Bioinformatic and experimental identification of microRNA in *Phytophthora infestans*

Guohong Cai, Lalit Ponnala, Qi Sun,
Jarek Pillardy, Bill Fry

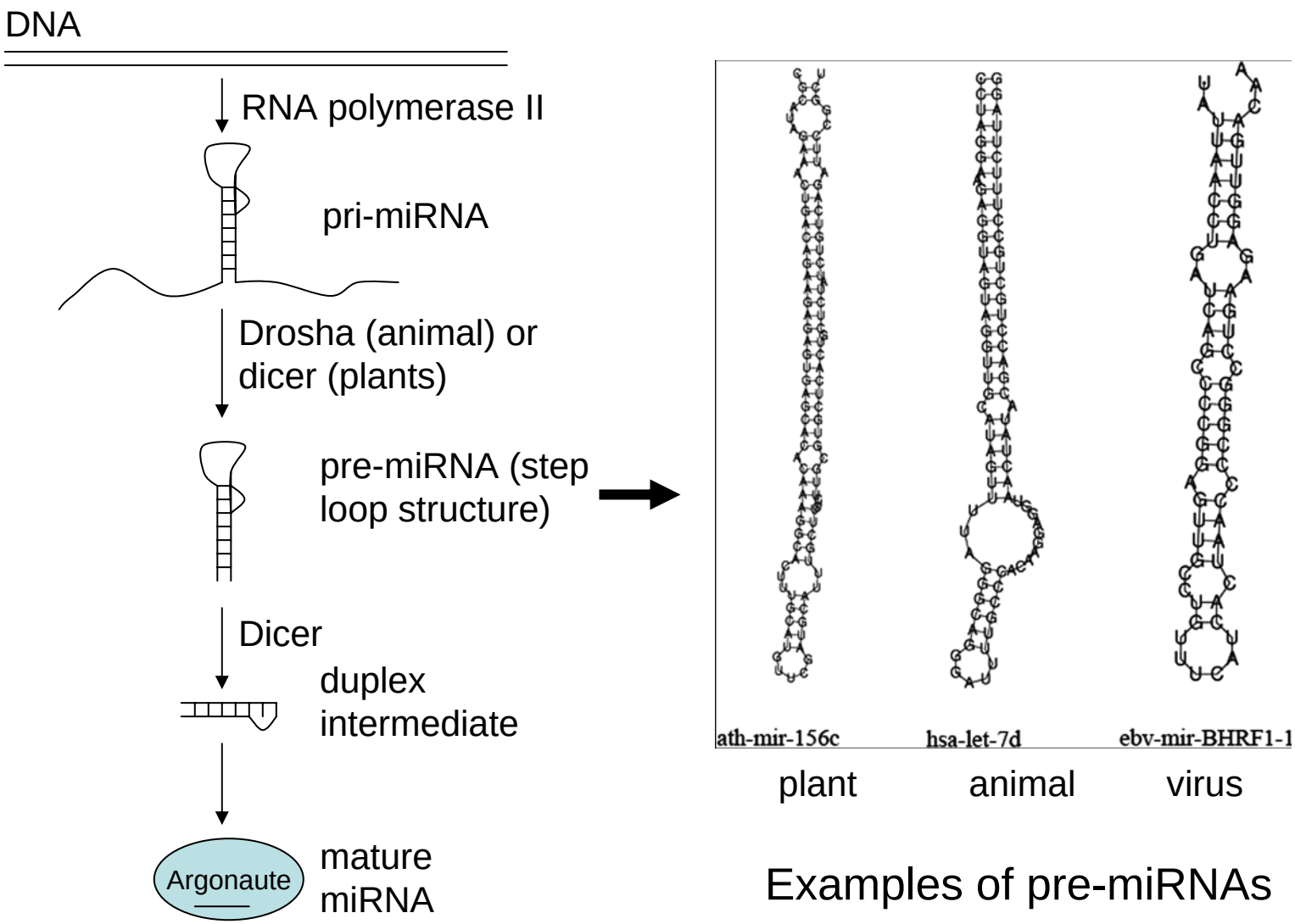
Department of Plant Pathology and
Center for Advanced Computing
Cornell University, Ithaca NY

