A longer paper gathers more citations

Researchers could garner more citations simply by making their papers longer, a study seems to imply.

In an analysis of 30,027 peer-reviewed papers published between 2000 and 2004 in top astronomy journals, astronomer Krzysztof Stanek of Ohio State University in Columbus found that the median number of citations increases with the length of the paper — from just 6 for papers of 2–3 pages to about 50 for 50-page papers[1].

"There is definitely too much obsession with citations and other indices."

There is, however, a limit to the benefits of size: citations start to tail off when papers reach lengths of 80 pages or so, perhaps because fewer people have the stamina to read them.

It is unexpected, says astronomer Jörg Dietrich of the European Southern Observatory headquarters in Germany, who recently conducted a similar analysis and found the same results but didn't publish them. "I expected that shorter papers would be cited more than longer ones," he says. "I assumed that people don't have the time to read long papers."

Papers of about 4 pages — the length of Letters in *Astrophysical Journal* and *Astronomy and Astrophysics*, which report brief summaries of work that is usually published in more detail later — fare better than papers 5–10 pages long. But brevity offers no such benefit for papers in the other two journals considered, *Astronomical Journal* and *Monthly Notices of the Royal Astronomical Society*, which do not have Letters.

Stanek says he can't explain the length effect. Dietrich thinks that longer papers are more useful because on average they contain more information. But he suspects that such papers are often not read in their entirety. A study of the propagation of citation errors has revealed that citations are often simply copied without being read[2].

Stanek does not intend to submit his study to a journal (although he encourages readers to "feel free to cite it as often as possible"). But it follows other recent investigations of citation statistics in astronomy[3,4], which looked at the effects of placing papers on the preprint server http://arxiv.org. On average this doubles the number of citations received[3], and Dietrich found that the timing of posting, and thus the paper's position in the mailing list, also had an important effect[4].

This apparent potential to rig citations

Gene chips unmask cryptic diseases

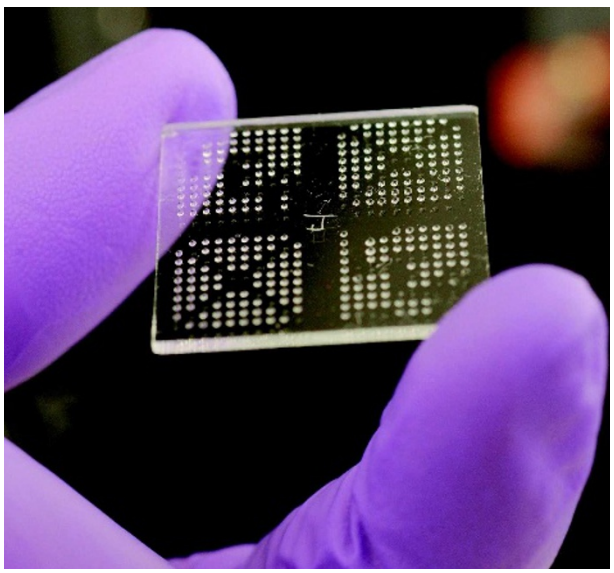
People with diverse symptoms including mental retardation, small head size, heart problems and cataracts have genomic rearrangements on the same region of chromosome 1, researchers reported last week.

A team led by genome scientist Evan Eichler at the University of Washington in Seattle linked deletions and duplications in a region of chromosome 1 that is 1.35 million DNA bases long to the abnormalities and cases of autism and learning difficulties[1]. Two studies published in July linked the same genomic region to schizophrenia[2,3].

The findings are just the latest in a spate of studies using microarrays, which measure genetic make-up and activity, to identify small DNA defects in patients with complex disorders. Doctors are already using the technique to diagnose genetic causes of unexplained developmental conditions after clinical tests have proven inconclusive, making microarray tests an early clinical success story in the genomic revolution.

"The technology is moving rapidly from research to clinical labs, and what's research information today is becoming part of our clinical practice tomorrow," says David Ledbetter, an expert in chromosomal abnormalities at Emory University in Atlanta, Georgia.

Microarrays are glass slides embedded with fragments of DNA selected from entire genomes. Scientists can detect abnormalities in a patient's DNA — such as losses or gains of hundreds or a thousand bases — by comparing it with the normal material on the microarray. These deletions or duplications, also called copy-number changes, can change the dosage of a gene. They were invisible to older techniques such as karyotype analysis and fluorescence *in situ* hybridization, both of which involve examining whole chromosomes.



PHOTOTAKE/BSIP/CHAGNON

Microarrays are an early genomic success story — the technology is already being used in the clinic.

Over the past two years, microarray-based DNA comparisons have become routine in clinics around the world. Microarray supplier Agilent, based in Santa Clara, California, estimates that the worldwide market in this form of microarray analysis approaches US$200 million and is growing rapidly.

But a new appreciation of the amount of genetic variation found in healthy people means it can be difficult to tell whether a particular change is causing a patient's symptoms.

In Eichler's study, for example, eight patients had spontaneous deletions in the chromosome 1 region, but nine patients inherited their abnormalities from a parent, some of whom showed no symptoms. It's not clear how rearrangements in the region, which contains nine genes, cause the many symptoms seen in the study.

