COMMENT

VISUALIZATION Lavish lab guide lays out steps to making images that wow **p.430** **ART** Infrared analysis shows Albrecht Dürer worked less on pictures destined for sale **p.431** **PUBLISHING** Ethics procedures should be required in papers, like other methods **p.432**

FUNDING Patients' priorities for schizophrenia research get UK support **p.432**

Face up to

Face up to false positives

Scientists and journals must work together to ensure that eye-catching artefacts are not trumpeted as genomic insights, says **Daniel MacArthur**.

hen a study of the genomes of centenarians reported genetic variants strongly associated with exceptional longevity¹, it received widespread media and public interest. It also provoked an immediate sceptical response from other geneticists. That individual genetic variants should have such large effects on a complex human trait was totally unexpected. As it turned out, at least some of the results from this study were surprising simply because they were wrong. In a retraction published a year later², the authors admitted to "technical errors" and "an inadequate quality control protocol". The work was later republished

in a different journal after heavy revision³. Few principles are more depressingly familiar to the veteran scientist: the more surprising a result seems to be, the less likely it is to be true. We cannot know whether, or why, this principle was overlooked in any specific study. However, more generally, in a world in which unexpected results can lead to high-impact publication, acclaim and headlines in *The New York Times*, it is easy to understand how there might be an overwhelming temptation to move from discovery to manuscript submission without performing the necessary data checks.

In fact, it has never been easier to generate high-impact false positives than in the genomic era, in which massive, complex biological data sets are cheap and widely available. To be clear, the majority of genomescale experiments yield real results, many of which would be impossible to uncover through targeted hypothesis-driven studies. However, hunting for biological surprises without due caution can easily yield a rich crop of biases and experimental artefacts, and lead to high-impact papers built on nothing more than systematic experimental 'noise'.

Flawed papers cause harm beyond their authors: they trigger futile projects, stalling the careers of graduate students and postdocs, and they degrade the reputation of genomic research. To minimize the

COMMENT

damage, researchers, reviewers and editors need to raise the standard of evidence required to establish a finding as fact.

Two processes conspire to delude ambitious genomicists. First, the sheer size of the genome means that highly unusual events occur by chance much more often than we would intuitively expect. The limited grasp of statistics that many biologists have and the irresistible appeal of biological findings that neatly fit the facts are a recipe for spurious findings.

Second, all high-throughput genomic technologies come with error modes and systematic biases that, to the unwary eye, can seem like interesting biology. As a result, researchers who are inexperienced with a technology — and some who should know better — can jump to the wrong conclusion.

Again, whether these factors play a part in any specific case is often impossible to know, but several high-profile controversies highlight the potential impact of chance and technical artefacts on genome-scale analyses. For instance, rare loss-of-function mutations in a gene called SIAE were reported to have a large effect on the risk of autoimmune diseases⁴. But a later, combined analysis of more than 60,000 samples⁵ showed no evidence of an association, suggesting that the finding in the original publication was down to chance. Key results in the retracted genetic analysis of longevity mentioned earlier¹ turned out to be errors that arose as a result of combining data from multiple genotyping platforms. And a study published last year that reported widespread chemical modification of RNA molecules⁶ was heavily criticized by experts, who argued that the majority of claimed modifications were, in fact, the product of known classes of experimental error⁷⁻⁹.

Resolving such controversies after results have been published can take time. Even after a strong consensus has emerged among experts in the field that a particular result is spurious, it can take years for that view to reach the broader research community, let alone the public. That provides plenty of opportunity for damage to be done to budding careers and to public trust.

REPLICATION AND REVIEWING

How can the frequency with which technical errors are trumpeted as discoveries be minimized? First, researchers starting out in genomics must keep in mind that interesting outliers — that is, results that deviate significantly from the sample — will inevitably contain a plethora of experimental or analytical artefacts. Identifying these artefacts requires quality-control procedures that minimize the contribution of each to the final result. Finding different ways to make data visual (including simply plotting results across the genome) can be more helpful than many researchers appreciate. The human eye, suitably aided,

428 | NATURE | VOL 487 | 26 JULY 2012 © 2012 Macmillan Publishers Limited. All rights reserved can spot bugs and biases that are difficult or impossible to see in massive data files. Crucially, genomicists should try to replicate technology-driven findings by repeating the study in new samples and using experimental platforms that are not subject to the same error modes as the original technology.

Stringent quality control takes time, a scarce resource in the fast-paced world of genomics. But researchers should weigh the risk of being scooped against the embarrassment of public retraction.

For 'paradigm-shifting' genomics papers, journal editors must recruit reviewers who have enough experience in the specific technologies involved to spot subtle artefacts. Often these will be junior researchers working in the trenches of quality control and manual data inspection. In addition to having the necessary experience, such reviewers often have more time for careful analysis than their supervisors.

Finally, the genomics community must take responsibility for establishing standards for the generation, quality control and statistical analysis of high-throughput data generated using new genomic technologies (a model that has generally worked well, for instance, in genome-wide association studies) and for responding rapidly to published errors. Traditionally, scientists wrote politely outraged letters to journals. Many now voice their concerns in online media, a more rapid and open way to ensure that the public view of a finding is tempered with appropriate caution. Such informal avenues for rapid postpublication discourse should be encouraged.

Nothing can completely prevent the publication of incorrect results. It is the nature of cutting-edge science that even careful researchers are occasionally fooled. We should neither deceive ourselves that perfect science is possible, nor focus so heavily on reducing error that we are afraid to innovate. However, if we work together to define, apply and enforce clear standards for genomic analysis, we can ensure that most of the unanticipated results are surprising because they reveal unexpected biology, rather than because they are wrong.

Daniel MacArthur is based at the Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. e-mail: macarthur@atgu.mgh.harvard.edu

- I. Sebastiani, P. *et al. Science* http://dx.doi.
- org/10.1126/science.1190532 (2010).
- 2. Sebastiani, P. et al. Science **333**, 404 (2011).
- 3. Sebastiani, P. et al. PLoS ONE 7, e29848 (2012).
- 4. Surolia, I. *et al. Nature* **466**, 243–247 (2010).
- 5. Hunt, K. A. et al. Nature Genet. 44, 3-5 (2012).
- 6. Li, M. et al. Science **333**, 53–58 (2011).
- Kleinman, C. L. & Majewski, J. Science 335, 1302 (2012).
- 8. Lin, W. et al. Science 335, 1302 (2012).
- Pickrell, J. K., Gilad, Y. & Pritchard, J. K. Science 335, 1302 (2012).