Background:

MicroRNAs (miRNAs) are:

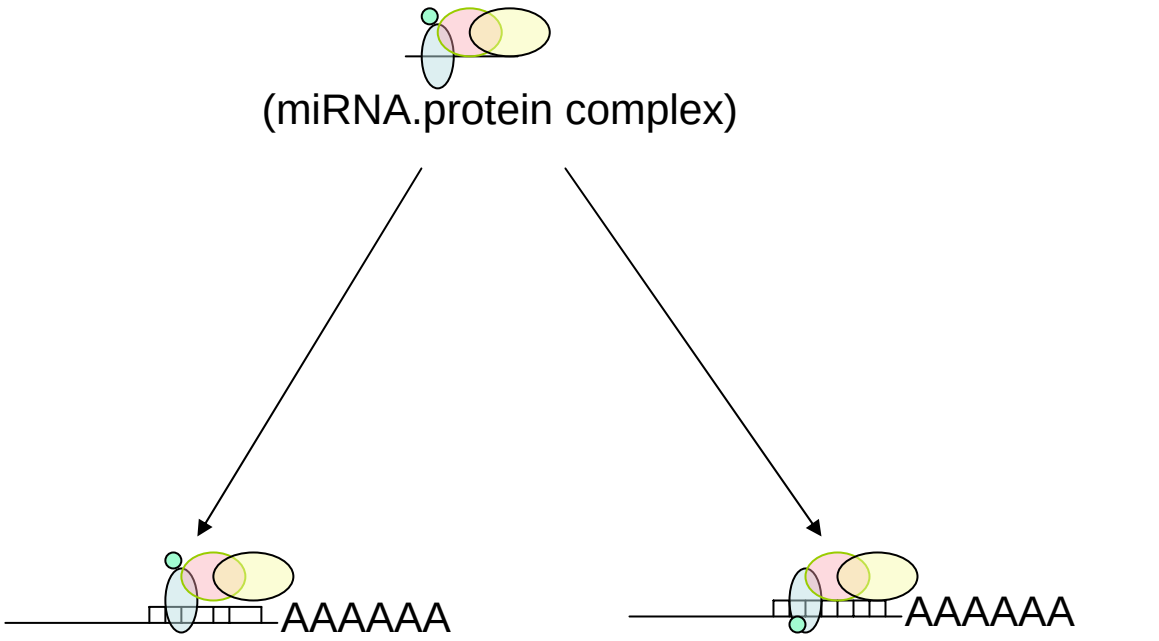
- Endogenous, small, noncoding RNAs;
- Typically 19-25 nt in length;
- The first member (lin4) identified in 1993 (Lee et a Cell75:843-854), but the term was introduced in 2001.
- Have been found in plants, viruses and animals.
- By August 2007, 5,071 miRNAs have been reported (http://microrna.sanger.ac.uk/sequences/).

Biogenesis of miRNAs



Functions of miRNA

Mechanisms:



Translation repression of mismatched mRNA

Degradation of mRNA with sufficient complemetarity

Functions:

Animals: Development, cancer and other disease pathogenesis;

Plants: Development, phase change (vegetative to reproductive), stress response;

Viruses: Regulation of virus and/or host gene expression, evasion of host immune system (e.g. human cytomegalovirus).

The importance of miRNA in gene regulation leads to explosion of studies trying to identify miRNAs in various organisms

Experimental approach:

Direct cloning and sequencing of small RNAs.

Difficult to find low abundant, tissue- or stage-specific miRNAs

Contamination of degrade mRNA and other non-coding RNAs.

Bioinformatic approach:

Rely on miRNA characteristics:
(1) conservation of miRNAs in related species;
(2) formation of stable stem–loop structure by pre-miRNAs;
(3) the presence of mature miRNAs in the stem and not in the loop of pre-miRNAs.

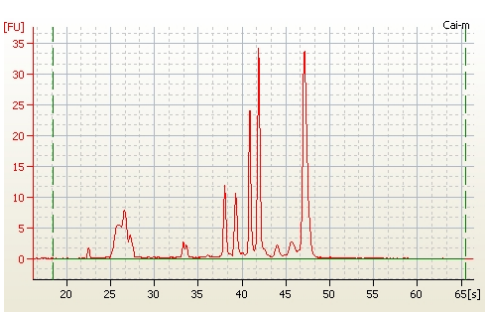
----Need experimental confirmation

These two approaches complement each other.

