PLOS Data Policy Prior to March 3, 2014

Manuscripts submitted to all PLOS journals prior to March 3, 2014 must adhere to the previous PLOS data policy provided below. These manuscripts were not required to include a Data Availability Statement with data availability information.

All manuscripts submitted after March 3, 2014 must include a Data Availability Statement with data availability information in adherence to the current PLOS Data Policy (journals.plos.org/plosbiology/s/data-availability).

Sharing of Materials, Methods, and Data

Publication is conditional upon the agreement of the authors to make freely available any materials and information described in their publication that may be reasonably requested by others for the purpose of academic, non-commercial research.

Availability of data and materials. PLOS is committed to ensuring the availability of data and materials that underpin any articles published in PLOS journals. PLOS's ideal is to make all data relevant to a given article and all readily replaceable materials immediately available without restrictions (while not compromising confidentiality in the context of human-subject research).

We appreciate, however, that this ideal is not yet the norm in all fields. We are therefore collaborating with a number of subject-specific initiatives in order to develop relevant policies. In the meantime, authors must comply with current best practice in their discipline for the sharing of data through databases: for example, deposition of microarray data in ArrayExpress or GEO; deposition of gene sequences in GenBank, EMBL or DDBJ; and deposition of ecological data in Dryad. We encourage all authors to comply with available field-specific standards for the preparation and recording of data; for more information, see below (section 8, Reporting Guidelines for Specific Study Designs). Where no field-specific database exists, authors can deposit data in Dryad.

Failure to comply with this policy will be taken into account when publication decisions are made. PLOS journal editors encourage researchers to contact them if they encounter difficulties in obtaining data or materials from articles published in PLOS journals. PLOS reserves the right to post corrections on articles, to contact authors' institutions and funders, and in extreme cases to withdraw publication, if restrictions on access to data or materials come to light after publication of a PLOS journal article.

Data for which public repositories have been established should be deposited before publication, and the appropriate accession numbers or digital object identifiers (DOIs) published with the paper.

If an appropriate repository does not exist, data should be provided in an open access institutional repository, a general data repository such as Dryad, or as Supporting Information files with the published paper. If none of these options is practical, data should be made freely available upon request.

The conclusions of a study must not depend solely on the analysis of proprietary data. If proprietary data were used to reach a conclusion, and the authors are unwilling or unable to make these data public, then the paper must include an analysis of public data that validates the conclusions so that others can reproduce the analysis and build on the findings. Any restrictions on the availability or use of datasets might be judged to diminish the significance of a paper and may therefore influence the decision about whether a paper should be published. These policies have been developed in accordance with the principles established in *Sharing Publication-Related Data and Materials* (National Academies Press, 2003).

Software. PLOS supports the development of open source software and believes that, for submissions in which software is the central part of the paper, adherence to appropriate open source standards will ensure that the submission conforms to (1) our requirements that methods be described in sufficient detail that another researcher can reproduce the experiments described, (2) our aim to promote openness in research, and (3) our intention that all work published in PLOS journals can be built upon by future researchers. Therefore, if new software or a new algorithm is central to a PLOS paper, the authors must confirm that the software conforms to the Open Source Definition (http://opensource.org/docs/osd/) have deposited the following three items in an open software archive, and included in the submission as Supporting Information:

- **The associated source code of the software described by the paper.** This should, as far as possible, follow accepted community standards and be licensed under a suitable license such as BSD, LGPL, or MIT (see http://www.opensource.org/licenses/alphabetical for a full list). Dependency on commercial software such as Mathematica and MATLAB does not preclude a paper from consideration, although complete open source solutions are preferred.

- **Documentation for running and installing the software.** For end-user applications, instructions for installing and using the software are prerequisite; for software libraries, instructions for using the application program interface are prerequisite.

- **A test dataset with associated control parameter settings.** Where feasible, results from standard test sets should be included. Where possible, test data should not have any dependencies — for example, a database dump.

Acceptable archives should provide a public repository of the described software. The code should be easy to locate and download without the requirement for creating user accounts, logging in or otherwise registering personal details. The repository must have been in existence for over five years or be hosting more than 1,000 projects. Examples of such archives are: SourceForge, Bioinformatics.Org, Open Bioinformatics Foundation (O|B|F), Google

Code, BerliOS Developer, Savannah, GitHub and the Codehaus. Authors should provide a direct link to the deposited software from within the paper.

Deposition with the journal and in an open source archive ensures that the original source associated with the paper is available as well as any enhancements made after the paper is published. An article can be considered for publication if it covers a well-established project that has been providing an open source code repository for an extended amount of time. A condition of acceptance is that the software can be run by reviewers accessing the public software and that the results presented in the paper are reproducible. The software need run on only one hardware-software platform in common use by the readership (including MATLAB), although it must run without dependencies on proprietary or otherwise unobtainable ancillary software. Articles describing software that requires access to databases and other resources whose persistence is not guaranteed (e.g. individual laboratory databases without funding support) will not be considered. In addition, the results described in the paper must be reproducible when peer reviewers, editors, or readers run the software on the deposited dataset and with the provided control parameters.

When the software or algorithm is not central to the paper, we nevertheless encourage authors to make all relevant materials freely available.