Supplemental Figures for:

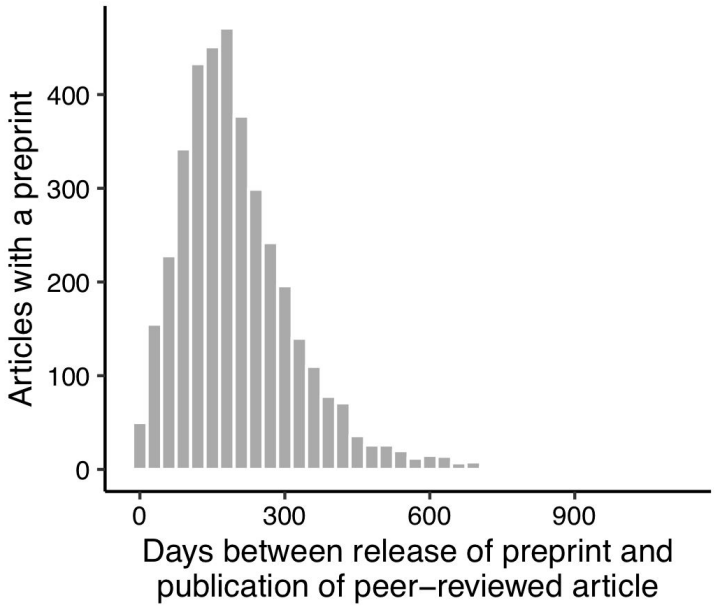
Releasing a preprint is associated with more attention and citations

Darwin Y. Fu[1] and Jacob J. Hughey[1,2,*]

[1]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee;
[2]Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee

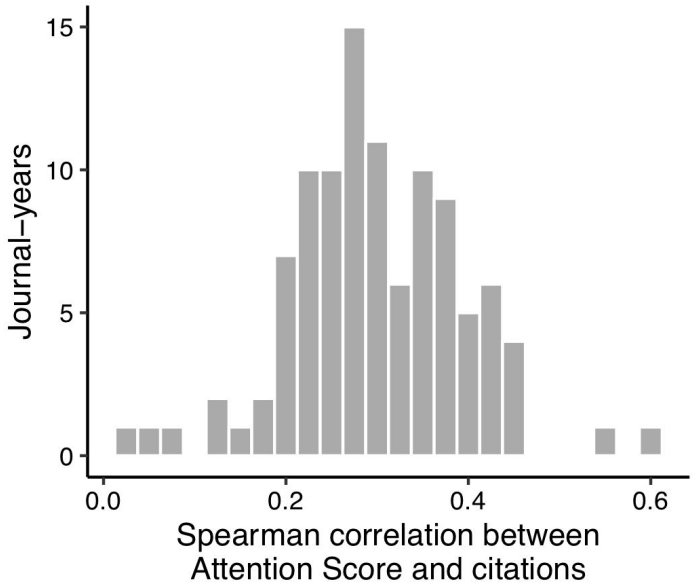
*To whom all correspondence should be addressed: jakejhughey@gmail.com

