

Лидеры консервативны и полезны

Leaders are not useless

Aleksei Shevkoplias, Zoe Chervontseva



Background

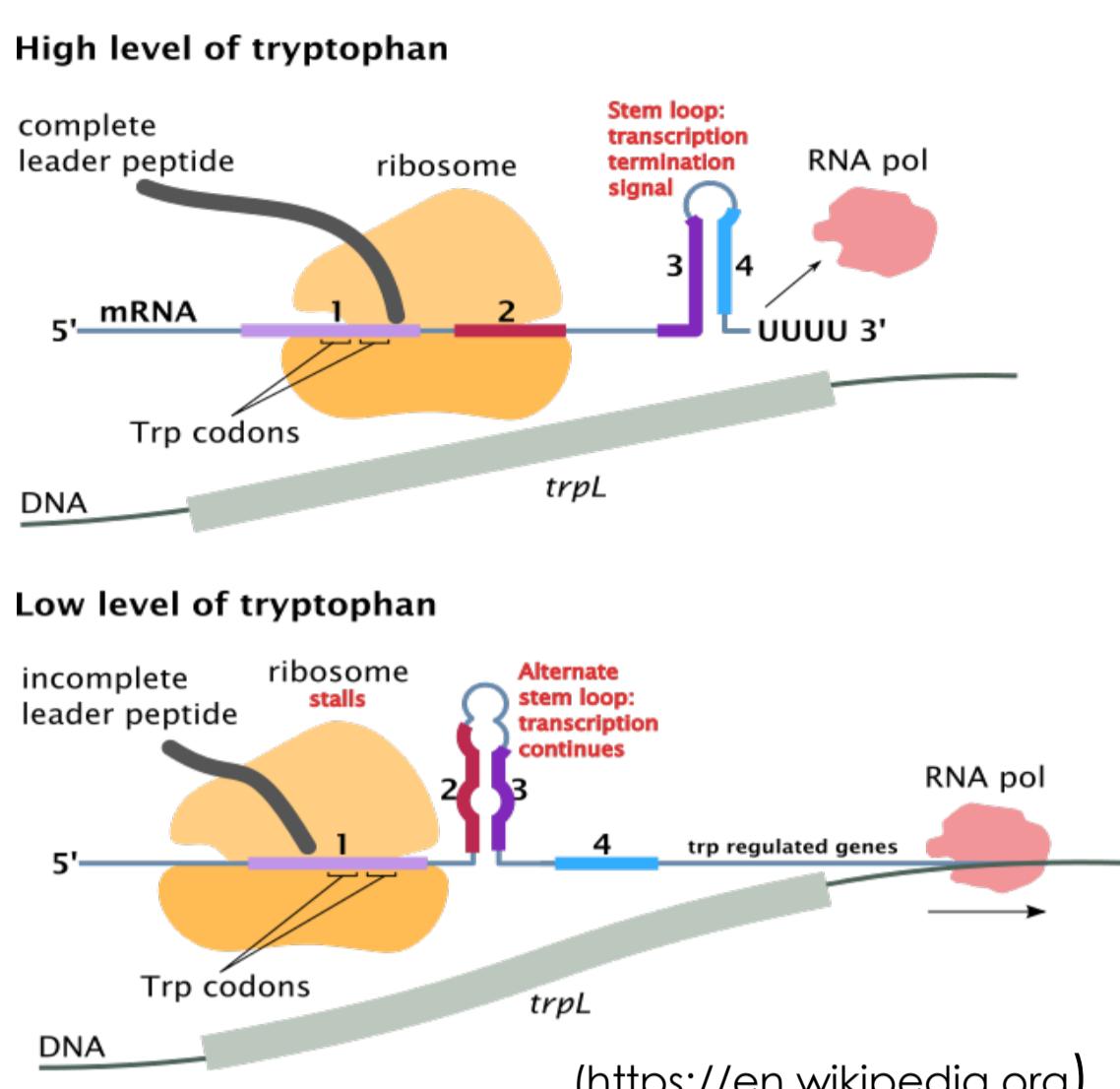


Figure 1. Transcriptional attenuation of the trp operon

Триптофановый оперон у некоторых бактерий регулируется транскрипционной аттенюацией, в ходе которой образуется так называемый триптофанный лидерный пептид. В учебниках написано, что он разлагается сразу же после трансляции, однако наши коллеги из Университета Юстуса Либига обнаружили, что у альфа-протеобактерии *Sinorhizobium meliloti* он взаимодействует с sRNA. Мы проанализировали лидерные пептиды других альфа-протеобактерий, чтобы понять, могут ли они иметь сходные дополнительные функции.

