

# New approach to 'spell checking' gene sequences

May 10 2012

---



(Phys.org) -- A PhD student from CSIRO and the University of Queensland has found a better way to 'spell check' gene sequences and help biologists better understand the natural world.

The student, Lauren Bragg, has contributed to the May issue of the prestigious journal *Nature Methods* highlighting her new approach and its software implementation called [Acacia](#).

Acacia analyses the output of next-generation gene sequencing instruments which read the four-letter alphabet of As, Cs, Ts and Gs – the 'bases' that code for DNA and spell out the genes of different living organisms. Acacia specifically applies to important parts of microbe

genes called amplicons.

Just as a computer spell checker finds typing errors in words, so Acacia finds errors in the DNA code of amplicon sequences produced during gene sequencing.

Acacia shows clear improvements over the two error-correction tools currently used by [biologists](#) for amplicon sequences and it's easier for biologists to use.

Ms Bragg's development of Acacia is part of the field of bioinformatics, a blend of computer science, statistics and biology. Despite her surname, however, she is modest about her achievements.

"It's exciting to be published in a journal like [Nature Methods](#) but I get more satisfaction from hearing how my software is helping biologists fix sequencing errors." she said.

Machine errors in the long lengths of A, C, G and T code can cause biologists to misinterpret which [genes](#) are there, or which microbial species might exist in an environmental sample from, say, a waste water treatment plant or from the ocean or even our guts.

Acacia works by using the statistical theory of likelihoods to analyse the code for DNA bases which may have been mistakenly added or deleted – common errors in [gene sequencing](#).

"The Nature article is our way of telling the international biology community that there's a new software tool they can use for error-correcting that's pretty easy to use, quick and reliable".

"That way, they won't think they've discovered a new microbe species when they haven't or overlooked one they should have found", she said.

The method, or algorithm, that Acacia uses took 18 months for Ms Bragg to fully develop and test.

Now it's nose to the grindstone to get the thesis done.

Provided by CSIRO

Citation: New approach to 'spell checking' gene sequences (2012, May 10) retrieved 26 April 2024 from <https://phys.org/news/2012-05-approach-gene-sequences.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.