

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Genome analysis without compare

ArticleInfo		
ArticleID	:	4947
ArticleDOI	:	10.1186/gb-spotlight-20040429-01
ArticleCitationID	:	spotlight-20040429-01
ArticleSequenceNumber	:	299
ArticleCategory	:	Research news
ArticleFirstPage	:	1
ArticleLastPage	:	3
ArticleHistory	:	RegistrationDate : 2004-4-29 OnlineDate : 2004-4-29
ArticleCopyright	:	BioMed Central Ltd2004
ArticleGrants	:	
ArticleContext	:	130594411

David Secko

Email: dmsecko@interchange.ubc.ca

A report in the April 29 [Nature](#) describes a novel method for detecting selection pressures on specific proteins using a single genome rather than comparing nucleotide sequences.

Measurement of selection pressure on proteins classically involves the [comparison of nucleotide sequences](#) from numerous individuals or species, requiring numerous homologous sequences. This approach can be limited by lack of suitable homologs, unfeasibility of sequencing additional genomes, or proteins lacking recognizable homologs.

"The article describes a new way to analyze DNA sequences to identify all of the genes in a genome that are under pressure both to evolve and not to evolve. The advantage here is that by just taking a single genome you're interested in and looking at that sequence, you can find a hidden footprint without comparing homologous genes," said [Joshua Plotkin](#), from the Bauer Center for Genomics Research and lead author of the study.

The method relies on a concept Plotkin's team termed "codon volatility." They hypothesize that proteins under strong selective pressures leading to amino acid substitutions have DNA sequences that allow small mutations to result in changes in the protein - for example, they are more volatile. In contrast, proteins not under pressure will be less volatile.

"The basic idea is really simple. By simply looking at the volatility, we are able to infer whether a gene has been under pressure to change or remain fixed," Plotkin told us.

Applying the method to the completed genomes of *Mycobacterium tuberculosis* and *Plasmodium falciparum* correlated high volatility with predicted surface antigens in both organisms. In addition, the team observed that the *M. tuberculosis* PE/PPE family - [previously shown](#) to have an elevated substitution rate - exhibited high volatility in their analysis. They concluded that codon volatility could detect selection across a single genome.

"Whenever I speak to people in molecular evolution and tell them of the method to look at a single genome sequence to figure out the rate of evolution, that is met with a lot of skepticism and surprise, because it doesn't seem reasonable without comparisons," said Plotkin.

[Jeff Thorne](#), associate professor at North Carolina State University, agreed that is not how researchers would normally think about detecting selection pressure. "It's novel. The idea that you can use one sequence to make some evolutionary inferences is a little unusual, because usually you think about having to do comparative studies," he said.

"But it's quite an interesting paper, because one limitation of comparative genomics is that to compare you need at least two of something, and there are a lot of situations where it's hard to get such a data set. This makes this new approach all the more interesting," said Thorne, who was not involved in the study.

Thorne cautioned that codon bias and usage may be affecting the codon volatility measurements and stressed the importance of determining these relationships.

Rasmus Nielsen, from Cornell University, said the relationship between volatility and amino acid substitution requires further study. "The interesting question is then to find out why the pattern of codon usage measured using volatility correlates with the rate of amino acid substitution. In other organisms, such as *Drosophila*, there has long been an ongoing debate regarding the causes of non-random codon usage. It is possible that some of models of explanation proposed in this literature might also explain the results by Plotkin and his collaborators," Nielsen, who was not involved in the study, told us.

"It's useful to know which genes are evolving quickly and which ones are evolving slowly, in part because the genes under pressure not to evolve in any organism are usually critical for that organism to survive. In addition, proteins in a genome that are evolving quickly are potentially antigens and thus vaccine targets," said Plotkin.

References

1. *Nature*, [<http://www.nature.com/>]
2. Inference of selection from multiple species alignments
3. Joshua Plotkin, [<http://www.cgr.harvard.edu/people/people.html>]
4. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains
5. Jeffrey L. Thorne, [<http://statgen.ncsu.edu/thorne/thorne.html>]
6. Rasmus Nielsen, [<http://www.genomics.cornell.edu/faculty/facultybio.cfm?netid=rn28>]