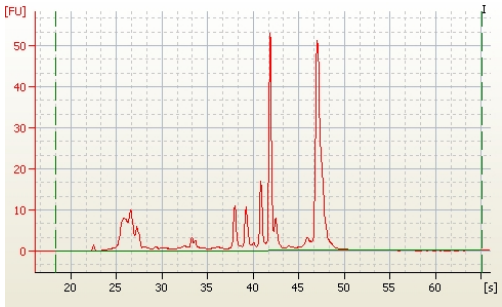
Our approaches:

I. Experimental approach

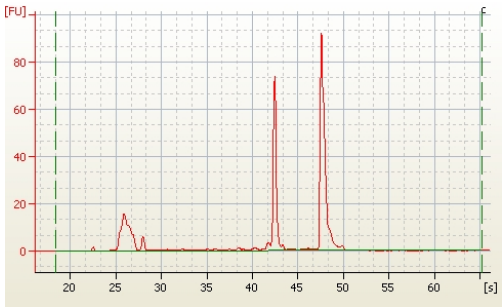
- Total RNAs were extracted from:



Mycelium from liquid culture



P. Infestans infected potato plants (no sporulation)



Healthy potato plants

- Small RNA libraries are being made at Biology Research Center in Cornell

- The libraries will be sequenced with Solexa machine. Results expected in 2-3 weeks.

With this approach we intend to identify:

•miRNAs in *P. infestans;*

•changes in miRNA profiles between mycelium in liquid medium and those *in planta*;

•miRNAs in potato;

•Changes in miRNA profiles between healthy and late blight infected potato.

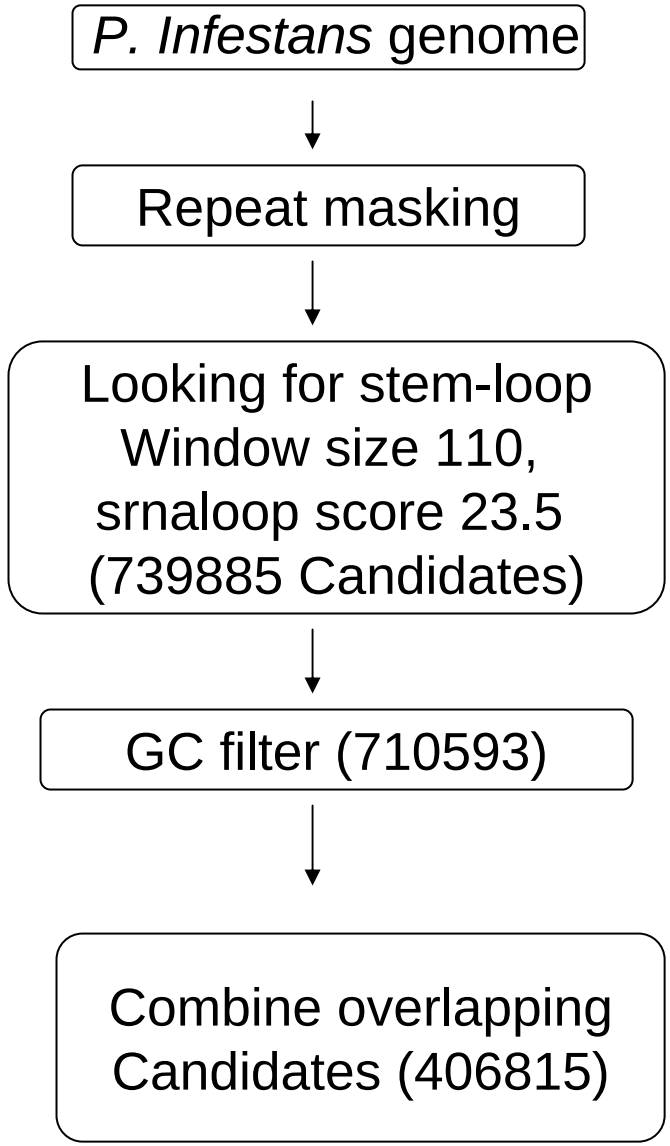
II. Bioinformatic approach:

