Same computational analysis, different miRNA target predictions

To the editor: Sethupathy et al.[1] compared various microRNA (miRNA) target-prediction programs and present a practical guide for predicting miRNA targets. The programs differ in their requirement for base-pairing of miRNA and target genes and implement similar but not the same criteria when cross-species conservation is applied. Consequently, these programs generate different sets of target genes for some, probably all, miRNAs. Sethupathy et al.[1] presented an online tool that takes sets of target genes predicted by the individual programs and presents compiled lists of the intersection of some programs, providing higher specificity, and the union of all programs, providing higher sensitivity for a particular miRNA (http://www.diana.pcbi.upenn.edu/cgi-bin/TargetCombo.cgi).

We used the same concept of intersection of some of the programs to obtain high specificity and a manageable number of target genes for subsequent experimental verification in the course of the research for one of our projects. We were particularly interested in miRNA has-miR-155 and some others, and we used only the new generation programs, namely PicTar, TargetScanS and miRanda. We noticed that the intersection of target gene sets we obtained for miR-155 differed strongly from that listed by Sethupathy et al.[1], with only one target in common (**Supplementary Table 1** online). Similarly the list of intersections and unions of other miRNAs, for example, miR-146 and miR-16 differed (data not shown).

A possible explanation for this discrepancy is differences in the versions of the mentioned prediction programs. We searched the websites of MiRanda (http://cbio.mskcc.org/cgi-bin/mirnaviewer/mirnaviewer.pl), TargetScanS (http://genes.mit.edu/tscan/targetscanS2005.html) and PicTar (http://pictar.bio.nyu.edu/cgi-bin/new_PicTar_vertebrate.cgi) for miR-155 targets. We are not clear whether the lists of predictions we used were consistent with the precompiled ones mentioned by Sethupathy et al.[1].

Note: Supplementary information is available on the Nature Methods website.

Xiequn Xu

Hartwell Center for Bioinformatics and Biotechnology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA.
e-mail: tony.xu@stjude.org

COMPETING INTERESTS STATEMENT
The author declares that he has no competing financial interests.

1. Sethupathy, P., Megraw, M. & Hatzigeorgiou, A.G. *Nat. Methods* **3**, 881–886 (2006).

Hatzigeorgiou replies: Xu noticed a discrepancy between his results for the unions and intersections of different target prediction programs and the precompiled data on the provided websource TargetCombo (http://www.diana.pcbi.upenn.edu/cgi-bin/TargetCombo.cgi).

He correctly points out that a possible explanation for this discrepancy is differences in the versions of the prediction programs. For the TargetCombo website, we used the precompiled lists provided by TargetScanS, PicTar and MiRanda. The discrepancy can be largely explained by two different sets of predictions provided on the MiRanda website (http://www.microrna.org/mammalian/index.html). One set of predictions is accessible by web search (http://cbio.mskcc.org/cgi-bin/mirnaviewer/mirnaviewer.pl) and another set, used for TargetCombo, is available for direct download in the form of Excel files.

We manually investigated several putative miR-155 targets that are discrepant between the two MiRanda prediction sets. For example, the ATP2B1 gene can be found as a predicted target of miR-155 in the downloadable Excel files but not in the web-searchable set of miRanda predictions. In contrast, the RREB1 gene can be found in the latter but not in the former set of predictions. Finally, genes such as CUGBP2 can be found in both prediction sets. This would explain many of the differences in the results presented by Xu and those on the TargetCombo website.

In the first step of the published analysis, which investigates the performance of different programs on experimentally supported targets, we manually inspected the web-searchable set of miRanda predictions (performed in April 2006). When preparing the TargetCombo website, we used the predictions provided in the miRanda Excel files (July 2006 version) instead because these facilitated a straightforward procedure for automatic data extraction. Additionally, between October 23, 2006 and November 15, 2006 there was a discrepancy on a limited set of MiRanda targets listed in TargetCombo, caused by a parsing error in the web site script. This error has been fixed and in no way affects the analysis as described in the publication. We have also added a cautionary note on the TargetCombo website that multiple versions of some programs, as well as differences in gene-identification systems, can indeed yield noticeably different prediction sets.

Artemis G Hatzigeorgiou

Center for Bioinformatics, Department of Genetics in the School of Medicine, and Department of Computer and Information Science, School of Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.
e-mail: artemis@pcbi.upenn.edu

COMPETING INTERESTS STATEMENT
The author declares that she has no competing financial interests.