

March, 2008

Attention Robert Brown, Accelrys Senior Director of Informatics Marketing:

In light of last summer's Accelrys corporate decision¹ to discontinue GCG (Genetics Computer Group) support, we, the undersigned biochemical, cellular, evolutionary, and genetics molecular scientists of the world, urge Accelrys to release GCG source code to a suitable repository. The specific arrangement doesn't really matter; the point being it is a huge injustice on Accelrys' part to deny continuing development of a vitally important package with over a 25-year legacy of worldwide usage.² Accelrys argues that the legal costs of code release are prohibitive due to the many original authors involved in the primary algorithms, and that much of the code base has been incorporated into their new Pipeline Pilot project. This may be so, and won't be argued against in this forum; however, we would like to argue for a compromise.

In our view, even the sole release of the license and source code to GCG's graphical user interface (GUI) and multiple sequence alignment editor SeqLab, would be a crucial step in continuing the legacy. Nearly all of the functionality of the component programs within the GCG package are available elsewhere, e.g. in EMBOSS³ — this was a primary argument in Accelrys' decision to discontinue GCG support — yet no other existing sequence analysis interface and editor has nearly the power and flexibility of SeqLab. Open source, external components can, and will, be incorporated into SeqLab menus, given its release.

SeqLab's predecessor, the Sun OS specific Genetic Data Environment⁴ (GDE) has been ported to both Mac OS X and Linux, and offers a possible alternative, yet in our opinion SeqLab has far more power, and ease of use. Another viable alternative is Manolo Gouy's SeaView⁵, yet even with recent enhancements, it can only tie in external multiple sequence alignment programs, not an entire suite of sequence analysis tools, nor does it offer sequence feature visualization or an organizational "list" mode, all powerful SeqLab capabilities.

We understand that Steven Smith, the original author of GDE and a member of the initial GCG team that built SeqLab, has expressed interest in continuing to develop the GCG product. If not possible or practical, Mr. Smith has offered to continue the development and support of the SeqLab GUI in much the same fashion as the original GDE tool had been. Therefore, **we hereby highly encourage Accelrys to release the GCG Package and/or SeqLab's license and code base to Steven Smith**, so that he can continue to update and support this valuable bioinformatics resource.

Notes:

¹ See e.g. <http://bio.fsu.edu/~stevet/GCG.death.pdf>.

² The original GCG paper [Devereux, J., Haeblerli, P., and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research*, **12**, 387–395] has been cited by 3,750 scientific articles; and not all authors cite the original paper, 5,246 articles merely cite the *Genetics Computer Group* or the *Wisconsin Package* without referring to the original paper. That is almost 9,000

citations since GCG's advent! To put this number into perspective, the original 1953 Watson and Crick *Nature* DNA structure article has been cited by 185 articles; the 2001 Venter et. al *Science* human genome paper has been cited 686 times (previous citation statistics from NCBI's PubMed Central, as of March 2008). ISI *Web of Science* numbers are even more impressive. 14,075 articles cite the original paper, as of March 2008 according to ISI. These numbers put the Devereux et al. paper in the top-ten most cited scientific papers of all time by several standards. Furthermore, until recently Accelrys advertised on its Web site that the package has been used at more than 950 institutions worldwide.

- 3 See <http://emboss.sourceforge.net/>.
- 4 Smith, S.W., Overbeek, R., Woese, C.R., Gilbert, W., and Gillevet, P.M. (1994) The Genetic Data Environment, an expandable GUI for multiple sequence analysis. *Computer Applications in the Biosciences*, **10**, 671–675. The original Sun OS version is available at <ftp://megasun.bch.umontreal.ca/pub/gde/>. See Linux and Mac OS X GDE ports at <http://www.bioafrica.net/GDElinux/index.html> and <http://www.msu.edu/~lintone/macgde/> respectively.
- 5 Galtier, N., Gouy, M. and Gautier, C. (1996) SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. *Computer Applications in the Biosciences*, **12**, 543–548. See <http://pbil.univ-lyon1.fr/software/seaview.html>.

Primary author:

Steven M. Thompson stevet@bio.fsu.edu Florida State University School of Computational Science

Undersigned petitioners: