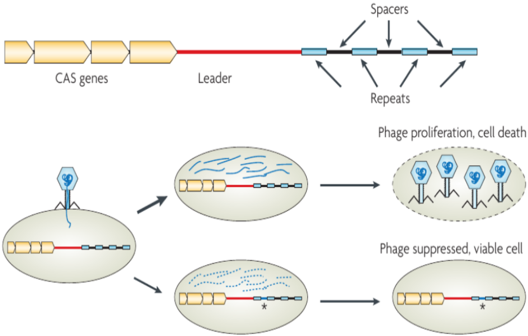
Analysis of CRISPR-Repeat Structure

Prof.Dr. Rolf Backofen
Dominic Rose
Sita Lange
Omer Alkhnbashi

Bioinformatics Group
University of Freiburg
Germany

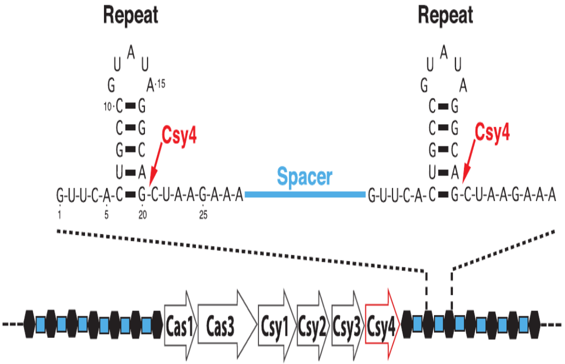
What is a CRISPR ?

- Clustered Regularly Interspaced Short Palindromic Repeats.
- CRISPR - a mechanism to keep the record of and destroy invasive elements.
- CRISPR - a novel class of repeats, separated by unique spacer sequences.

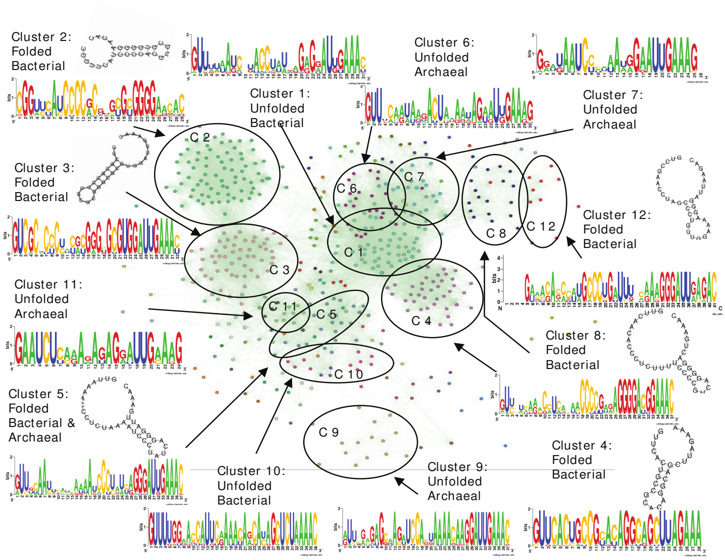


What are we aiming at ?

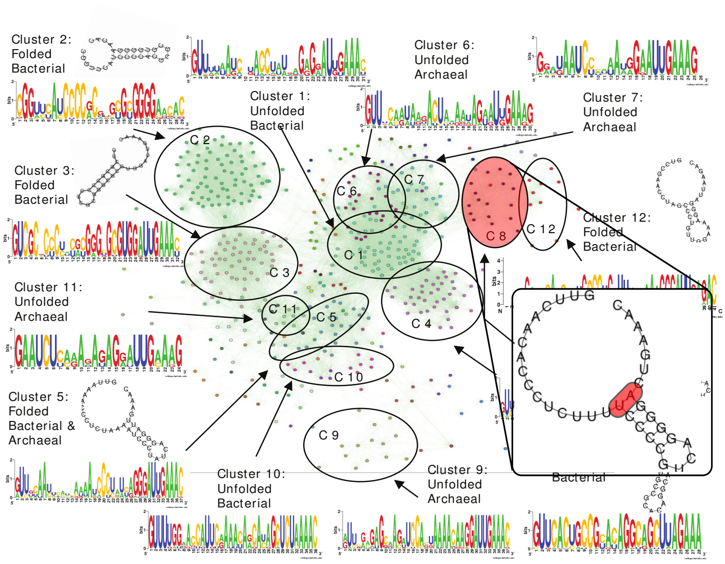
- Repeat structure.
- How to efficiently identify such a structure ?



