The Prevalence of Inappropriate Image Duplication in Biomedical Research Publications

Elisabeth M. Bik[1], Arturo Casadevall*[2,3], Ferric C. Fang*[4]

[1]Department of Medicine, Division of Infectious Diseases, Stanford School of Medicine, Stanford, California, USA

[2]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

[3]Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

[4]Departments of Laboratory Medicine and Microbiology, University of Washington School of Medicine, Seattle, Washington, USA

*contributed equally to this work

Corresponding author: Elisabeth M. Bik, VAHCS Palo Alto, 3801 Miranda Ave, Palo Alto, CA 94034. Telephone: +1 650-493-5000, extension 63163, eliesbik@stanford.edu

Condensed title: **Image duplication prevalence in biomedical papers**

Number of characters: 25,711

ABSTRACT

Inaccurate data in scientific papers can result from honest error or intentional falsification. This study attempted to determine the percentage of published papers containing inappropriate image duplication, a specific type of inaccurate data. The images from a total of 20,621 papers in 40 scientific journals from 1995-2014 were visually screened. Overall, 3.8% of published papers contained problematic figures, with at least half exhibiting features suggestive of deliberate manipulation. The prevalence of papers with problematic images rose markedly during the past decade. Additional papers written by authors of papers with problematic images had an increased likelihood of containing problematic images as well. As this analysis focused only on one type of data, it is likely that the actual prevalence of inaccurate data in the published literature is higher. The marked variation in the frequency of problematic images among journals suggest that journal practices, such as pre-publication image screening, influence the quality of the scientific literature.

Keywords: Research misconduct, ethics in science, biomedical research, peer review

INTRODUCTION

Inaccuracies in scientific papers have many causes. Some result from honest mistakes, such as incorrect calculations, use of the wrong reagent, or improper methodology (Casadevall et al., 2014). Others are intentional and constitute research misconduct, including situations in which data are altered, omitted, manufactured or misrepresented in a way that fits a desired outcome. The prevalence of honest error and misconduct in the scientific literature is unknown. One review estimated the overall frequency of serious research misconduct, including plagiarism, to be 1% (Steneck, 2006). A meta-analysis by Fanelli, combining the results of 18 published surveys, found that 1.9% of researchers have admitted to modification, falsification or fabrication of data (Fanelli, 2009).

There is also little firm information on temporal trends regarding the prevalence of error and misconduct. Research error and misconduct have probably always existed. Even scientific luminaries such as Darwin, Mendel, and Pasteur have been accused of manipulating or misreporting their data (Lüscher, 2013; Geison, 1996). However, the perception of error and misconduct in science has been recently magnified by high profile cases and a sharp rise in the number of retracted manuscripts (Alberts et al., 2015). In recent years, retractions have increased at a rate that is disproportionately greater than the growth of the scientific literature (Steen et al., 2013). Although this could be interpreted as an increase in problematic papers, the actual causes may be more complex and could include a greater inclination by journals and authors to retract flawed work (Steen et al., 2013). Retractions are a poor indicator of error because most retractions result from misconduct (Fang et al., 2012), and many erroneous

62 studies are never retracted (Casadevall et al., 2014). In fact, only a very small fraction of the
63 scientific literature has been retracted. As of April 2016, the PubMed bibliographic database
64 listed 8,735 retracted publications among more than 25 million articles (0.035%).

65 Concerns about misconduct have been accompanied by increasing concerns about the
66 reproducibility of the scientific literature. An analysis of 53 landmark papers in oncology
67 reported that only 6 could be reproduced (Begley and Ellis, 2012), and other pharmaceutical
68 industry scientists have also reported low rates of reproducibility of published findings, which in
69 some cases led to the termination of drug development projects (Prinz et al., 2011). In the field
70 of psychology, less than half of experimental and correlational studies are reportedly
71 reproducible (Open Science Collaboration, 2015). Inaccurate data can result in societal injury.
72 For example, a now-retracted study associating measles vaccination with autism continues to
73 resonate and may be contributing to low vaccination rates (Flaherty, 2011). Corrosion of the
74 literature, whether by error or misconduct, may also impede the progress of science and
75 medicine. For example, false leads may be contributing to increasing disparities between
76 scientific investment and measurable outcomes, such as the discovery of new pharmacological
77 agents (Bowen and Casadevall, 2015).

78 In this study we sought to estimate the prevalence of a specific type of inaccurate data
79 that can be readily observed in the published literature, namely inappropriate image
80 duplication. The results demonstrate that problematic images are disturbingly common in the
81 biomedical literature and may be found in approximately 1 out of every 25 published articles
82 containing photographic image data, in particular Western blots.

83

RESULTS

Papers containing inappropriately duplicated images

A total of 20,621 research papers containing the search term "western blot" from 40 different journals and 14 publishers were examined for inappropriate duplications of photographic images, with or without repositioning or evidence of alteration (Table S1). Of these, 8,138 (39.8%) were published by a single journal (*PLOS ONE*) in 2013 and 2014; the other 12,483 (60.5%) papers were published in 39 journals spanning the years 1995-2014 (Fig. 1). Overall, 782 (3.8%) of these papers were found to include at least one figure containing inappropriate duplications.