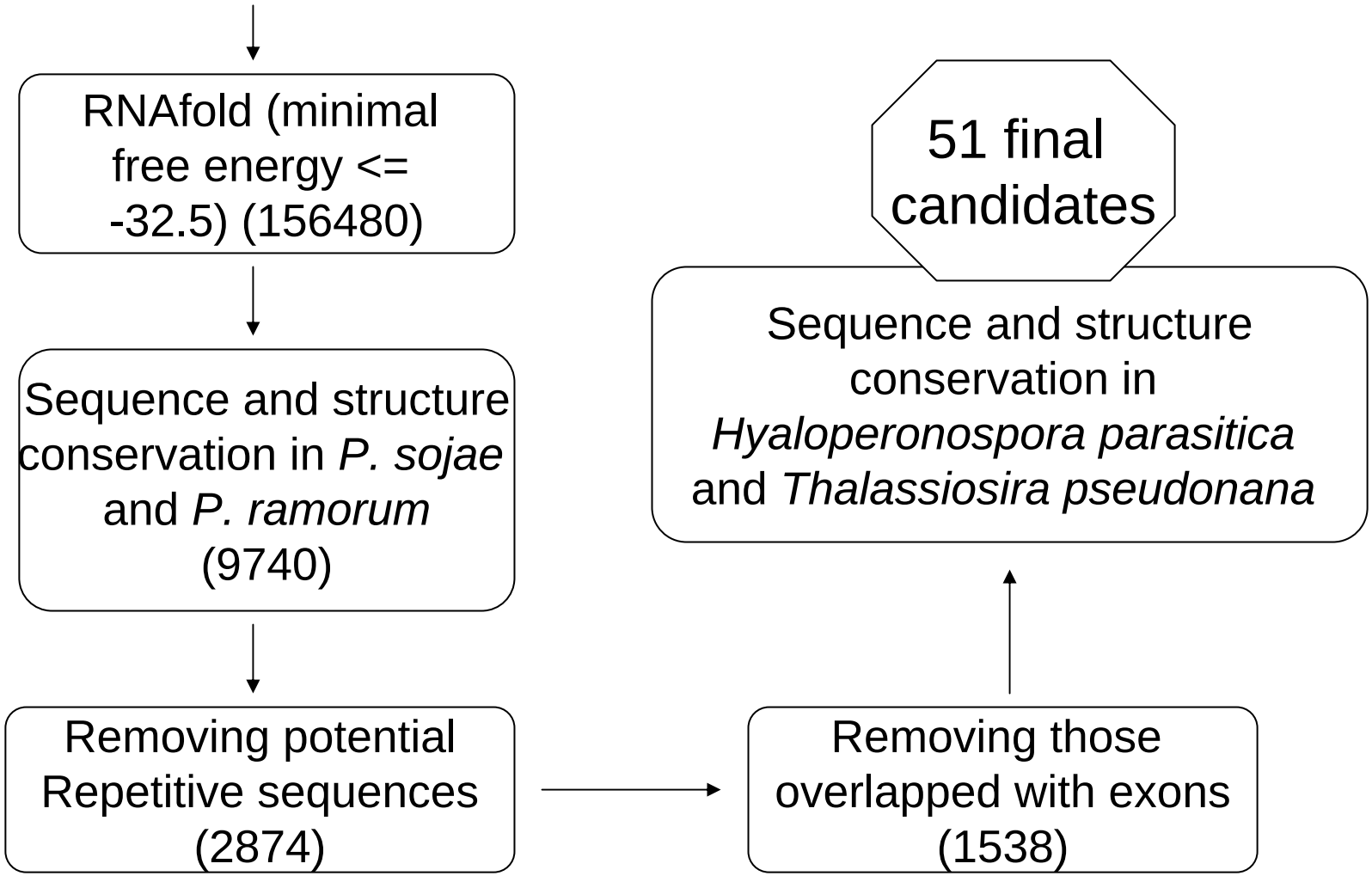
In collaboration of Computational Biology Service Unit of Cornell Theory Center:

