

Accepted Manuscript

Accurate Annotation of Protein-Coding Genes in Mitochondrial Genomes

Marwa Al Arab, Christian Höner zu Siederdisen, Kifah Tout, Abdullah H. Sahyoun, Peter F. Stadler, Matthias Bernt

PII: S1055-7903(16)30259-7

DOI: <http://dx.doi.org/10.1016/j.ympev.2016.09.024>

Reference: YMPEV 5637

To appear in: *Molecular Phylogenetics and Evolution*

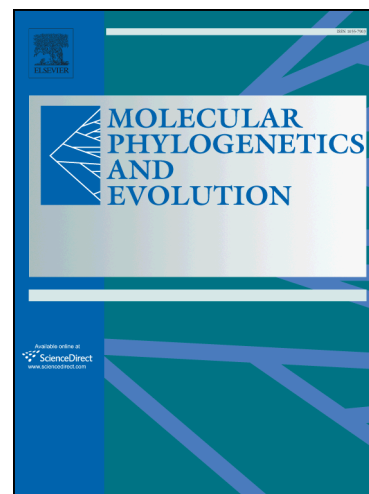
Received Date: 30 March 2016

Revised Date: 29 August 2016

Accepted Date: 25 September 2016

Please cite this article as: Arab, M.A., Höner zu Siederdisen, C., Tout, K., Sahyoun, A.H., Stadler, P.F., Bernt, M., Accurate Annotation of Protein-Coding Genes in Mitochondrial Genomes, *Molecular Phylogenetics and Evolution* (2016), doi: <http://dx.doi.org/10.1016/j.ympev.2016.09.024>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Accurate Annotation of Protein-Coding Genes in Mitochondrial Genomes

Marwa Al Arab^{a,b,h}, Christian Höner zu Siederdisen^{a,b,c}, Kifah Tout^h,
Abdullah H. Sahyoun^{a,b,h,i}, Peter F. Stadler^{a,b,c,d,e,f,g}, Matthias Bernt^{a,i,*}

^a*Bioinformatics Group, Department of Computer Science University of Leipzig,
Härtelstraße 16-18, D-04107 Leipzig, Germany.*

^b*Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstraße 16-18,
D-04107 Leipzig, Germany.*

^c*Institute for Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090
Wien, Austria.*

^d*Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, D-04103 Leipzig,
Germany.*

^e*Fraunhofer Institut für Zelltherapie und Immunologie, Perlickstraße 1, D-04103 Leipzig,
Germany.*

^f*Center for non-coding RNA in Technology and Health, University of Copenhagen,
Grønnegårdsvej 3, DK-1870 Frederiksberg C, Denmark.*

^g*Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501*

^h*Doctoral School of Science and Technology, AZM Center for Biotechnology Research,
Lebanese University, Tripoli, Lebanon*

ⁱ*TRON - Translational Oncology at the University Medical Center of the Johannes
Gutenberg University Mainz gGmbH, Mainz, Germany.*

^j*Parallel Computing and Complex Systems Group, Department of Computer Science,
University of Leipzig, Augustusplatz 10 D-04103 Leipzig, Germany*

Abstract

Mitochondrial genome sequences are available in large number and new sequences published nowadays with increasing pace. Fast, automatic, consistent, and high quality annotations are a prerequisite for downstream analyses. Therefore, we present an automated pipeline for fast *de-novo* annotation of mitochondrial protein-coding genes. The annotation is based on enhanced phylogeny-aware Hidden Markov models (HMMs). The pipeline builds taxon-specific enhanced multiple sequence alignments (MSA) of already annotated sequences and corresponding HMMs using an approximation of the phylogeny.

*Corresponding author

Email addresses: marwa@bioinf.uni-leipzig.de (Marwa Al Arab),
choener@bioinf.uni-leipzig.de (Christian Höner zu Siederdisen), ktout@ul.edu.lb
(Kifah Tout), abdullah.sahyoun@tron-mainz.de (Abdullah H. Sahyoun),
studla@bioinf.uni-leipzig.de (Peter F. Stadler), bernt@informatik.uni-leipzig.de
(Matthias Bernt)

Download English Version:

<https://daneshyari.com/en/article/5592552>

Download Persian Version:

<https://daneshyari.com/article/5592552>

[Daneshyari.com](https://daneshyari.com)