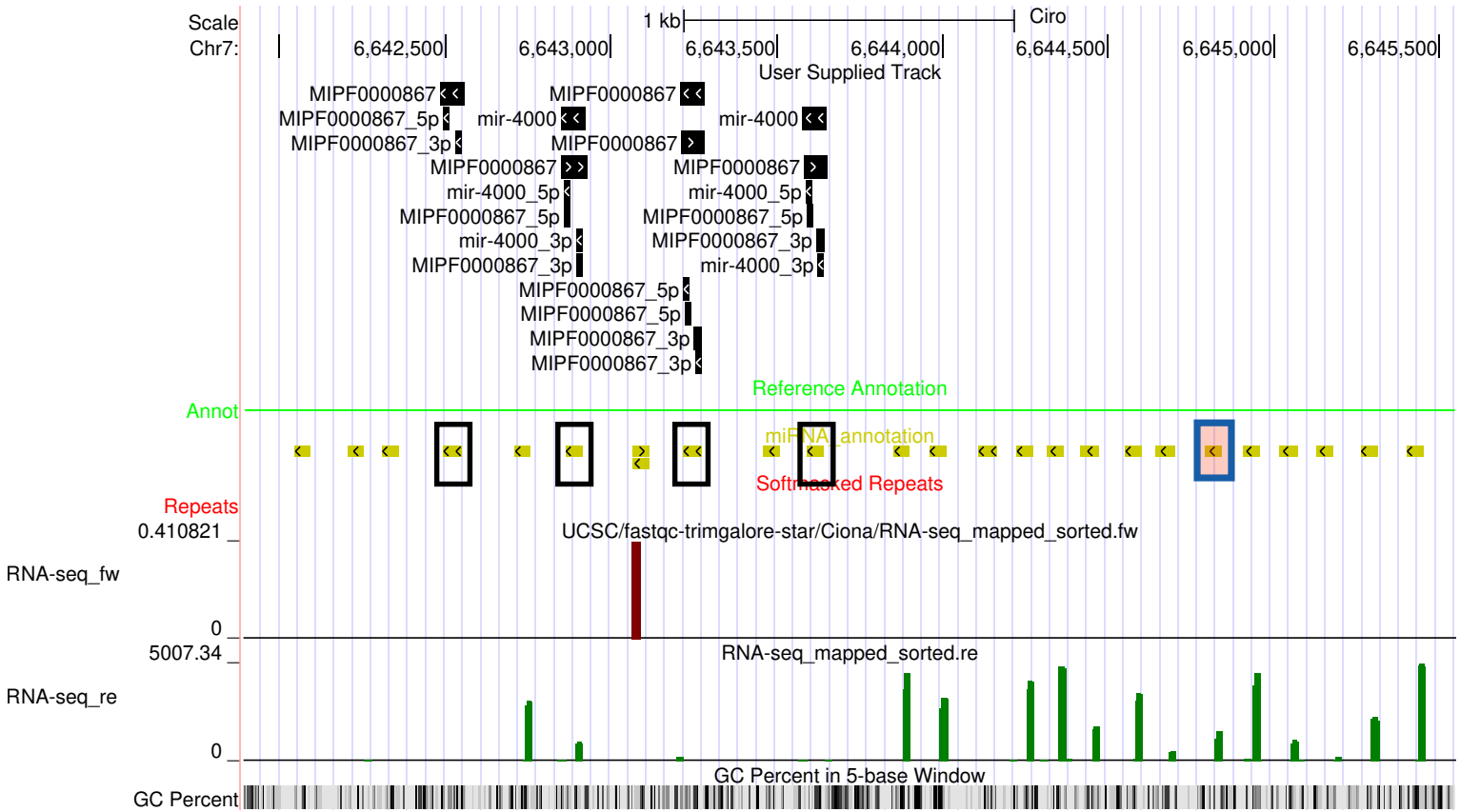
Results: Canonical refining



- Mostly short candidates: maturation by-products?/isomirs? ($\sim 77.4\%$)