Figure S1



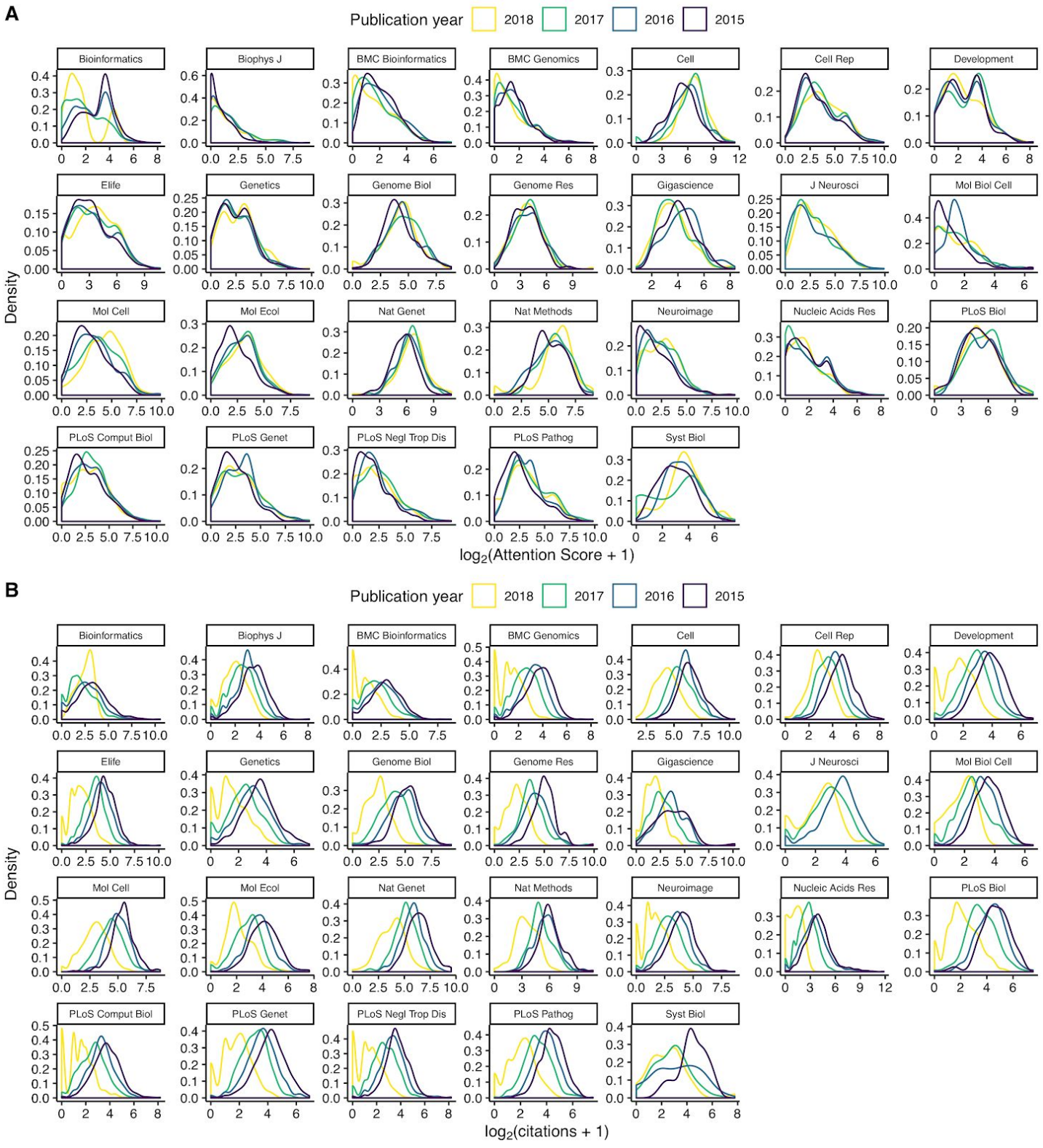
Histogram of the number of days by which release of the preprint preceded publication of the peer-reviewed article, including articles from all journals.

Figure S2



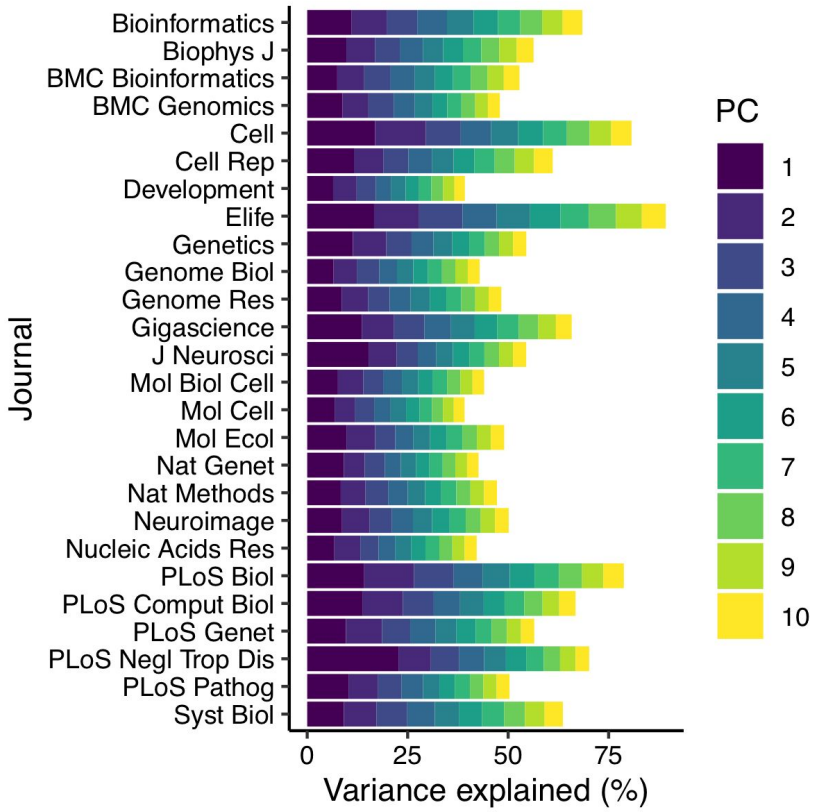
Histogram of Spearman correlation between Altmetric Attention Score and number of citations for each combination of journal and publication year.

