Automating the Classification of

Authorship & Acknowledgement

Nic Weber
nmweber@illinois.edu
@nniiicc

Andrea Thomer
akthom2@illinois.edu
@_an_dre_a

Motivation

> Explore the automation of classifying acknowledgment & authorship

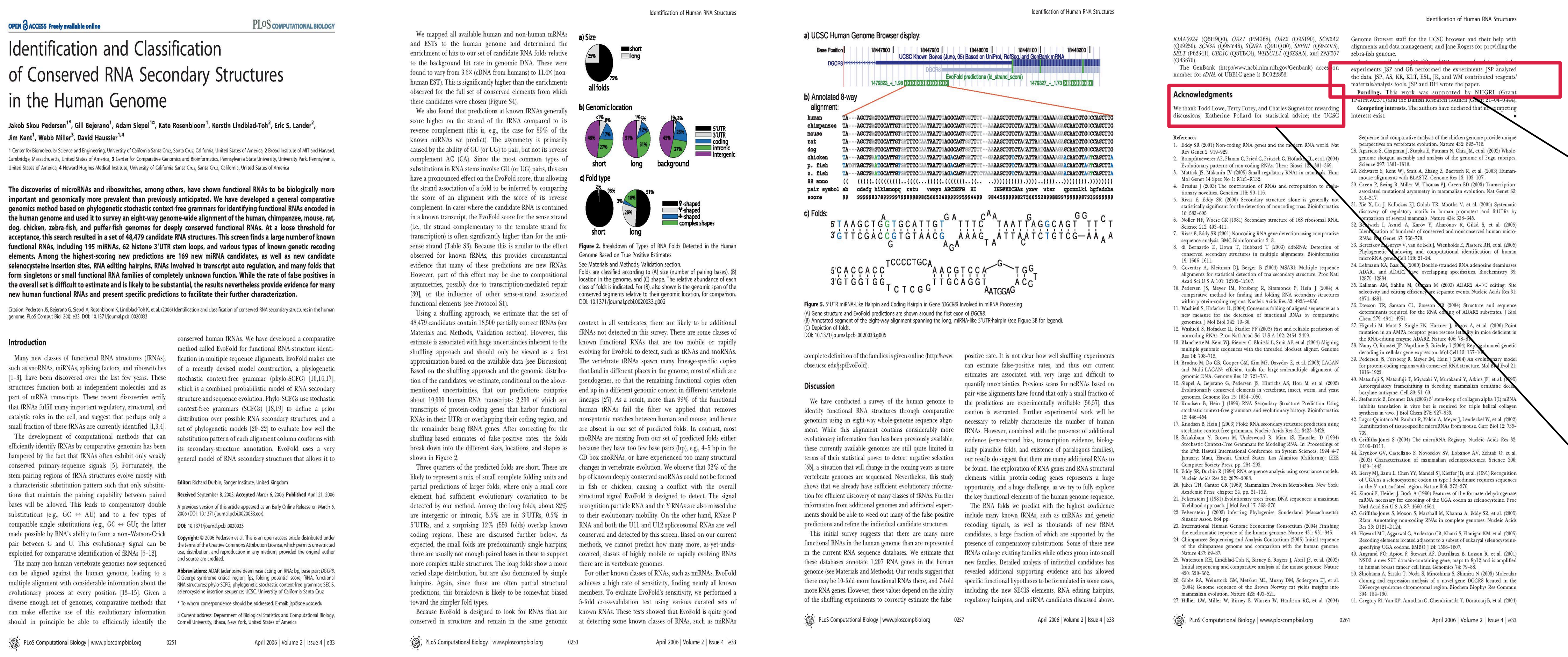
> Bootstrap the use of existing ontologies to increase the reliability of our own classifications

Data

> Corpus of articles from the field of Bioinformatics (n= 9741)

> Extracted authorship statements and acknowledgments (see below) for each article

> Manually classified a subset (n = 300) of each paratext using the Scholarly Contributions and Roles Ontology (Shotton and Peroni, 2013)



Authorship

Author contributions. JSP, GB, and DH conceived and designed the experiments. JSP and GB performed the experiments. JSP analyzed the data. JSP, AS, KR, KLT, ESL, JK, and WM contributed reagents/materials/analysis tools. JSP and DH wrote the paper.

Acknowledgement

We thank Todd Lowe, Terry Furey, and Charles Sugnet for rewarding discussions; Katherine Pollard for statistical advice; the UCSC Genome Browser staff for the UCSC browser and their help with alignments and data management; and Jane Rogers for providing the zebra-fish genome.

Automation

> Using our manual classifications as training data, we attempted to use Stanford's etcML to automate the classifications of each

Full results are available at

http://dx.doi.org/10.6084/m9.figshare.928642

Works Cited

Shotton, D. and Peroni, S. (2013). SCoRO, the Scholarly Contributions and Roles Ontology. Retrieved on Nov 25, 2013 from: http://www.essepuntato.it/lode/http://purl.org/spar/scoro