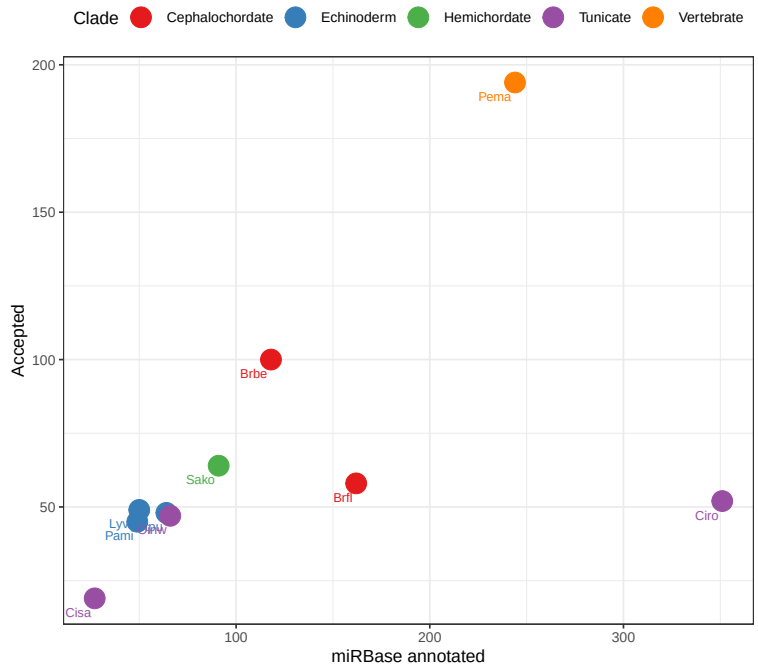
Results: Going back to Chr7 cluster

Canonical de-novo families: One additional candidate on Chr7 : (



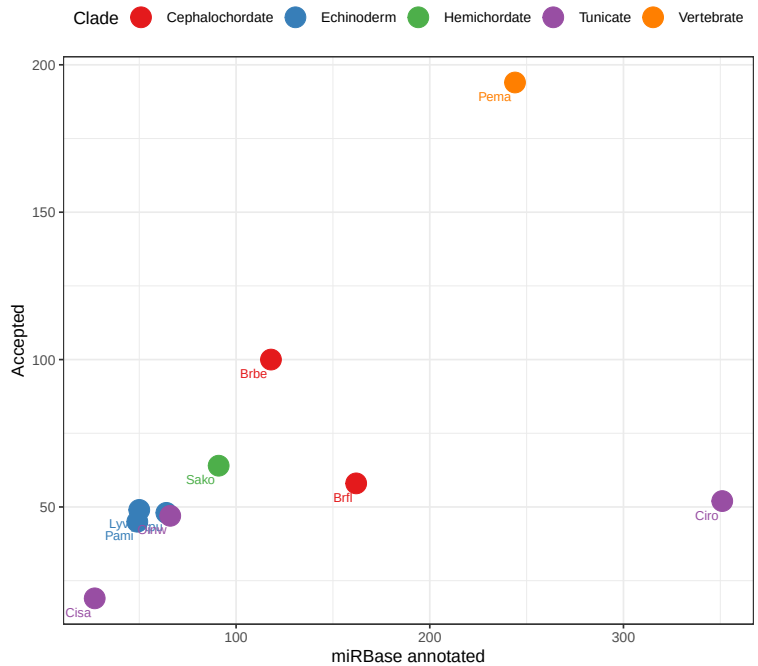
Why $\sim 82\%$ loci were not detected?

