

ScienceWatch Home
Interviews

Featured Interviews

Author Commentaries

Institutional Interviews

Journal Interviews

Podcasts

Analyses

Featured Analyses

What's Hot In...

Special Topics

Data & Rankings

Sci-Bytes

Fast Breaking Papers

New Hot Papers

Emerging Research Fronts

Fast Moving Fronts

Research Front Maps

Current Classics

Top Topics

Rising Stars

New Entrants

Country Profiles

About Science Watch

Methodology

Archives

Contact Us

RSS Feeds

scienceWATCH.com

TRACKING TRENDS & PERFORMANCE IN BASIC RESEARCH



Interviews

Analyses

Data & Rankings

2008 : July 2008 - Fast Moving Fronts : Ewan Birney

FAST MOVING FRONTS - 2008

July 2008



Ewan Birney talks with *ScienceWatch.com* and answers a few questions about this month's Fast Moving Front in the field of Computer Science.


Article: GeneWise and genomewise

Authors: Birney, E;Clamp, M;Durbin, R

Journal: GENOME RES, 14 (5): 988-995 MAY 2004

Addresses: European Bioinformat Inst, Wellcome Trust Genome Campus, Cambridge CB10 1SA, England.

European Bioinformat Inst, Cambridge CB10 1SA, England.

Wellcome Trust Sanger Inst, Cambridge CB10 1SA, England.

[+enlarge image](#)
SW: Why do you think your paper is highly cited?

GeneWise is a very robust method which has been around for about 10 years now. Only about five years ago did we publish it correctly, which is outlined in this paper. It handles many of the scenarios people want to deal with robustly—such as that of homology-based gene structure prediction in the presence of draft sequence data, i.e., where sequencing error occurs at an appreciable rate.

Having been around for a while, a lot of groups are comfortable with its method and so it also gets used a little out of place; I am embarrassed to say that one person even runs it "because they like the way it formats the output" and not for the algorithm. Bioinformatics tools always end up having this behavior of groups being comfortable in how they run and quirks of the program meaning that once successful they hang around for far longer than you expect. But genomewise is a fundamentally correct algorithm (if rather slow).

SW: Would you summarize the significance of your paper in layman's terms?

We present two algorithms in this paper: GeneWise, which predicts gene structure using similar protein sequences, and genomewise, which provides a gene structure final parse across cDNA- and EST-defined spliced structure. GeneWise is heavily used by the Ensembl annotation system and many other places worldwide; Genomewise has never really been used for a serious length of time.

Finding genes in genomes is a complex task, and one of our most informative pieces of information is related genes in other organisms. GeneWise takes a related gene (in fact its protein) and places it onto a new genome, accounting for splicing and sequencing errors. The algorithm was developed from a principled combination of hidden Markov models (HMMs). The genomewise algorithm is highly accurate and can provide both accurate and complete gene structures when used with the correct evidence.

SW: How did you become involved in this research and were there any particular problems encountered along the way?

I was originally trained (in my undergraduate days) as a Biochemist, but moved quickly into bioinformatics. I published my first set of programs (Pairwise and Searchwise) in 1994, while I was an undergraduate at Oxford, and GeneWise is really this program written correctly.

I benefitted greatly from a year at Adrian Krainer's Lab at **Cold Spring Harbor Laboratory** (CSHL), time spent with Toby Gibson (at EMBL) and also at Iain Campbell's lab at Oxford. I did my Ph.D. with Richard Durbin at the Sanger Institute, and have collaborated with him since that time.

In 2000, I joined the European Bioinformatics Institute (EBI) as a Team Leader, and I am now a Senior Scientist in the European Molecular Biology Laboratory (EMBL is the parent organization of EBI) and as part of the EBI Senior Management. I am best known as the head of the EBI side of the Ensembl project and my group has recently merged with **Rolf Apweiler's** group spanning DNA and protein sequence data. My own research continues to be focused on algorithms in bioinformatics.

SW: Where do you see your research leading in the future?

GeneWise is being replaced progressively with more modern schemes, in particular **Exonerate** from Guy Slater in my group. Exonerate basically can do most things that genewise does but 1,000-fold faster. However, there is a lot of resistance, for understandable reasons, for moving away from code systems which work in robust pipelines, which is why genewise keeps on being used.

Ewan Birney, Ph.D.
Head of Nucleotide Data
European Bioinformatics Institute (EBI)
Wellcome Trust Genome Campus
Hinxton, Cambridge, UK

Keywords: GeneWise, genewise algorithm, genomewise, gene structure, Ensembl annotation system, hidden Markov models, European Bioinformatics Institute (EBI), European Molecular Biology Laboratory, Ensembl project, Rolf Apweiler, DNA, protein sequence data, Exonerate, Guy Slater.



PDF

[back to top](#)

2008 : July 2008 - Fast Moving Fronts : Ewan Birney

[Scientific Home](#) | [About Scientific](#) | [Site Search](#) | [Site Map](#)

[Copyright Notices](#) | [Terms of Use](#) | [Privacy Statement](#)