Figure S3



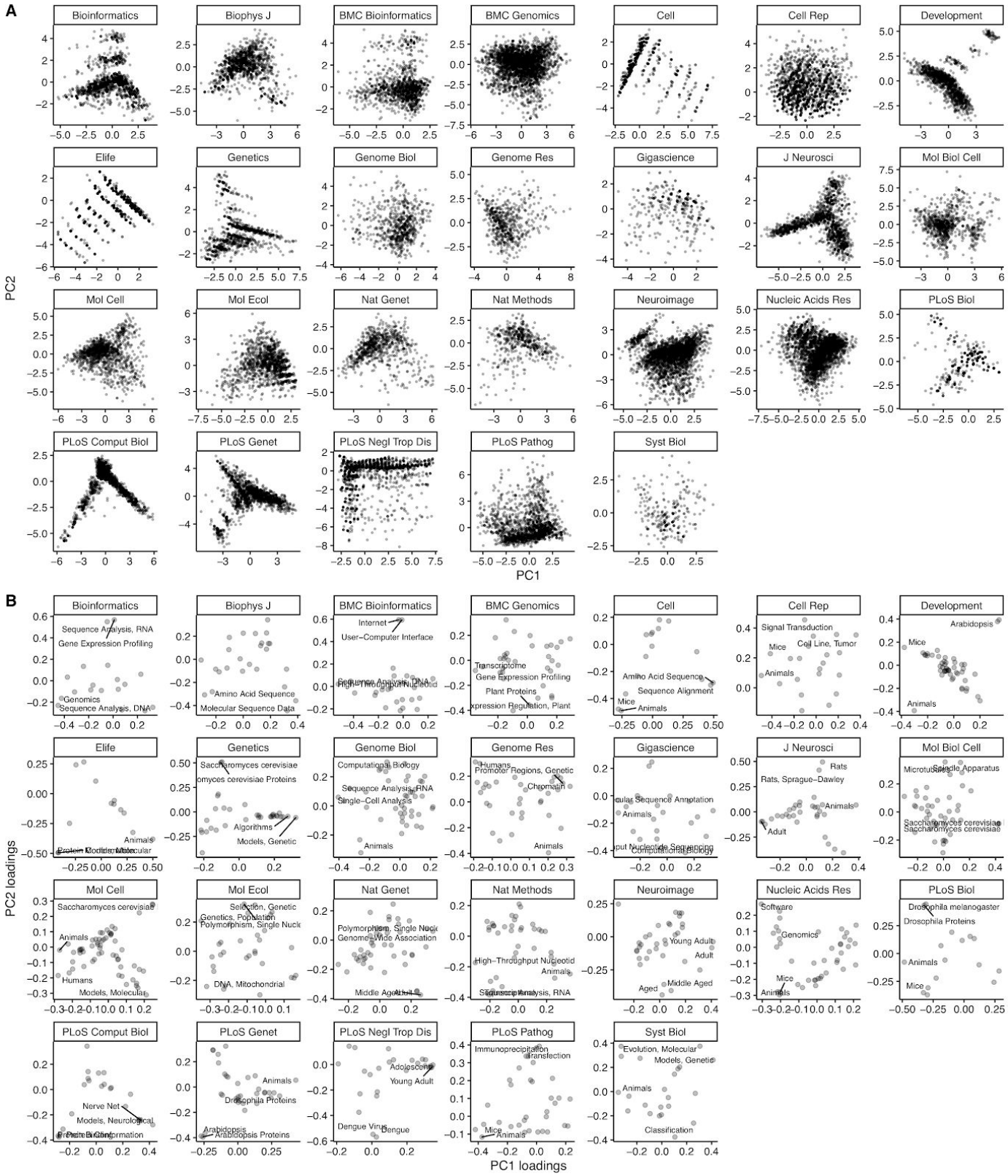
Kernel density estimates of **(A)** log2(Attention Score + 1) and **(B)** log2(citations + 1) for articles in each journal and each publication year. Estimates were computed using the default settings of the geom_density function of the ggplot2 R package.