Results

The *trpL* peptides form clusters by similarity, which are associated with the systematic position

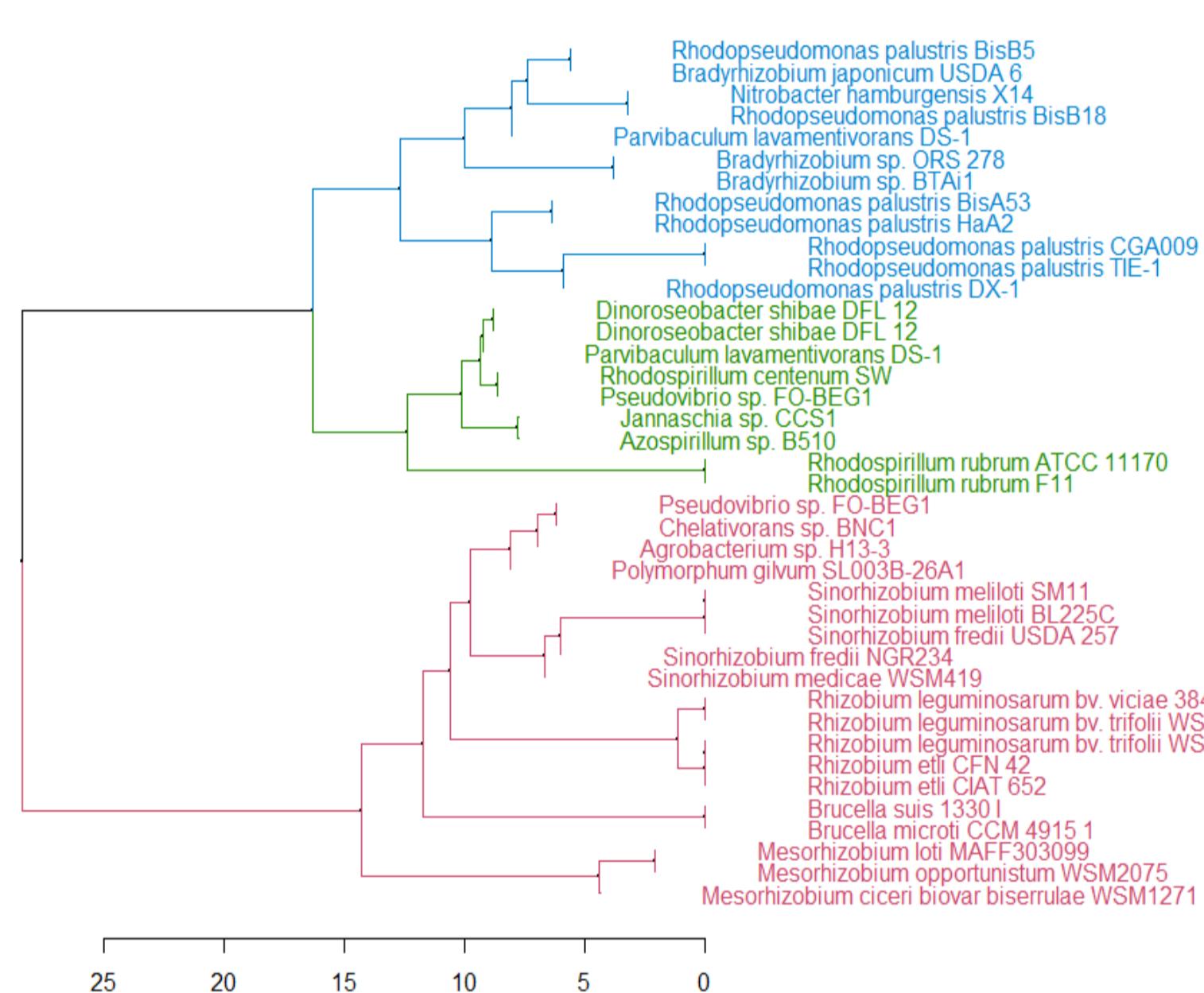


Figure 2. Dendrogram based on the *trpL* similarity. Only peptides with two tryptophans in a row are considered.