Results: miRNA annotation on chordates



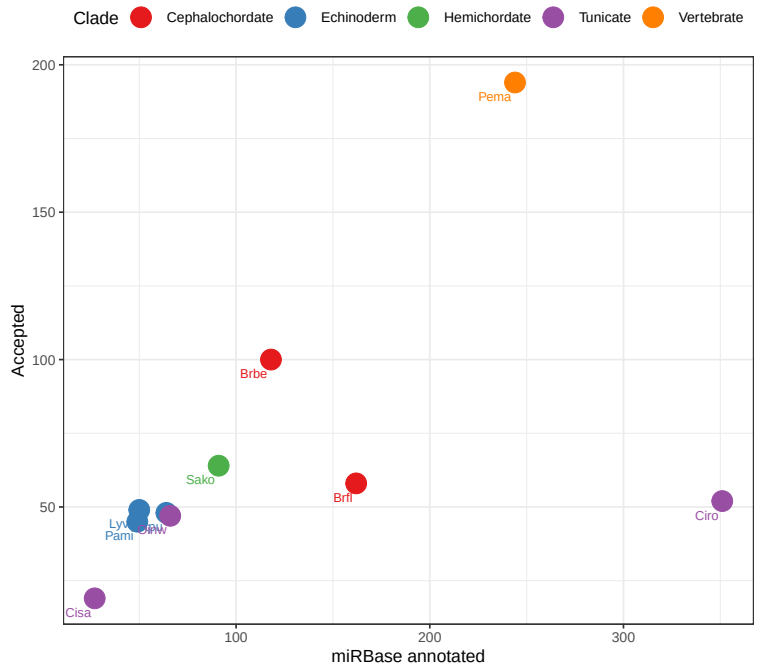
- *C. robusta* annotations: 351 loci

Results: miRNA annotation on chordates



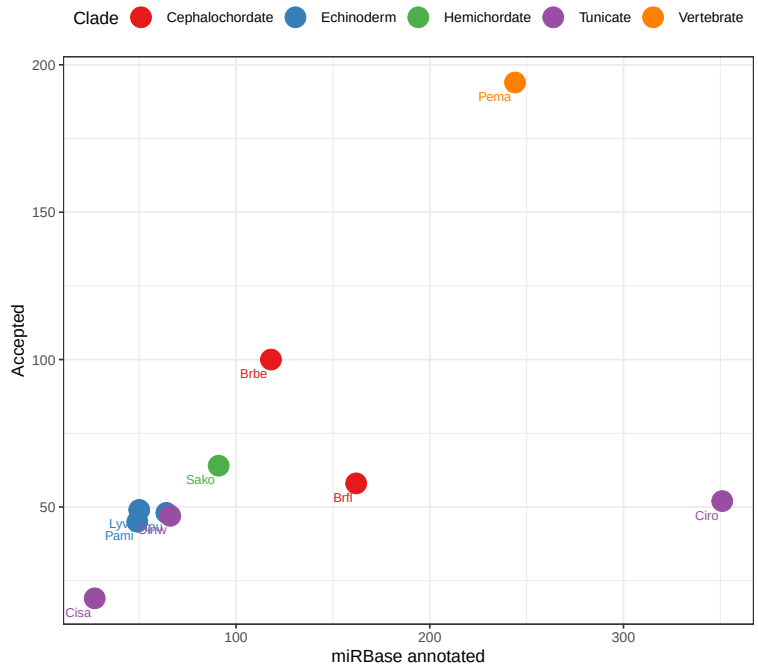
- *C. robusta* annotations: 351 loci
- With Fam.: 18.5% / No Fam: 81.5%.

Results: miRNA annotation on chordates



- *C. robusta* annotations: 351 loci
- With Fam.: 18.5% / No Fam: 81.5%.
- 13× than sister specie: *Ciona savignyi*