Figure S4



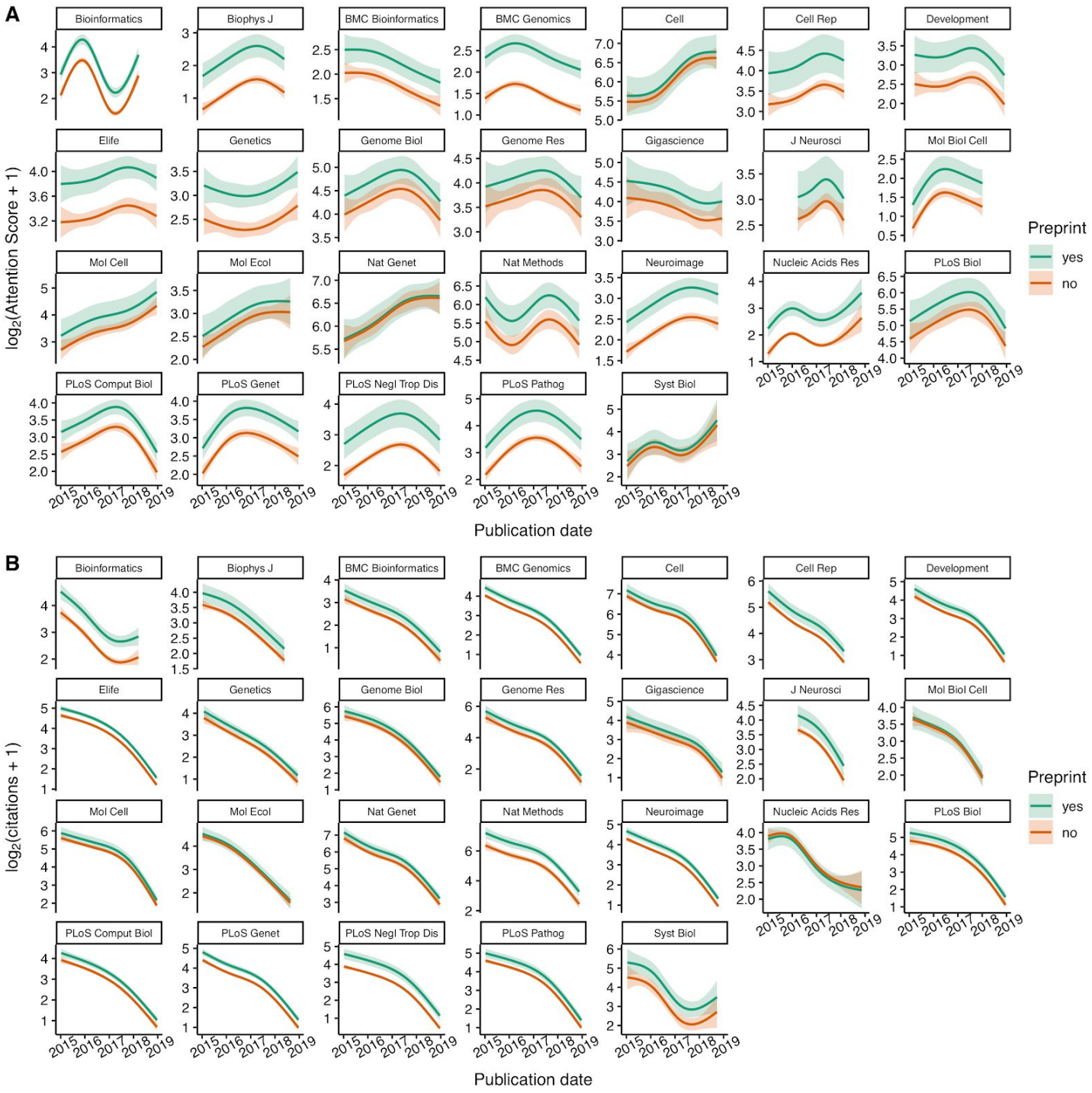
Percentage of variance explained by the top ten principal components of MeSH term assignments for each journal.