Besides tryptophan, positively charged amino acids (mainly arginine) are conserved within the clusters

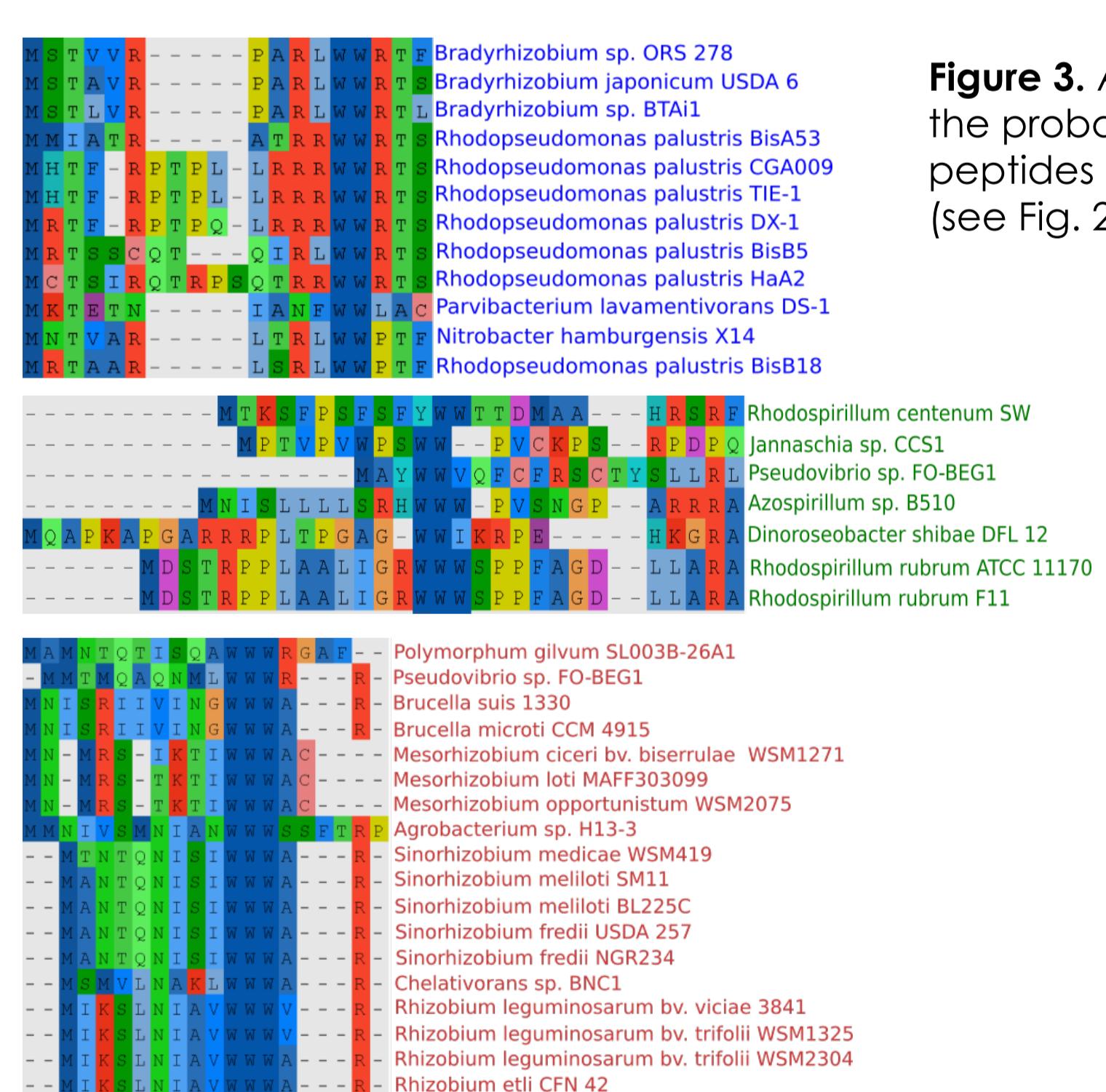


Figure 3. Alignment of the probable *trpL* peptides in the clusters (see Fig. 2)

