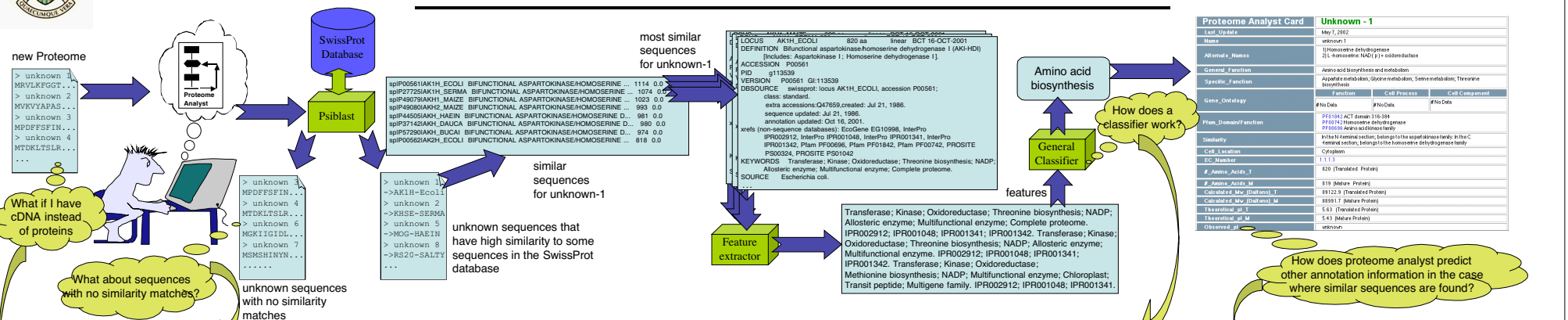




An Overview of Proteome Analyst

Roman Eisner, Brett Poulin, Duane Szafron, Paul Lu, Russell Greiner, David Wishart and Bahram Habibi-Nazhad
University of Alberta
www.cs.ualberta.ca/~bioinfo/PA

Proteome Analyst has received Significant funding from PENCE The Canadian Protein Engineering Network



Low Similarity Sequences! What do we do?

Analysis techniques are being developed based on motifs

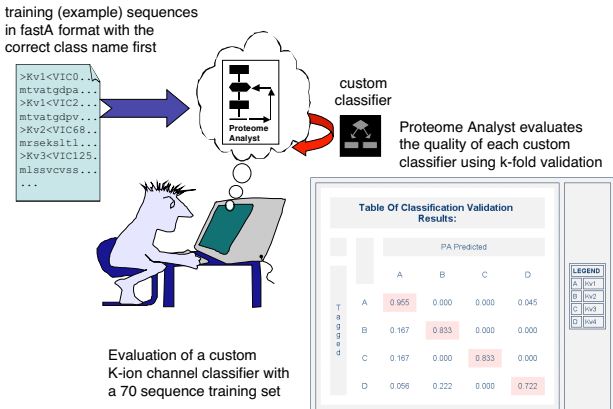


cDNA?

No problem. Proteome Analyst will pre-translate to protein.

```
> 2,14,73
3' End, Reading Frame = 3
CCCTGACAGTGGTGAACAGTACGTGATCCAAAGATTGGCCAGAGGGGAGGCGAGGAGGGTTCCTCC
GTCCACAGTGGTGGAGCGAGCGGGGGGCTGGGGCGGGGCGGCTCCACACACAGCCAGCGGCGCGCA
CGAGGGGGGCTCCCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA
ACCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG
H Q D C E A A V N R D Q I M L E Y A V Y V L S
T P T T A C T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T
M S Y P F D R D V A L K N F A E R F L B O S R
A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G
E E R E B A E R L M F L D R G L T A M E C A L H E R
A C C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C
T O N A V T C T C T G A T G A C A G A G A G A G A G A G A G A G A G A G A G A G A G A
I E R T V L I R S O V A I R S O V A I R S O V A I R S O V A I R S O V A I R S O V A I
A G A T G G G C C A C A A T T G A T T G G A G A G A T T C T T T C A C A G A C A C C C T
E W G C Q V L A A A R V F S
A T T A A A A T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T
C T T T A C T T G T C A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G
A A C T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T
```

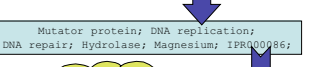
Creating a Custom Classifier



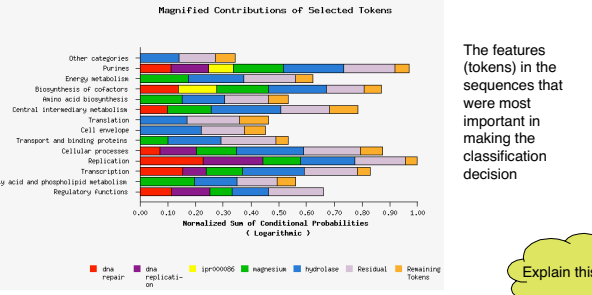
Classifiers

- Classifier Algorithms (Machine Learning)
 - Naive Bayes
 - Tree-Augmented Naive Bayes (TAN)
 - Neural Networks
 - Decision Trees
 - Support Vector Machine (SVM)
- Feature Selection - Wrapping

Classifier Decisions



Classifier Decisions Explained



Acknowledgements

Thanks to the 2001 summer students who worked on Proteome Analyst:

- Cynthia Luk
- Samer Nassar
- Kevin McKee
- Shannon Brooks

Thanks to some of the micro-biologists who tested Proteome Analyst and provided significant feedback:

- Warren Gallin
- Kathy Magor
- Chris Upton

Predicting other annotation information when similar sequences are found