Classification of inappropriately duplicated images

Problematic images were classified into three major categories: simple duplications, duplications with repositioning, and duplications with alteration.

- Category I: Simple Duplications. Figures containing two or more identical panels, either within the same figure or between different figures within the same paper, purporting to represent different experimental conditions, were classified as simple duplications. The most common examples in this category were beta-actin loading controls that were used multiple times to represent different experiments or identical microscopy images purporting to be obtained from different experiments. For papers containing such figures, the methods and results were reviewed to establish that the duplicated figures were indeed re-used for different experiments. The re-use of loading controls in

106 different figures obtained from the same experiment was not considered to be a
107 problem. Examples of simple duplication are shown in Fig. 2.

108 - Category II: Duplication with Repositioning. This category included microscopic or blot
109 images with a clear region of overlap, where one image had been shifted, rotated, or
110 reversed in respect to the other. Fig. 3 shows examples of duplicated figures with
111 repositioning.

112 - Category III: Duplication with Alteration. This category consisted of images that were
113 altered with complete or partial duplication of lanes, bands, or groups of cells,
114 sometimes with rotation or reversal in respect to each other, within the same image
115 panel or between panels or figures. This category also includes figures containing
116 evidence of "stamping" in which a defined area is duplicated multiple times within the
117 same image, "patching" in which part of an image is obscured by a rectangular area of
118 different background, and FACS images sharing conserved regions and other regions in
119 which some data points have been added or removed. Examples of duplicated images
120 with alteration are shown in Fig. 4.

121

122 Two additional types of image modification were not scored as problematic, although they may
123 represent questionable research practices by current standards and would not be accepted by
124 certain journals (i.e., *Journal of Cell Biology*):

125 - Cuts. Abrupt vertical changes in the background signal between adjacent lanes in a blot
126 or gel suggest that the lanes were not next to each other in the original gel. These

splices are of potential concern but do not necessarily indicate inaccurate data representation.

- Beautification. Part of the background of a blot or gel where no band of interest is expected may show signs of patching, perhaps to remove a smudge or stain. This is not considered to represent best practice according to contemporary guidelines for data presentation (Rossner and Yamada, 2004) but does not necessarily indicate inaccurate data representation.

Although researcher intent could not be definitively determined in this study, the three categories of duplicated images were felt to have different implications with regard to the likelihood of scientific misconduct. Category I (simple duplication) images are most likely to result from honest errors, in which an author intended to insert two similar images but mistakenly inserted the same image twice. Alternatively, simple duplications may result from misconduct, for example, if an author intentionally recycled a control panel from a different experiment because the actual control was not performed. Category II (duplication with repositioning) and category III (duplication with alteration) may be somewhat more likely to result from misconduct, as conscious effort would be required for these actions.

In our study, a paper was classified as containing an inappropriate duplication when at least one category I, II, or III problem was identified. Papers were classified according to the highest category of duplicated image; e.g., a paper containing both category I and category II images was classified as a category II paper.

147 Among the 782 problematic papers found in this study, 230 (29.4%) contained simple
148 duplications, 356 (45.5%) contained duplicated images with repositioning, while the remaining
149 196 (25.1%) contained duplicated figures with alteration.
150
151 **Temporal trends in image duplication**
152 To investigate the prevalence of image duplications and alterations over time, we plotted the
153 percentage of papers containing inappropriate image duplication as a function of publication
154 year (Fig. 5). The percentage of papers with image duplications appeared to be relatively low (<
155 2%) from 1995-2002, with no problematic images found among the 194 papers screened from
156 1995. However, a sharp increase in the percentage of papers with duplicated images was
157 observed in the year 2003 (3.6%), after which the percentages have remained close to or above
158 4%. This pattern remained very similar when only a subset of 16 journals for which papers were
159 scanned from all 20 years was considered, except for a decline in the duplications found in
160 2014, the last year of our screen (2.2%) (Fig. 5).
161
162 **Correlation of impact factor with image duplication**
163 Substantial variation in the prevalence of papers with image duplication was observed among
164 the 40 journals investigated. In *PLOS ONE*, from which the largest number of papers was
165 screened, 4.3% of the papers were found to contain inappropriately duplicated images,
166 whereas the percentage of papers with image duplication ranged from 0.3% (*Journal of Cell*
167 *Biology)* to 12.4% (*International Journal of Oncology*) among the other journals, with a mean of
168 4.2%. Hence, even though *PLOS ONE* was the journal that provided the largest set of papers

