

Primary research on existing data

Articles in the Analysis format report primary research carried out on publications, datasets or research practices. We see Analyses as a way to generate new hypotheses, test data integrity and promote research integration.

We have recently seen rapid growth of research capacity in next-generation sequencing, high-throughput genotyping and other intensive data-collection methods. Funding bodies are increasingly insisting that data producers make a commitment to ensuring the accessibility and usability of resource-generating projects so to enable not only the grantees' own research but also that of the entire genomics community. These practices, together with collaborative and consortium science, have led to increases in the number and quality of experimental research publications. After having received grants for primary research and with a responsibility to write up their experimental findings, it is not surprising that data producers may have little time to pause and reanalyze their own results. However, we encourage them to do so and also to encourage other data users to look at data 'not created here'.

For this reason, we think it is necessary to draw attention to a number of recent articles published in this journal in the Analysis format (*Nat. Genet.* **39**, 17–23, 2007; **40**, 141–147, 2008; **40**, 827–834, 2008; **40**, 499–507, 2008; **41**, 149–155, 2008). These have addressed various topics, from the epigenetic effects of chromatin on transcriptional regulation *in cis* and *in trans* in yeast to the discovery of an embryonic stem cell–like transcriptional signature in the transcriptomes from poorly differentiated tumors from various tissues of origin.

Research articles are not ideal vehicles for analysis in large numbers, so special tools and practices may need to be created in order to extract essential details and comparable results. For example, curated results may need to be aggregated in a suitable database, as was the case for the Alzheimer's disease and schizophrenia databases (www.alzgene.org/, www.szgene.org/). These databases enable analyses that could not be directly and

iteratively performed on the corpus of published research itself. In the process of curation, the type of tool can be made generic to allow it to address related problems. At the same time, the difficulties encountered by the curator can be fed back to journals to help inform future guidelines for making the essential details of the paper easier to extract for new research purposes. Databases and tools add a dynamic dimension to the Analysis, and they make possible interactive reanalysis of the literature that the original article lacks.

Reanalysis is a direct measure of utility to other researchers, which is the basis for reputation and for a high level of citation. So there is potentially value in Analyses looking at the accessibility of data, the adequacy of metadata and protocols, and the reproducibility of specific results. Such studies—if transparently and responsibly conducted—can be more useful in guiding excellent research practices than any number of standard-setting documents because they directly demonstrate the effects of data formatting, accessibility and precisely described decision making on other researchers' ability to use the published results and data.

Our first few Analyses have set a high standard for originality, and we will need to be very selective, editorially prioritizing for review only those articles distinguished by the quality and rigor of their analytical methods as well as the robustness and utility of their new conclusions. We do not intend the format to be a formula for publication of open-ended meta-analyses, however useful that toolset can be for arriving at translational recommendations. We therefore recommend that prospective authors send a presubmission enquiry to the editorial team via our online manuscript tracking system so that we can advise them at an early stage on the timeliness and likelihood of success of their Analysis proposal. ■