

Genetic information: Codes and enigmas

Helen Pearson¹

1. Helen Pearson is a reporter for *Nature* based in New York.

Abstract

There's more than one way to read a stretch of DNA, finds Helen Pearson — and we need to understand them all.

"It was known that they were a little acquainted but not a syllable of real information could be procured as to what the truly was..." Reduce it to just a sequence of letters, and even a delicate phrase from Jane Austen's *Emma* becomes virtually impenetrable gobbledygook. So it was something of a triumph for Simon Shepherd when, in 2001, an algorithm he had written reconstructed all of *Emma*, word for separated word, from just such an uninterrupted string, despite being unacquainted with English vocabulary or syntax. The software worked out which groupings of letters were most likely to appear together, and thus have distinct meanings.

<http://www.nature.com/nature/journal/v444/n7117/full/444259a.html>