Cluster Analysis of CRISPR Repeats (Kunin et al. 2007)



Cluster Analysis of CRISPR Repeats (Kunin et al. 2007)



Cluster 2:
Folded
Bacterial

Cluster 6:
Unfolded
Archaeal

Cluster 1:
Unfolded
Bacterial

Cluster 7:
Unfolded
Archaeal

Cluster 3:
Folded
Bacterial

Cluster 12:
Folded
Bacterial

Cluster 11:
Unfolded
Archaeal

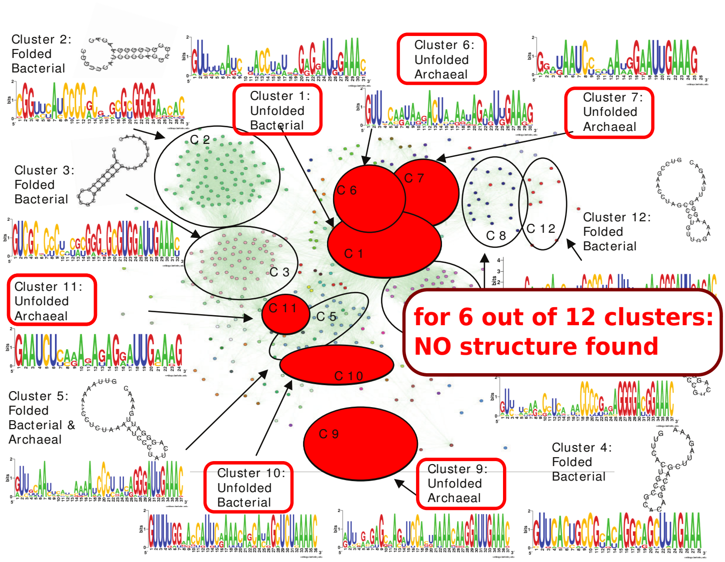
Cluster 5:
Folded
Bacterial &
Archaeal

Cluster 10:
Unfolded
Bacterial

Cluster 9:
Unfolded
Archaeal

Bacterial

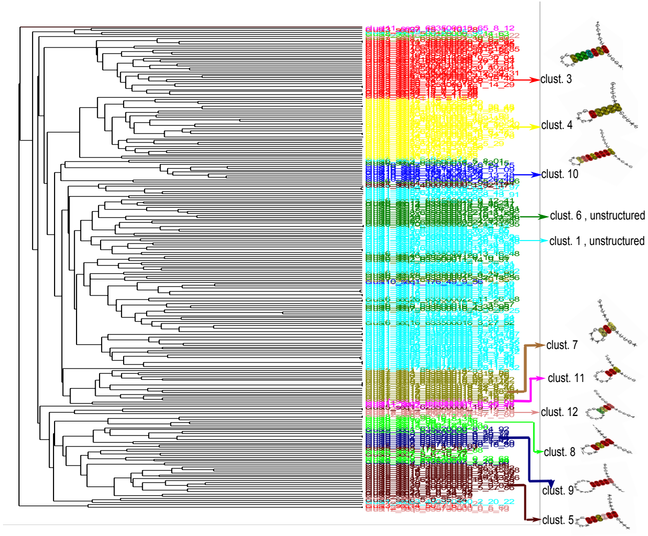
Cluster Analysis of CRISPR Repeats (Kunin et al. 2007)



for 6 out of 12 clusters:
NO structure found

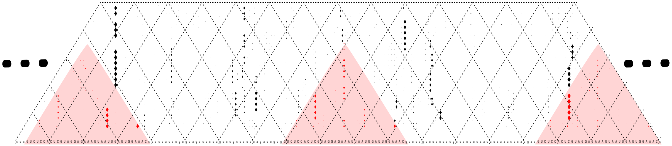
Our initial Analysis using LocARNA

- 11 of 12 clusters analyzed.
- Improved result: 9 of 11 structures found.



