Jonathan Wells

@Jonathan N Wells
@jonathan-wells
@JN_Wells

email: jnw72@cornell.edu
mobile: +1 (607) 262-0386

Technical Skills

- Proficient in Python, Bash, and R, and have played around with several other languages, e.g. C, Julia, Rust.
- Comfortable working in high-performance computing environments. Willing and able to pick up cloud-based services.
- Up to date with best practice in software development, inc. version control (regular Git user), test-driven development (in Python), workflow management systems (basic familiarity with Snakemake), etc.
- 10+ years of experience with large-scale genome mining, genome annotation, sequence analysis and phylogenetics - I have a strong theoretical and practical understanding of these topics.
- Broad experience working with a wide range of data types, including but not limited to NGS sequencing (inc. single cell), high-throughput proteomics and structural data.
- Basic familiarity with modern ML approaches, particularly in the area of natural language processing. Eager to gain more hands-on experience in this field.
- 5+ years working directly with bench scientists – I enjoy collaborating in the design and interpretation of wet-lab experiments.
- Long-term mentoring experience from undergraduate to graduate level (9 individuals, mean duration 8 months)
- Able to communicate scientific findings rapidly and effectively through a range of media (see selected publications below).

Professional Experience

2023-	Research Associate in Professor Cédric Feschotte's group at Cornell University, with time divided between a continuation of my prior work (70%) and new work on HIV (30%).
2019-2023	Independent HFSP Postdoctoral Fellow in Professor Cédric Feschotte's group at Cornell University, studying genomic conflict between transposable elements and animals, with a focus on the evolution of genomic defense systems.
2017-2018	Postdoctoral Associate in Dr Joseph Marsh's group at the University of Edinburgh. This was a continuation of my PhD work and was focused on understanding the regulatory mechanisms behind protein complex assembly, from pre-transcriptional to post-translational.

Education and Professional Development

2014-2017	PhD in Computational Biology, University of Edinburgh
2012-2013	MSc in Mechanistic Biology, 1st class with honors, University of Sheffield
2009-2012	BSc in Molecular Biology, Class 2.i, University of Sheffield

2023	Natural Language Processing Specialization, Coursera, Online
2022	Machine Learning Specialization, Coursera, Online
2019	Workshop on Molecular Evolution, MBL Woods Hole, MA
2019	Probability: The Science and Uncertainty of Data, MITx, Online

2018	Tree building: Advanced concepts and practice of phylogenetic analysis, Faro, Portugal (EMBO course)

Grants and Awards

2019-2023	Human Frontier Science Program – Long term fellowship: $300,000
2019	MBL Woods Hole – Molecular Evolution course stipend: $800
2015	Biochemical Society bursary to attend Translation UK 2015: ~$500
2012-2013	Biotechnology and Biological Sciences Research Council stipend: ~$24,000

Selected Publications*

***** Google Scholar for complete list.

1. **Wells, J. N.** et al. Transposable elements drive the evolution of metazoan zinc finger genes. *Genome Res* **33**, 1325-1339 (2023).

   Combined extensive data mining and annotation (4k+ animal genomes) with comparative phylogenetics and functional genetics in zebrafish embryos to demonstrate that zinc finger genes (the largest family of animal transcription factors) have evolved to silence transposable elements during embryogenesis and beyond.

2. **Wells, J. N.** & Feschotte, C. A Field Guide to Eukaryotic Transposable Elements. *Annu Rev Genet* **54**, 539-561 (2020).

   Comprehensive review of eukaryotic transposable elements. Ann. Rev. Genetics' most downloaded article, 2020-present.

3. **Wells, J. N.**, Gligoris, T. G., Nasmyth, K. A. & Marsh, J. A. Evolution of condensin and cohesin complexes driven by replacement of Kite by Hawk proteins. *Current Biology* **27**, R17-R18 (2017).

   Developed a novel pipeline combining HMM to HMM sequence searches with the Markov clustering algorithm to resolve evolutionary relationships between proteins with alignments in the "twilight zone" of sequence similarity (20-35% identity). Used this to discover the origin of a family of protein cofactors essential to the emergence of eukaryotic condensin and cohesion protein complexes.

4. McShane, E., Sin, C., Zauber, H., **Wells, J. N.** et al. Kinetic analysis of protein stability reveals age-dependent degradation. *Cell* **167**, 803-815 (2016).

   Part of an interdisciplinary group using mass-spectrometry to understand the degradation of mammalian proteins. We found that ~40% of proteins are initially degraded much faster than expected based on later rates. Using data on protein complexes mined from mass-spec and structural sources, I demonstrated that this phenomenon was explained by rapid degradation of unbound protein complex subunits, followed by subsequent stabilization in mature complexes.