Figure S5



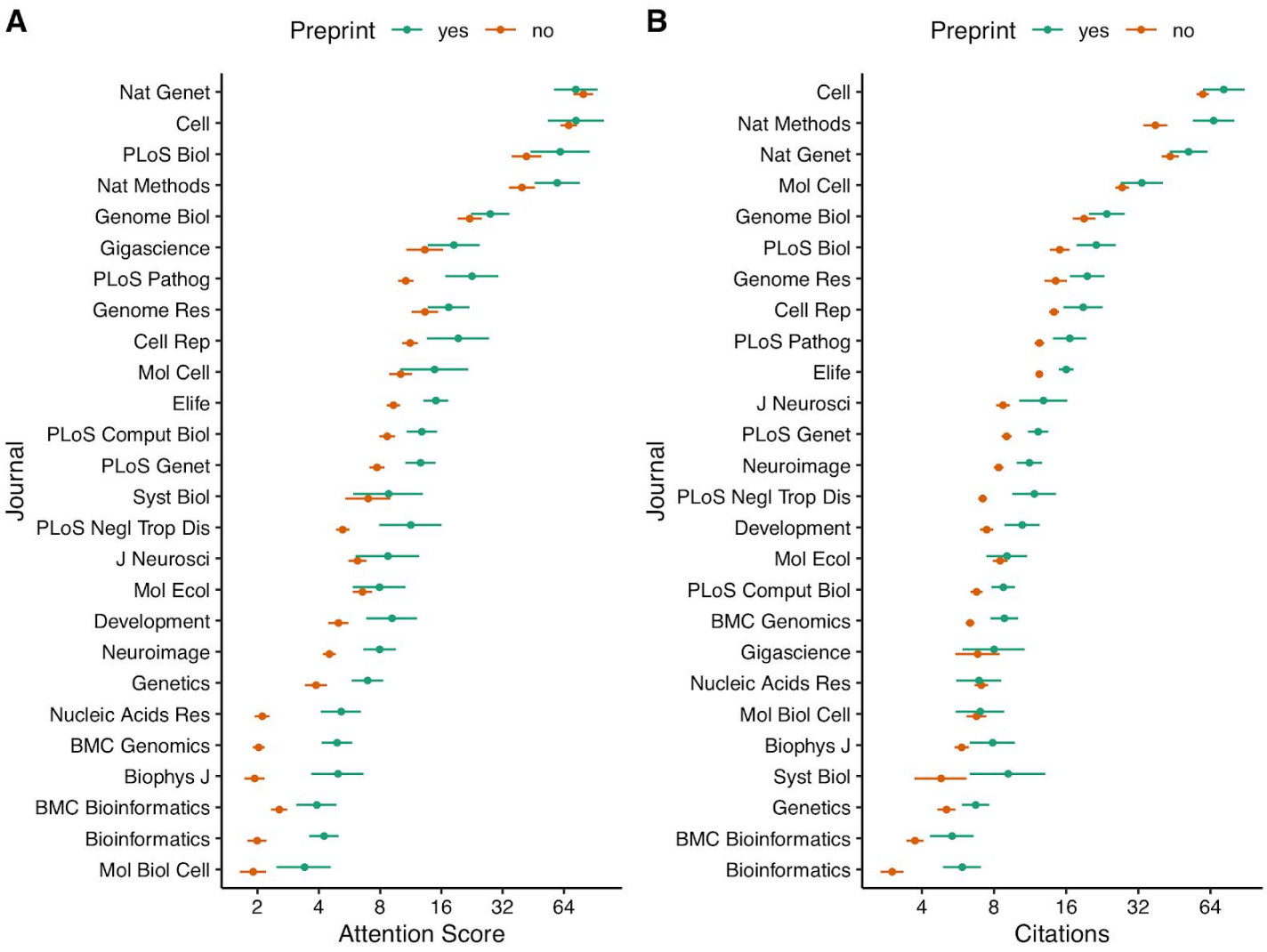
Approximating scientific subfield of peer-reviewed articles using MeSH term assignments. **(A)** Scores and **(B)** loadings for the top two principal components of MeSH term assignments for each journal. In (A), each point corresponds to an article. In (B), each point corresponds to a MeSH term and the two MeSH terms with the highest absolute loading for each PC are labeled.

