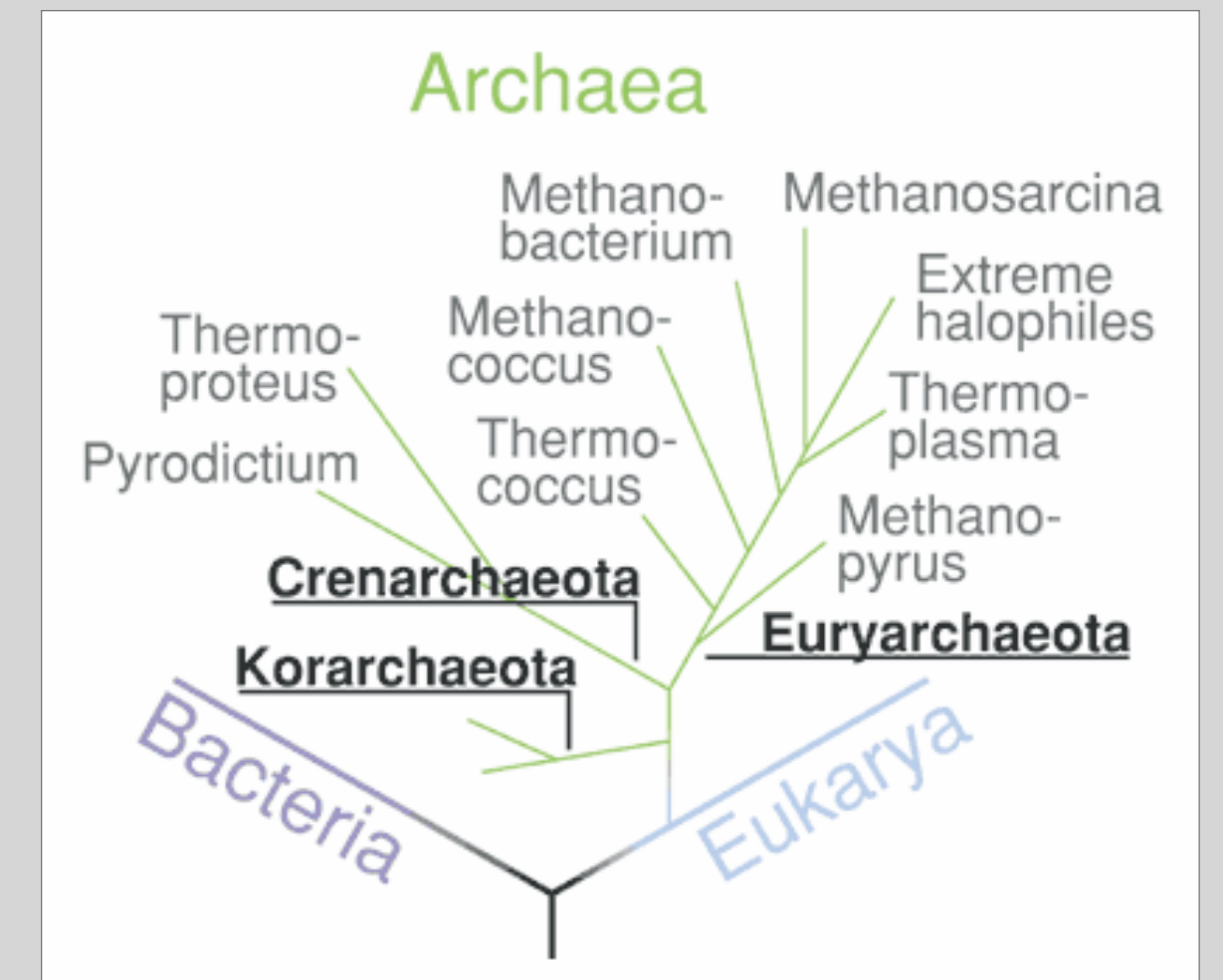




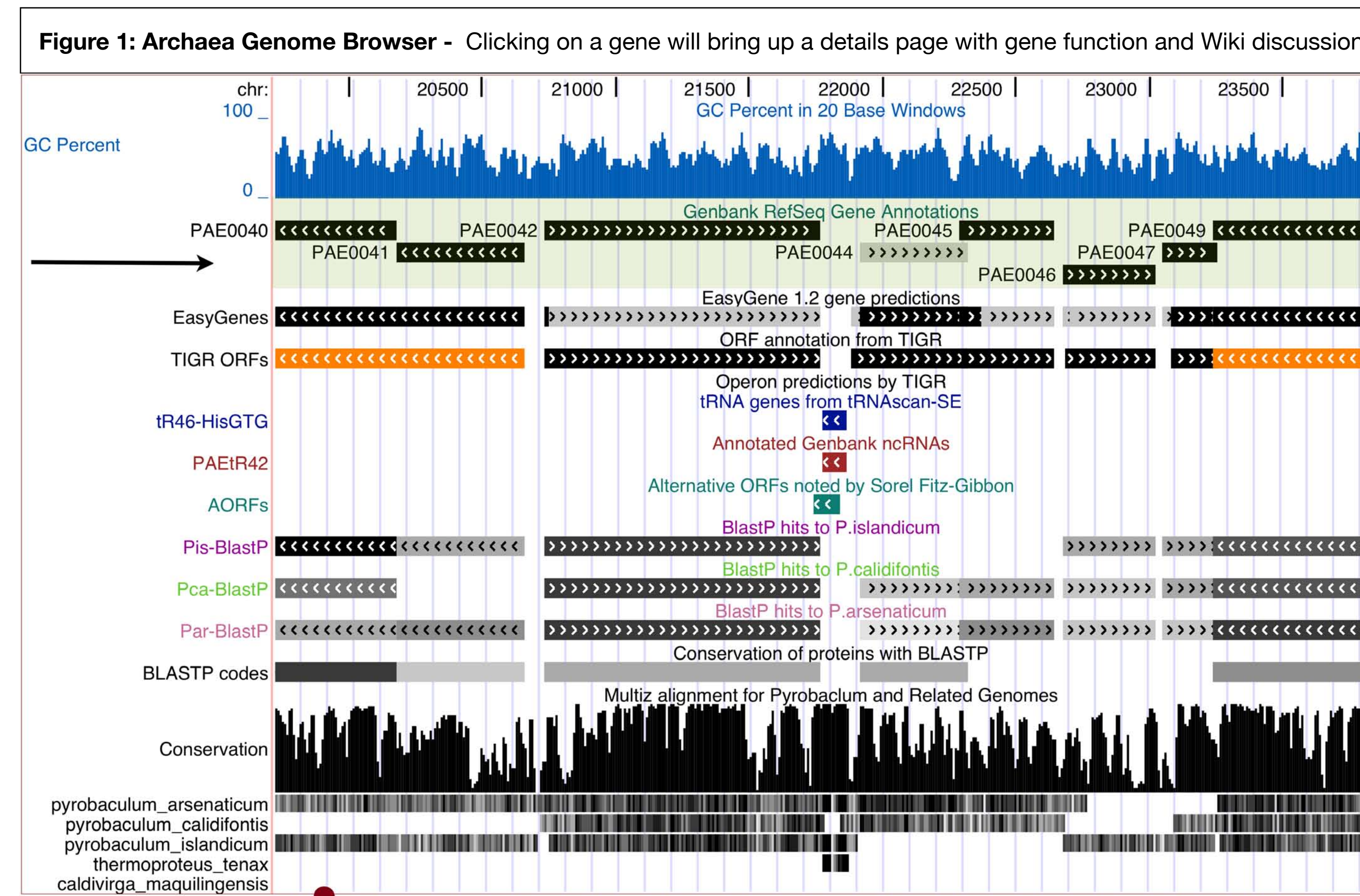
Enhancing and Developing Approaches to Gene Annotation

SURF-IT Elisabeth Kain, Mentor: Todd Lowe
Department of Biomolecular Engineering,
University of California, Santa Cruz



Introduction:

- After sequencing an entire genome, pertinent information about raw DNA sequences must be added to genome databases to describe different regions and identify gene function within genetic code.
- This process can be automated through developing better algorithms to process data thus reducing the need for manual tasks, and allowing more time for investigation of results.
- Our research focuses specifically on the study of Archaeal genomes through both computational and biochemical techniques.



Tools - Finding tRNA Introns:

- RNA plays an important role in protein synthesis and other chemical activities in the cell.
- tRNA? Transfer RNA is a class of RNA that carries amino acids into ribosomes and bonds with messenger RNA (mRNA) for protein production.
- tRNA can contain introns or possible gene insertions within a sequence.
- Searching for these introns within multi-aligned genomes can help detect where specific evolutionary events may have occurred, and how different species evolved.