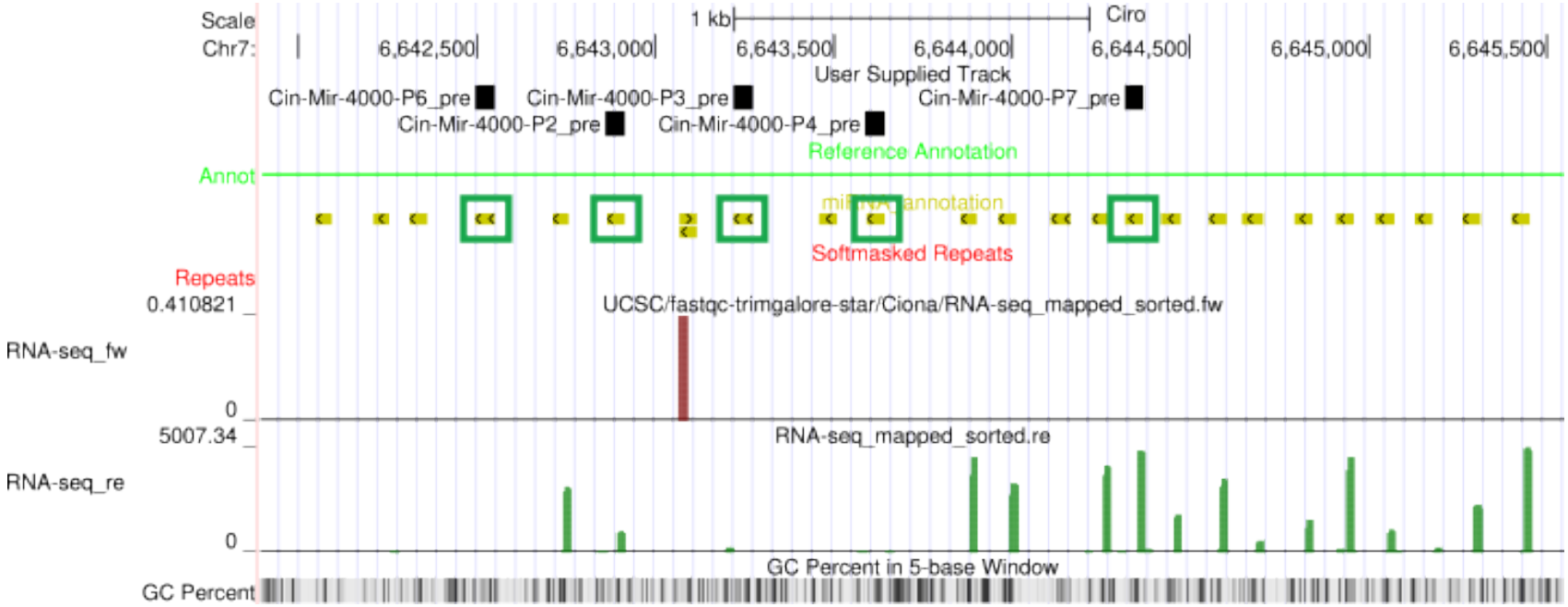
Results: miRNA annotation on chordates



- *C. robusta* annotations: 351 loci
- With Fam.: 18.5% / No Fam: 81.5%.
- 13× than sister specie: *Ciona savignyi*
- Higher loci number than coelacanth and lancelet (!).

Results: Discovering canonical miRNAs

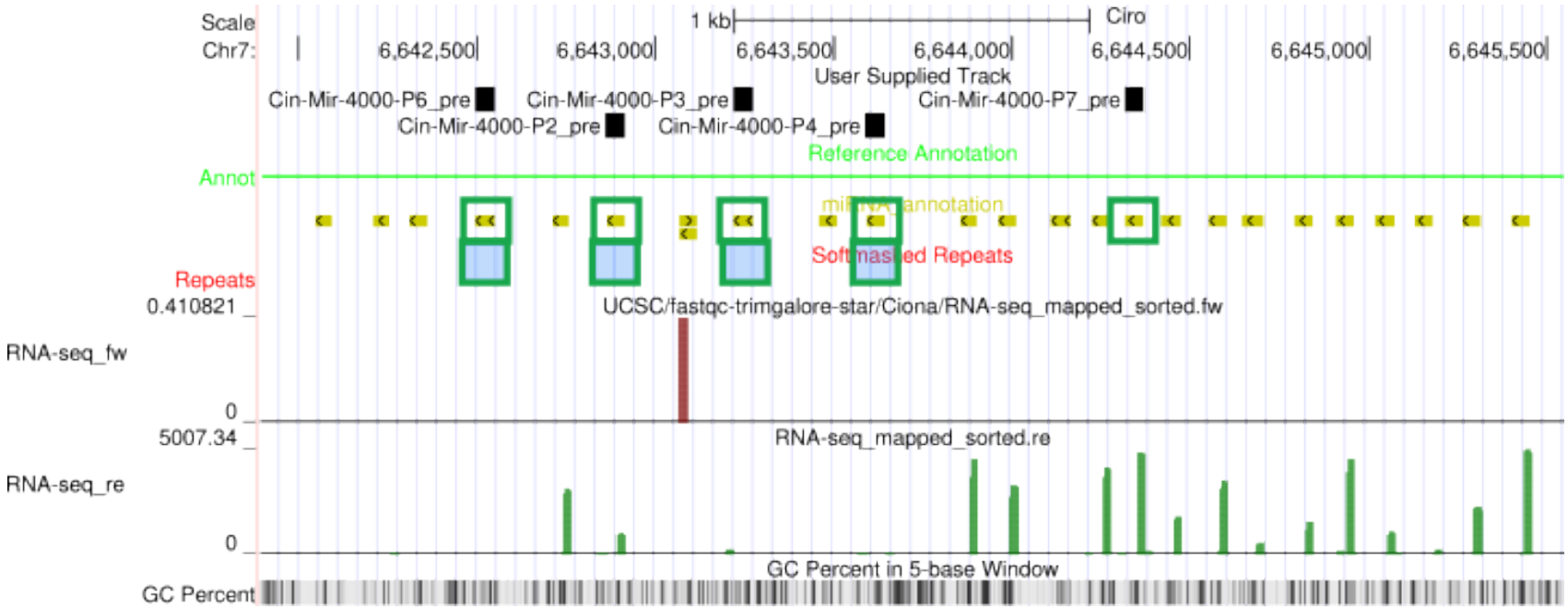
Finding canonical de-novo families: Structural assessment



- Canonical miRNAs represent $\sim 21\%$ of cluster.

