Identification of Conserved Regions in CRISPR protein family

Christine Baek[1] christib@andrew.cmu.edu
Qi Chu[1] qchu@andrew.cmu.edu
Yanyu Liang[1] yanyul@andrew.cmu.edu
[1]Computational Biology, Carnegie Mellon University

Abstract

The abstract goes here.

I. Introduction

TODO: Christine : brief intro

A. CRISPR/Cas

TODO: Christine writes about CRISPR bg Subsection text here.

B. Past Approaches

TODO: Qi : summarize the HMMer approach?

C. Approaches in This Paper

TODO: everyone?

D. Goal of Paper

TODO: Yanyu : short discussion of our goal in this paper

II. Methods

We used 3 differenr approaches .. blahblahblah. All used the same .fa sequence, etc. etc. talk about the data itself here, and why we used 3 different methods. **TODO:** Christine : fill in this section, and why protein sequence

A. Data Retrieval

TODO: Qi : talk about source of data

B. Sequence Alignment using Dynamic Programming

1) Model & Algorithm Overview: Semiglobal ailgnment using Needleman-Wunsch [3] and Local Alignment using Smith-Waterman Algorithm [4]
2) Pros and Cons: of using the method - what is it capable of, what are the limitations ?
3) Protocol: implementation details - justifications for decisions you made when you ran the experiment, parameters, etc.
4) Analysis: Discuss METHOD for analysis, not the actual result/analysis itself.

C. Gibbs Sampling

1) Model & Algorithm Overview: talk about overview of what the method does (method itself, not in detail of how you used it)
2) Pros and Cons: of using the method - what is it capable of, what are the limitations ?
3) Protocol: implementation details - justifications for decisions you made when you ran the experiment, parameters, etc.
4) Analysis: Discuss METHOD for analysis, not the actual result/analysis itself.

D. HMM

1) Model & Algorithm Overview: talk about overview of what the method does (method itself, not in detail of how you used it)

2) Pros and Cons: of using the method - what is it capable of, what are the limitations ?
3) Protocol: implementation details - justifications for decisions you made when you ran the experiment, parameters, etc.
4) Analysis: Discuss METHOD for analysis, not the actual result/analysis itself.

III. Results

Individual results for each
TODO: put in appropriate figures, and any analysis for each

A. Sequence Alignment using Dynamic Programming

B. Gibbs Sampling

C. HMM

IV. Conclusion

TODO: we need to do this one together
TODO: please include any other resources or papers you referenced

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.
[2] K. S. Makarova, et al., *An updated evolutionary classification of CRISPR-Cas systems* http://dx.doi.org/10.1038/nrmicro3569, 28 September 2015
[3] Saul B. Needleman, Christian D. Wunsch, *A general method applicable to the search for similarities in the amino acid sequence of two proteins* http://www.sciencedirect.com/science/article/pii/0022283670900574, 28 March 1970
[4] Smith, Temple F. & Waterman, Michael S., *Identification of Common Molecular Subsequences* Journal of Molecular Biology. 147: 195-197. doi:10.1016/0022-2836(81)90087-5. PMID 7265238., 1981