Figure S6



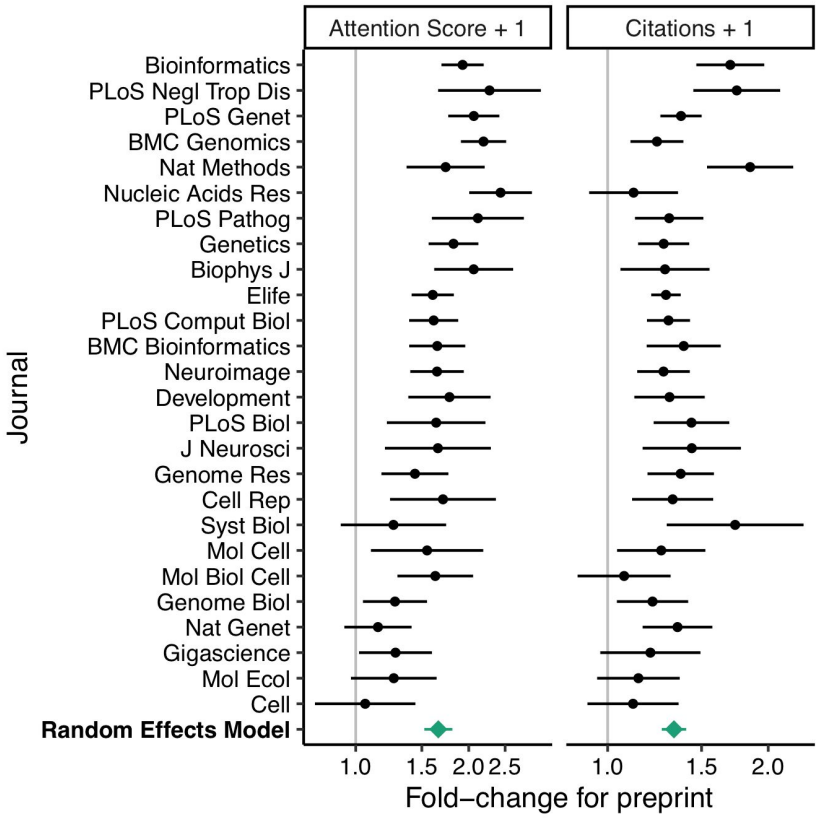
Visualizing the regression fits. Lines show predicted mean **(A)** log2(Attention Score + 1) and **(B)** log2(citations + 1) by journal, preprint status, and publication date, assuming the mean value (i.e., zero) for each of the top ten PCs of MeSH term assignments. Ribbons show 95% confidence intervals. Fits are only shown for date ranges for which articles in a given journal had peer review histories and MeSH terms.

Figure S7



Extending the linear regression model with the number of days by which a preprint preceded its peer-reviewed article. Plots show predicted mean and 95% CI for (A) Altmetric Attention Score and (B) number of citations of a hypothetical article published on January 1, 2017, either with a preprint released 180 days prior or without a preprint. In each plot, journals are sorted by mean of the predicted means between articles with and without a preprint.

Figure S8



Associations between releasing a preprint and the Attention Score and citations of the peer-reviewed article, excluding the PCs of MeSH term assignments from the regression. Error bars indicate 95% CIs. Journals are sorted by mean lower bound of the 95% CI of log2 fold-change. Bottom row shows estimates from random effects meta-analysis, in which releasing a preprint was associated with a 1.66 times higher Attention Score + 1 (95% CI 1.52 to 1.81) and 1.33 times more citations + 1 (95% CI 1.26 to 1.40) of the peer-reviewed article.