Such findings complicate the use of microarray tests in prenatal diagnostics, an area that is proceeding cautiously. On 1 October, doctors will begin enrolling 4,000 patients in a study funded by the US National Institute of Child Health and Human Development in Bethesda, Maryland, to

OFFSHORE DRILLING: BLACK GOLD?
The United States debates its exploration of the sea for oil and gas.
www.nature.com/news

simply by adjusting a paper's size, place and timing of exposure sounds worrying at a time when citation statistics are increasingly being used, both formally and informally, as indicators of performance. Stanek points out that the '*h* index'[5], a measure of the cumulative impact of a researcher's published output, has become a fetish among some scientists. "You will become obsessed with it," he advises young scientists in his paper, and will start "telling other astronomers that your *h* is bigger than their *h*".

"There is definitely too much obsession with citations and other indices," Stanek says. This is partly because they are easy to calculate regardless of whether they actually mean very much. But he confesses to using them himself.

Stanek treats his results semi-humorously, using them with earlier studies to draw up a set of guidelines for how graduate students might manage their publications and citations to greatest advantage as their careers progress. "Make sure you submit your papers to [Arxiv] just after 4 p.m. US Eastern time on Wednesday," he suggests, for example.

"Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely."

But his study highlights some important questions. One is whether — in the face of new dissemination channels such as preprint servers and an increased sensitivity to citation indices — it is realistic to regard citations as an accurate measure of achievement.

Another question is how long a paper ought to be. If length really does matter, will that encourage researchers simply to inflate their results unnecessarily? Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely. "Most astronomical publications in the nineteenth century were very short observation reports, a few paragraphs at most," says Dietrich. But Stanek has no problem with length, saying that he prefers papers to "be as self-contained as possible — and I have seen a lot of short papers that are not".

Yet Dietrich thinks the risk of encouraging people to inflate their papers with waffle is small. "Writing a bloated paper is considerably more work, and incurs the risk of diluting an interesting work to the point that readers don't find the interesting aspects. Also, referees and editors are usually very critical of bloat. The best tactic is still to write a good paper, and a bloated paper is not a good paper." ■

Philip Ball

1. Stanek, K. Z. Preprint at http://arxiv.org/abs/0809.0692 (2008).
2. Simkin, M. V. & Roychowdhury, V. P. *Complex Syst.* **14,** 269 (2003).
3. Schwarz, G. J. & Kennicutt, R. C. Jr Preprint at http://arxiv.org/abs/astro-ph/0411275 (2004).
4. Dietrich, J. P. Preprint at http://arxiv.org/abs/0712.1037 (2007).
5. Hirsch, J. E. *Proc. Natl Acad. Sci. USA* **102,** 16569–16572 (2005).

compare microarray-based tests with traditional prenatal diagnosis techniques.

Scientists worldwide are trying to decrease uncertainty by pooling their samples in databases of copy-number aberrations, including Canada's Database of Genomic Variants, DECIPHER in the United Kingdom, and the European Union-funded ECARUCA. The databases help scientists link rare DNA changes to conditions ranging from autism and schizophrenia to kidney disease.

Eventually, scientists and doctors hope to understand why changes in gene copy number cause disease. But Eichler warns that this will require studies with perhaps tens of thousands of patients, as well as consultations with patients and their families for follow-up analyses, something that isn't part of most large genome-wide association studies today.

It will also require US geneticists to share samples as freely as their European colleagues, says Eichler, who is grateful to his European collaborators, but laments his US colleagues' reluctance to pool their resources on a large scale. Ledbetter is trying to coax US geneticists to share with the help of a grant from the American College of Medical Genetics Foundation in Rockville, Maryland. Eichler hopes the plan will succeed: "We are going to need a sea change." ■

Erika Check Hayden

1. Mefford, H. C. *et al. New Engl. J. Med.* Advance online publication doi:10.1056/NEJMoa0805384 (2008).
2. Stefansson, H. *et al. Nature* **455,** 232–236 (2008).
3. International Schizophrenia Consortium *Nature* **455,** 237–241 (2008).

SNAPSHOT
Long-lost antcestor

This pale-yellow, eyeless creature is so bizarre that naturalist E. O. Wilson named it "the ant from Mars". *Martialis heureka*, a native of the Brazilian Amazon, is the founding member of a new subfamily of ants, which split off from the ant family tree early in its evolution.

"It could represent a relict species that retained some ancestral morphological characteristics," says discoverer Christian Rabeling, a graduate student in integrative biology at the University of Texas in Austin.

Ants evolved from wasps, so it was long assumed that any living ancestral species would be wasp-like and similar to a Cretaceous ant fossil discovered in the 1960s by Wilson and his colleagues. But *Martialis* stunned entomologists by looking completely different — genetic analysis confirms that it doesn't fit into the known taxonomy of ants (C. Rabeling *et al. Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.0806187105; 2008). It has long, delicate mouthparts, for munching soft invertebrates perhaps. And, compared with its sturdy front legs, the rear two sets are thin and spindly (the three other legs in the specimen shown were lopped off for DNA analysis). "It doesn't even look like it could walk at all," says Brian Fisher, an ant expert and curator of entomology at the California Academy of Sciences in San Francisco.

Amber Dance

C. RABELING/UNIV. TEXAS, AUSTIN