Leader peptides do not necessarily have two tryptophans in a row



Figure 4. Sequence of a possible *trpL* peptide from several *Acetobacter pasteurianus* strains. Consider the «WLLW» subsequence

On the dendrogram for the peptides with at least one tryptophan, most of the peptides with two consecutive tryptophans (except those which are 3 aa long) form a cluster (yellow) with the above-mentioned *Acetobacter pasteurianus* peptides.

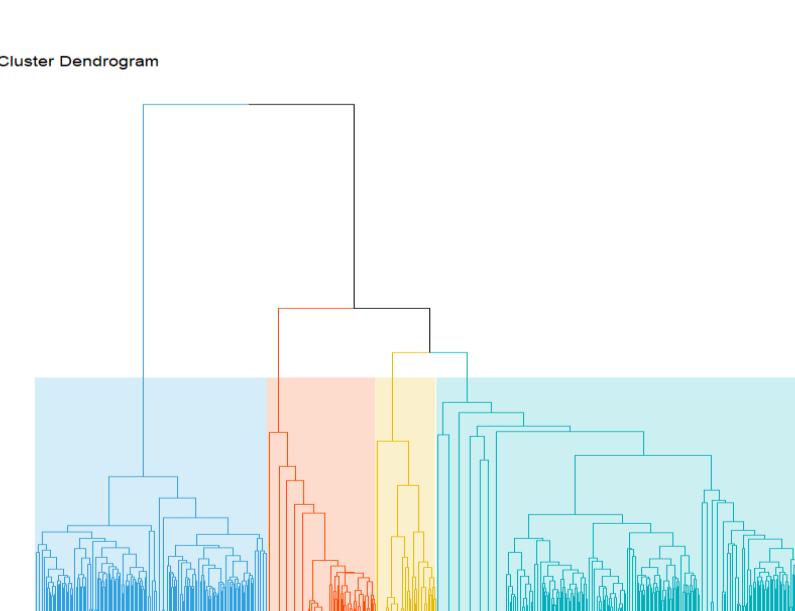


Figure 5. Dendrogram based on the similarity of candidate *trp* leaders with at least one tryptophan

Лаборатория сравнительной и функциональной геномики
Laboratory of Comparative and Functional Genomics

Materials and Methods

What we have done:

Что мы сделали:

- Genomes of the Alphaproteobacteria were selected from the database.
Из базы данных были отобраны гены альфа-протеобактерий.
- Genomes with annotated operons were selected.
В проекте использовались гены с аннотированными оперонами.
- Operons containing genes with tryptophan synthesis-related keywords were selected.
Были выбраны опероны, названия генов которых содержат слова, связанные с синтезом триптофана.
- 300 bp upstream sequences were collected and translated into possible peptides.
Были выписаны пептиды, вероятно закодированные в 300 нуклеотидах до начала оперонов.
- Peptides with at least one tryptophan and then with at least two tryptophans in a row were aligned with each other using the Smith-Waterman algorithm.
Сначала пептиды с хотя бы одним триптофаном, а затем только пептиды с двумя триптофанами подряд были выравнены друг с другом с помощью алгоритма Смита-Ватермана.
- Relying on the pairwise alignment scores, dendograms were built using the R packages cluster, factoextra, magrittr, ggplot2, colorspace, dendextend.
Опираясь на очки попарных выравниваний, мы построили дендрограммы, используя R.

- A script predicting candidate hairpins was written. Then the candidate attenuators were selected manually from all predicted hairpins. The website <http://www.e-rna.org/r-chie> was used to visualize them.
С помощью скрипта были предсказаны возможные шипы, затем из них отобрали возможные аттенюаторы, которые визуализированы на вышеуказанном сайте.