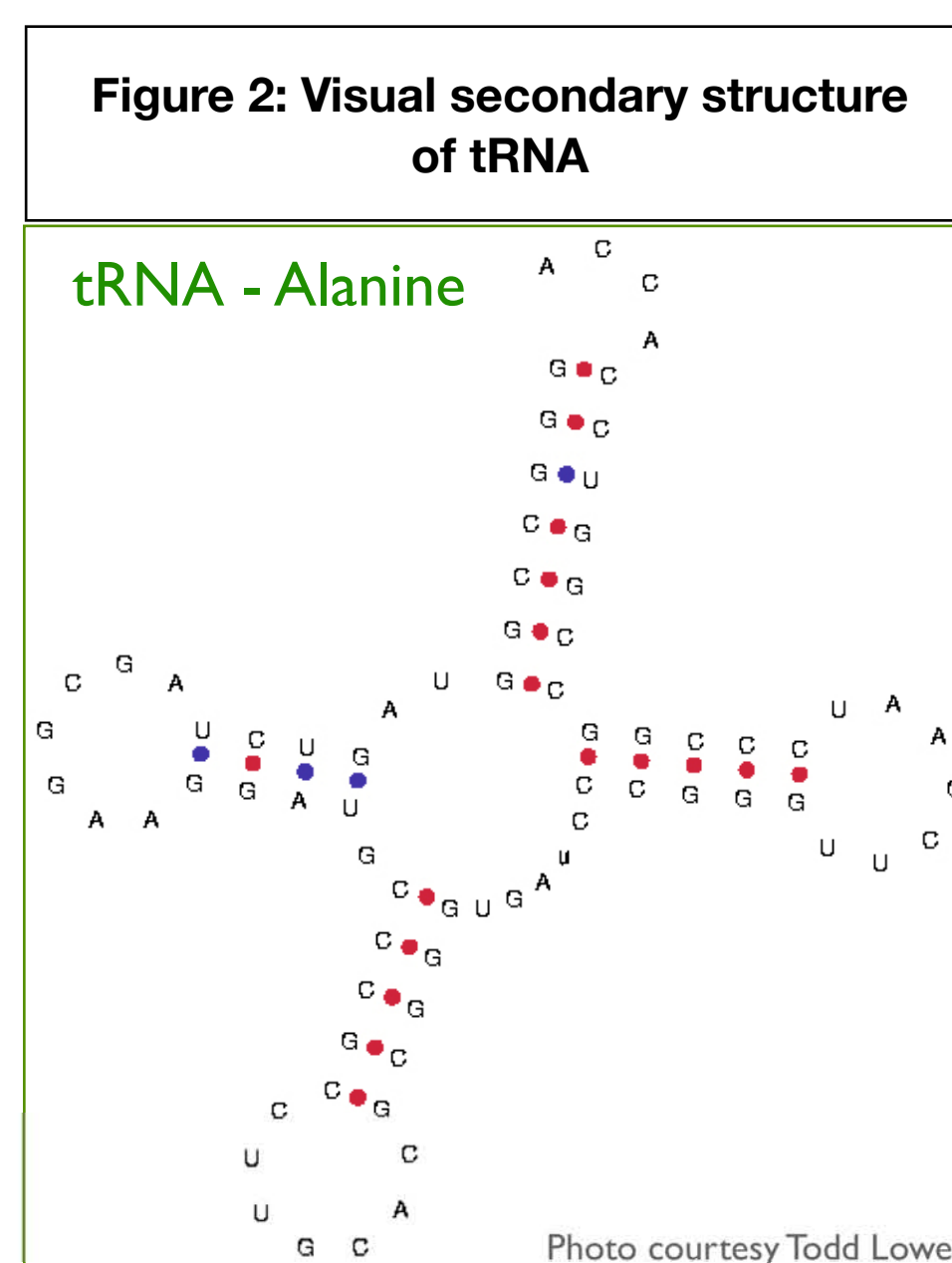


Figure 3: Sample tRNA Intron Finder Output - 5 Introns were found in this example tRNA as highlighted below.

Intron Number	Species to Compare Against	<i>P.aerophilum</i>	<i>P.arsenicum</i>	<i>P.islandicum</i>
1		x	-	x
2		x	-	-
3		x	x	-
4		-	x	-
5		-	x	x

Genome: *Pyrobaculum aerophilum*

tRNA: tR46-HisGTG

Multi-aligned View of Sequences:

```

pyrAer1  GCCCCGGTAGCTCAGTCCGGCAGAG.....CGGCGG.....GCTGCAG.....ACCCGTAGGTCCCGGTTCAAAGGCGGGCTTTGCCGCCAATCGATCCCGCGGGGGCTCC
pyroArse0 GCCCCGGTAGCTCAGTCCGGCAGAGTGCAGAGTGCACCCGGAGAGCGGGCGG.....GCTGCAG.....ACCCGTAGTCCCGGTTCAA.....ATCCCGCGGGGGCTTC
pyroIsla0 GCCCCGGTAGCTCAGTCCGGCAGAG.....CGGCGGAGGACCGCCAGTCCGCCACGCTGCAGGCTACGGCAGCCACCCGTAGTCCCGGTTCAA.....ATCCCGCGGGGGCTCC

```

Community-based methods:



● Integrating the UCSC Genome Browser Wiki allows scientists worldwide to collaboratively contribute their findings to a moderated discussion of genes in any species. Under development for both Archaeal and Human Genome Browsers.
<http://genomewiki.cse.ucsc.edu>

- Previously, if a user clicked on a gene to find out more information, most data would be listed as “hypothetical” or nonexistent.
- Our work helps researchers to communicate their ideas and have them displayed on these pages.

Conclusions & Future Work

The development of these web-based tools help to better determine how Archaea are able to live in such extreme environments, and improve current annotation methods.

- Allow priority sorting in wiki annotations and strive to set new annotation standards in genomic publications.
- Analysis of the tRNA intron scanner in comparison with complementary tools.
- Develop more utilities and custom tracks in the Archaeal Browser: <http://archaea.ucsc.edu>



This work was completed as part of UCSC's SURF-IT summer undergraduate research program, an NSF CISE REU Site. This material is based upon work supported by the National Science Foundation under Grant No. CCF-0552688.