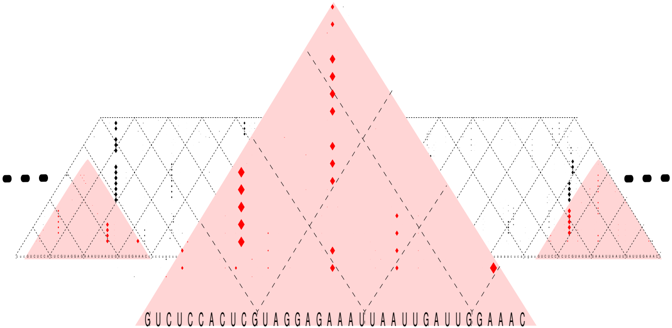
What is the Structure of a Repeat in Context ?

- Subsequence of 3 repeats.



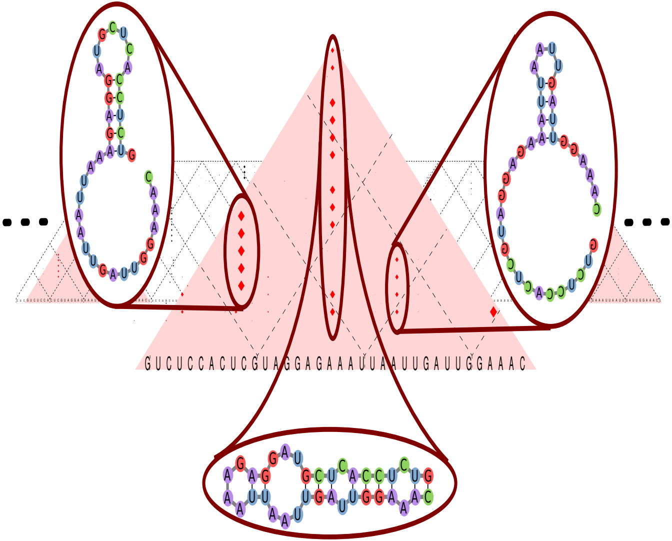
What is the Structure of a Repeat in Context ?

- Subsequence of 3 repeats.



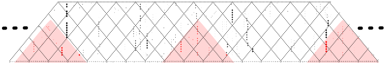
What is the Structure of a Repeat in Context ?

- Subsequence of 3 repeats.

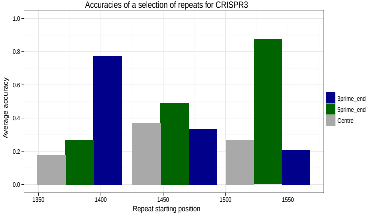


Which of the Structure is Correct ?

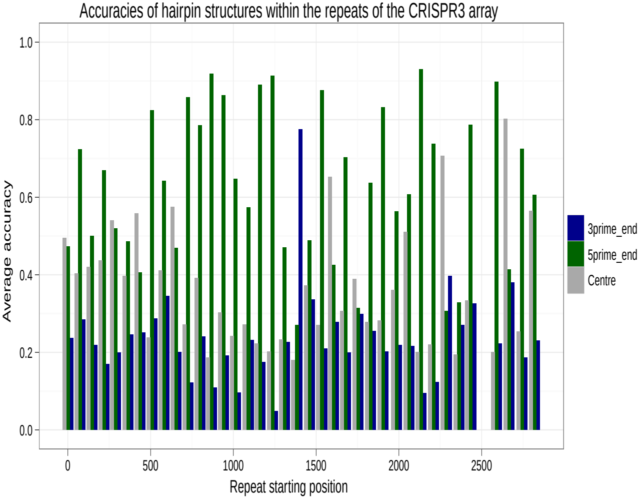
- Observation: different qualities ⇒ different processing order ?



- New quantitative measurement: expected accuracy.



All CRISPR Repeats



Acknowledgements

- Prof.Dr. Rolf Backofen.
- Dominic Rose.
- Sita Lange.

Thank you!