What we have not considered:

Что мы (пока) не учли:

- We have not used the most recent genome database.
Используемой базе данных несколько лет. Недавно секвенированные гены не были проанализированы.
- It's likely not all operons are annotated.
Наверняка не все опероны аннотированы.
- Operons of the tryptophan metabolism also contain some of these keywords.
Некоторые опероны, связанные с метаболизмом триптофана, тоже попали в выборку.
- If several starts were possible in one frame, only the distal one was considered.
Если в рамке считывания были возможны несколько старт-кодонов, то учитывался только первый из них.
- Scores and penalties were not rationally designed for short peptides.
Очки и штрафы не были рационально определены для коротких пептидов.

Our method of identifying possible attenuators is not well-defined and therefore probably is not exact.
Наш метод выбора аттенюаторов неточен.

A number of scripts were written ad hoc in Python. Some figures were processed in Inkscape. The poster is done in OpenOffice Impress

36 *trp* attenuators with possibly functional leader peptides were predicted (with different degrees of probability)

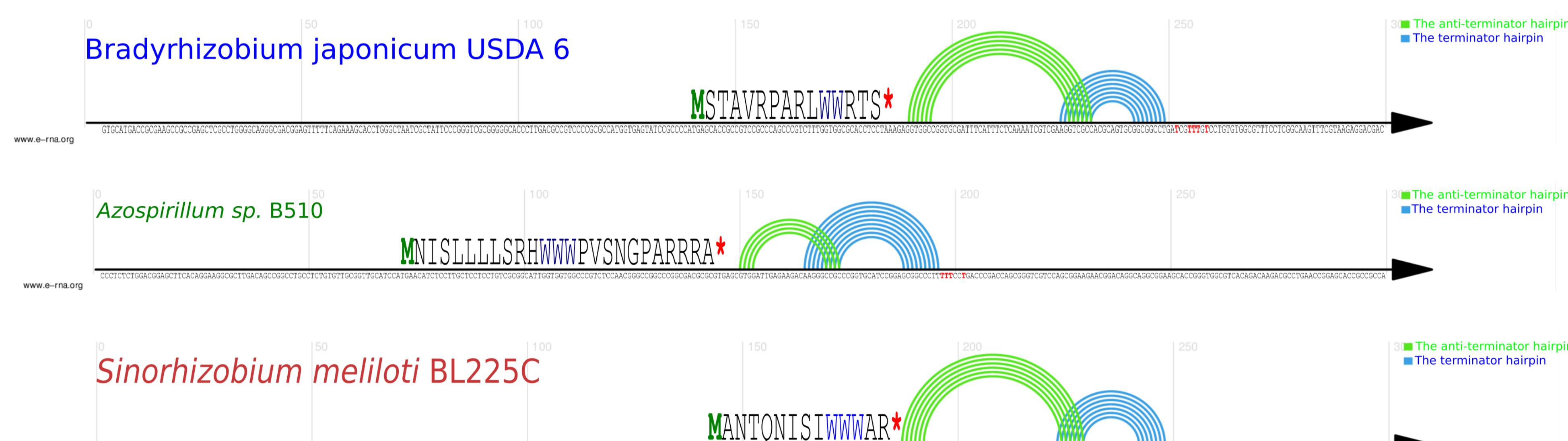


Figure 6. Typical for the clusters (see Fig. 2) regions of 300 bp upstream of the tryptophan operon. Possible attenuation hairpins and leader peptides are shown.

Conclusions

In some species of the Rhizobiales order (Alphaproteobacteria), *trpL* peptides tend to have positively charged amino acids in particular positions. We know that even one arginine residue can make a peptide bind to small RNA (Mikolaj Olejniczak, personal communication). So our prediction is that the *trpL* peptide has its own function in at least some Rhizobiales species. That inspires for further research.

Полученные результаты позволяют нам предположить, что по крайней мере в некоторых бактериях лидерный пептид триптофанового оперона имеет самостоятельные функции. Это вдохновляет на дальнейшие исследования.

References

- Abstracts of the 5th MEETING of Regulating with RNA in Bacteria & Archaea
- A. Vitreschak et al. Attenuation regulation of amino acid biosynthetic operons in proteobacteria: comparative genomics analysis, FEMS Microbiology Letters 234(2004) 357-370

