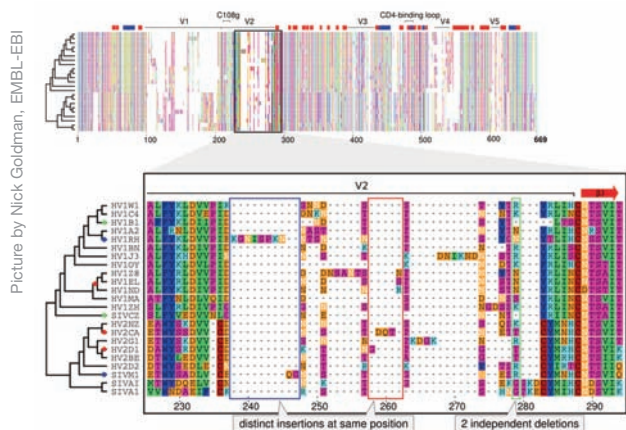


Embargoed until Thursday, 19 June 2008, 19:00 BST

Scientists fix bugs in our understanding of evolution

A new computational tool allows the most accurate insights into evolution ever



Sequence alignment according to the new, phylogeny-aware method.

Hinxton, 20 June 2008 – What makes a human different from a chimp? Researchers from the European Molecular Biology Laboratory’s European Bioinformatics Institute (EMBL-EBI) have come one important step closer to answering such evolutionary questions correctly. In the current issue of *Science* they uncover systematic errors in existing methods that compare genetic sequences of different species to learn about their evolutionary relationships. They present a new computational tool that avoids these errors and provides accurate insights into the evolution of DNA and protein sequences. The results challenge our understanding of how evolution happens and suggest that sequence turnover is much more common than assumed.

“Evolution is happening so slowly that we cannot study it by simply watching it. That’s why we learn about the relationships between species and the course and mechanism of evolution by comparing genetic sequences,” says Nick Goldman, group leader at EMBL-EBI.

The four letter code that constitutes the DNA of all living things changes over time; for example individual or several letters can be copied incorrectly (substitution), lost (deletion) or gained (insertion). Such changes can lead to functional and structural changes in genes and proteins and ultimately

to the formation of new species. Reconstructing the history of these mutation events reveals the course of evolution.

A comparison of multiple sequences starts with their alignment. Characters in different sequences that share common ancestry are matched and gains and losses of characters are marked as gaps. Since this procedure is computationally heavy multiple alignments are often built progressively from several pairwise alignments. It is impossible, however, to judge if a length difference between two sequences is a deletion in one or an insertion in the other sequence. For correct alignment of multiple sequences, distinguishing between these two events is crucial. Existing methods, that fail to do that, lead to a flawed understanding of the course of evolution.

“Our new method gets around these errors by taking into account what we already know about evolutionary relationships,” says Ari Löytynoja, who developed the tool in Goldman’s lab. “Say we are comparing the DNA of human and chimp and can’t tell if a deletion or an insertion happened. To solve this our tool automatically invokes information about the corresponding sequences in closely related species, such as gorilla or macaque. If they show the same gap as the chimp, this suggests an insertion in humans.”

Findings achieved with the new technique suggest that insertions are much more common than assumed, while the frequency of deletions has been overestimated by existing methods. A likely reason for these systematic errors of other techniques is that they were originally developed for structural matching of protein sequences. The focus of molecular biology is shifting, however, and understanding functional changes in genomes requires specifically designed methods that consider sequences’ histories. Such approaches will likely reveal further bugs in our understanding of evolution in future and might challenge the conventional picture of sequence evolution.

Source Article

A. Löytynoja & N. Goldman. Phylogeny-aware Sequence Alignment Prevents Systematic Error and Bias in Evolutionary Analyses, *Science*, 20 June 2008

Contact:

Louisa Wright PhD, EMBL-EBI Scientific Outreach Officer, Hinxton, UK, Tel: +44 1223 494665, www.ebi.ac.uk, louisa@ebi.ac.uk
 Anna-Lynn Wegener, EMBL Press Officer, Heidelberg, Germany, Tel: +49 6221 387 452, www.embl.org, wegener@embl.de

About EBI:

The European Bioinformatics Institute (EBI) is part of the European Molecular Biology Laboratory (EMBL) and is located on the Wellcome Trust Genome Campus in Hinxton near Cambridge (UK). The EBI grew out of EMBL's pioneering work in providing public biological databases to the research community. It hosts some of the world's most important collections of biological data, including DNA sequences (EMBL-Bank), protein sequences (UniProt), animal genomes (Ensembl), three-dimensional structures (the Macromolecular Structure Database), data from microarray experiments (ArrayExpress), protein-protein interactions (IntAct) and pathway information (Reactome). The EBI hosts several research groups and its scientists continually develop new tools for the bio-computing community.

About EMBL:

The European Molecular Biology Laboratory is a basic research institute funded by public research monies from 20 member states (Austria, Belgium, Croatia, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom) and associate member state Australia. Research at EMBL is conducted by approximately 80 independent groups covering the spectrum of molecular biology. The Laboratory has five units: the main Laboratory in Heidelberg, and Outstations in Hinxton (the European Bioinformatics Institute), Grenoble, Hamburg, and Monterotondo near Rome. The cornerstones of EMBL's mission are: to perform basic research in molecular biology; to train scientists, students and visitors at all levels; to offer vital services to scientists in the member states; to develop new instruments and methods in the life sciences and to actively engage in technology transfer activities. EMBL's International PhD Programme has a student body of about 170. The Laboratory also sponsors an active Science and Society programme. Visitors from the press and public are welcome. www.embl.org

Policy regarding use

EMBL press releases may be freely reprinted and distributed via print and electronic media. Text, photographs & graphics are copyrighted by EMBL. They may be freely reprinted and distributed in conjunction with this news story, provided that proper attribution to authors, photographers and designers is made. High-resolution copies of the images can be downloaded from the EMBL web site: www.embl.org