<u>Lalit Ponnala</u>

Qi Sun

Jarek Pillardy



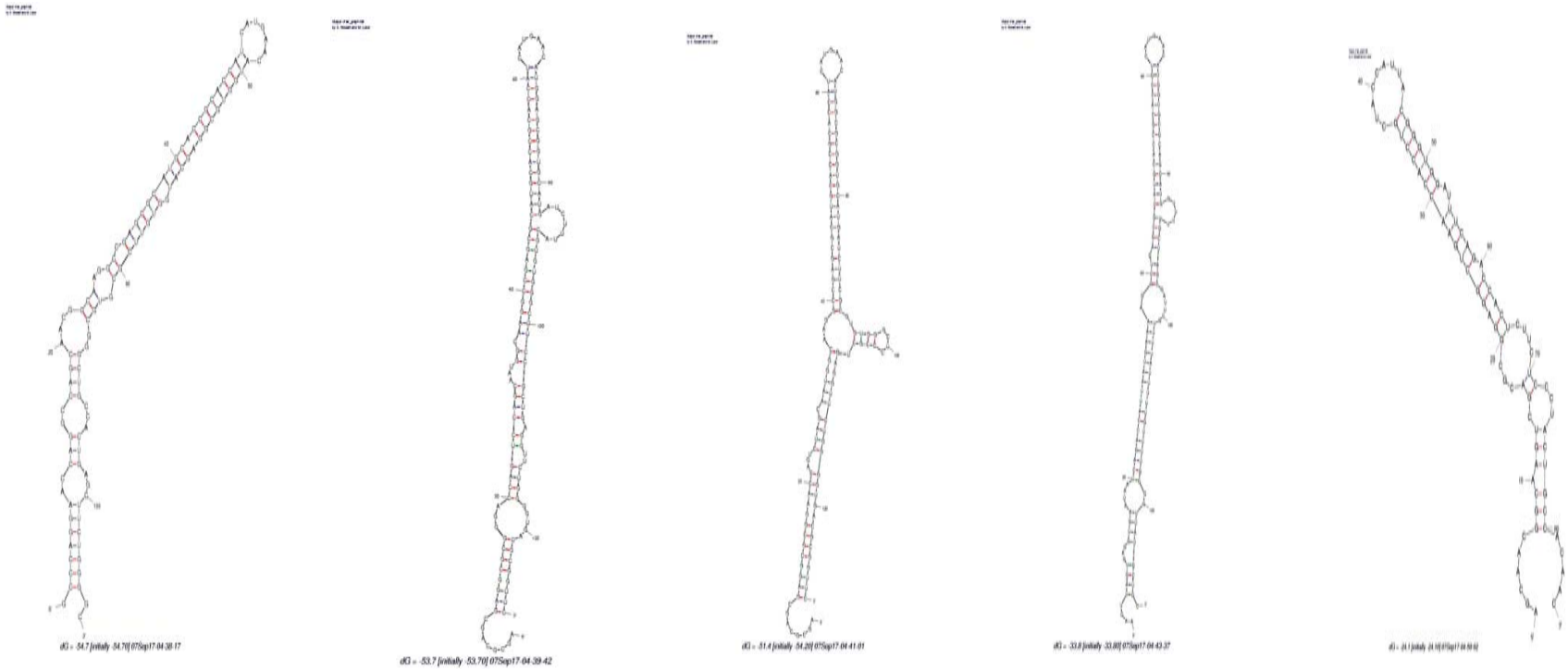

RNAfold (minimal free energy <= -32.5) (156480)
Sequence and structure conservation in *P. sojae* and *P. ramorum* (9740)
Removing potential Repetitive sequences (2874)
Removing those overlapped with exons (1538)
Sequence and structure conservation in *Hyaloperonospora parasitica* and *Thalassiosira pseudonana*
51 final candidates

Example of the candidates:

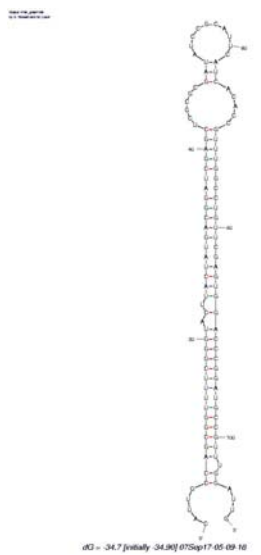
A candidate was identified on supercontig 1.605 (starting from nucleotide 9212)