Results: Discovering canonical miRNAs

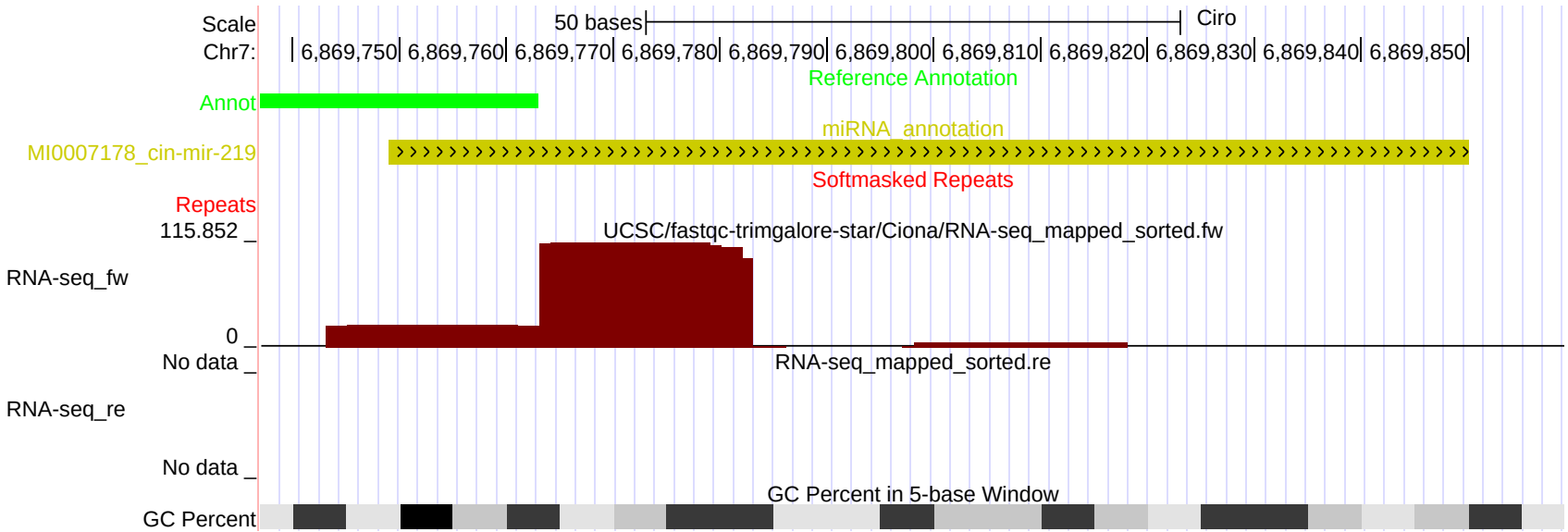
Finding canonical de-novo families: Structural assessment



- Canonical miRNAs represent $\sim 21\%$ of cluster.
- Reasons: Position of mature sequences on loop regions, structural constrains, miRNA-offset.

Results: Discovering canonical miRNAs

Finding canonical de-novo families: Structural assessment



- Canonical miRNAs represent $\sim 21\%$ of cluster.
- Reasons: Position of mature sequences on loop regions, structural constrains, miRNA-offset.

Conclusions

- Non-trivial miRNA annotation (i.e simplified chordate $\sim$ 120 Mb genome).
- Need to extend `miRNAture` & `MIRfix` to non-canonical miRNAs/expression products associated to microRNA maturation.
- Iterative models for miRNAs: expression patterns (blocks, clusters), mature definition, mature position, structural assessment, evolutionary information (+ synteny?).

Acknowledgements

- Prof. Dr. Peter F. Stadler.
- Dr. Jörg Fallmann.
- To your attention.

Appendix: miRNA annotation on chordates

Source	Specie	Ann.	Supp.	Novel	Mapped	Filt.	Accepted
miRBase	*S. kowalevskii*	91	83	0	72	8	64 (70.3%)
	*S. purpuratus**	64	53	6	54	6	48 (75.0%)
	*P. miniata**	49	58	0	51	6	45 (91.8%)
	*B. floridae**	162	90	0	67	9	58 (35.8%)
	*C. robusta**	351	141	14	130	78	52 (14.8%)
	L. variegatus	50	NA	NA	50	1	49 (98.0%)
	B. belcherei	118	NA	NA	110	10	100 (84.7%)
	C. savignyi	27	NA	NA	19	0	19 (70.4%)
	O. dioica	66	NA	NA	47	0	47 (71.2%)
	*P. marinus**	244	NA	NA	238	44	194 (79.5%)