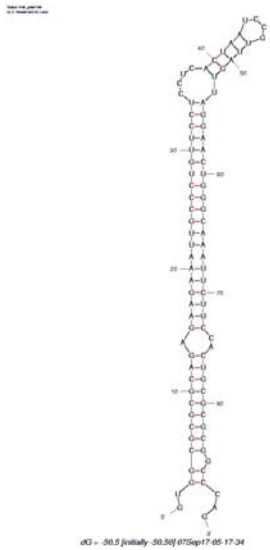
Hits in:

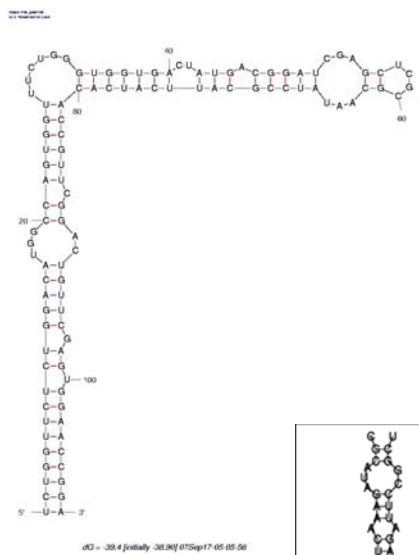
P. ramorum *P. sojae* *H. parasitica* *T. pseudonana*

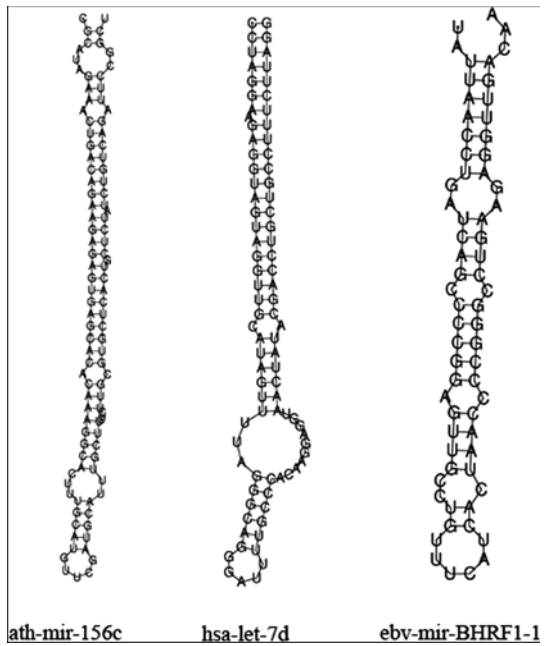


More examples:









Candidates from bioinformatic approach will be confirmed by:

- Comparing with data from sequencing approach;
- Northern blotting;
- RT-PCR.

miRNA data from sequencing approach and those from confirmed bioinformatic candidates will be used to optimize the bioinformatic approach:

- In the current approach, parameters were borrowed from those used in animals and plants. Usually the more stringent parameters were chosen.
- Should potential repetitive sequences be excluded?
- Should coding region be excluded? What about those overlapping with coding region but on the other strand?
- What's the best window size, srnaloop score, mfe, etc?

Future direction:

- Prediction of miRNA targets and function;
- Experimental characterization.

Additional interest: Identification and annotation of core components of RNA silencing pathway:

- Drosha, dicer-like proteins, agonaute proteins (agonaute subfamily and piwi-subfamily).
- We are especially interested in retrotransposon silencing in *P. infestans*.
- Initial survey showed that *P. sojae* and *P. ramorum* lacks piwi subfamily of agonaute proteins.
- Piwi proteins bind repeat-associate-siRNAs (rasiRNAs) in flies and are responsible for retrotransposon silencing. Evidence also showed that piwi proteins are involved in retrotransposon silencing in mammals.
- Retrotransposons seem to be very active in *P. infestans* and other oomycetes. Is that related to the lack of piwi proteins?
- Our sequencing of small RNAs will also provide sequence information of rasiRNAs in *P. infestans*