169 evaluated in this study, it is not an outlier with regard to inappropriately duplicated images
170 relative to the other journals examined. To assess the possibility that journals with higher
171 impact factors might be better at detecting problematic images and/or authors might be more
172 careful in preparing images for publication in such journals, the relationship between the
173 prevalence of image duplication and journal impact factor was examined (Fig. 6). For this
174 analysis, only papers published between 2005 and 2014 were included, because the prevalence
175 of problematic images was lower in older publications, and older papers were only evaluated
176 for selected journals. A negative correlation between image duplication and journal impact
177 factor was observed (Pearson's correlation; p-value 0.019), with the lowest percentage of
178 problematic images found in journals with high impact factors. The prevalence of image
179 duplication in 12 open access journals was not significantly different from that in 28 non-open
180 access journals ($p = 0.38$, chi-squared test).
181
182 **Country of origin of papers containing image duplication**
183 To determine whether inappropriate image duplication was more frequent in some countries
184 than in others, the country of origin for each of the 348 papers from *PLOS ONE* containing
185 duplicated images was compared to the country of origin for all papers published by that
186 journal during the same time interval that were included in our search. In cases where the
187 authors of a paper were affiliated with institutions in multiple countries, all countries were
188 taken into account. A majority of the 8,138 screened papers published in *PLOS ONE* during the
189 16-month study period from 2013-2014 were affiliated with China (26.2%) and the US (40.9%)
190 (Fig. 7). However, papers from China had a 1.89-fold higher probability of containing

problematic images than would have been predicted from the frequency of publication (chi-squared test, p-value < 0.001), while papers from the US had a lower probability (0.60-fold) (chi-squared test, p-value < 0.001). Other countries with a higher-than-predicted ratio of papers containing image duplication were India (1.93) and Taiwan (1.20), whereas the prevalence of image duplication was lower than predicted in papers from the UK (0.47), Japan (0.26) and Germany (0.34).

Number of authors and image duplication

Errors or misconduct might be predicted to be more frequent in papers with fewer authors, due to reduced scrutiny. The mean number of authors per paper for the 781 papers containing inappropriate image duplication was 7.28. No significant difference between the mean number of authors for papers with or without image duplication was found ($p > 0.1$).

Problematic images in multiple papers by the same author

Our dataset of 782 problematic papers contained 28 papers (i.e., 14 pairs) of papers with a common first author. To determine whether authors of papers containing inappropriate image duplication were more likely to have published additional papers containing image duplication, we screened other papers written by the first and last authors of 559 papers (all from unique first authors) identified during initial screening. This analysis encompassed 2,425 papers, or a mean of 4.3 additional papers for each primary paper. In 217 cases (38.8%), at least one additional paper containing duplicated images was identified. In total, 269 additional papers containing duplicated images (11.1%) were found out of the 2,425 papers in the secondary

213 dataset. The percentage of papers with duplicated images in the secondary dataset was
214 significantly higher than that of the first dataset (11.1% vs. 3.8%, chi-squared test: $p<0.001$),
215 indicating that other papers by first or last authors of papers with duplicated images have an
216 increased probability of also containing duplicated images.

217 **DISCUSSION**

218

219 The quality and integrity of the scientific literature has been increasingly questioned
220 (Begley and Ellis, 2012; Fang et al., 2012; Ioannidis, 2005; Collins and Tabak, 2014). The present
221 study attempts to empirically determine the prevalence of one type of problematic data,
222 inappropriate image duplication, by visual inspection of electrophoretic, microscopic and flow
223 cytometric data from more than 20,000 recent papers in 40 primary research journals. The
224 major findings of this study are: (1) figures containing inappropriately duplicated images can be
225 readily identified in published papers through visual inspection without the need for special
226 forensic software methods or tools; (2) approximately 1 of every 25 published papers contains
227 inappropriately duplicated images; (3) the prevalence of papers with inappropriate image
228 duplication rose sharply after 2002 and has since remained at increased levels; (4) the
229 prevalence of inappropriate image duplication varies among journals and correlates inversely
230 with journal impact factor; (5) papers containing inappropriately duplicated images originated
231 more frequently from China and India and less frequently from the US, UK, Germany, Japan, or
232 Australia; and (6) other papers by authors of papers containing inappropriately duplicated
233 images often contained duplicated images as well. These findings have important implications
234 for the biomedical research enterprise and suggest a need to improve the literature though
235 greater vigilance and education.

236 The finding that figures with inappropriate duplications can be readily identified by
237 simple inspection suggests that greater scrutiny of publications by authors, reviewers, and
238 editors might be able to identify problematic figures prior to publication. The *Journal of Cell*

239 *Biology* was among the first to call attention to the problem of figure alteration in manuscripts
240 (Rossner, 2002), and this journal instituted a policy to carefully inspect all manuscripts for
241 image manipulation prior to publication (Rossner and Yamada, 2004). The low prevalence of
242 problematic images in this journal (0.3%) suggests that these measures have been effective.
243 The *EMBO Journal*, which was not part of the present study, has also instituted a manual
244 screening process for aberrant images (Pulverer, 2014). Our findings are consistent with the
245 notion that greater scrutiny by journals can reduce the prevalence of problematic images.
246 However, this is likely to require a concerted effort by all journals, so that authors of papers
247 with problematic data do not simply avoid publication in venues that employ rigorous screening
248 procedures.

249 The prevalence of papers containing inappropriate image duplication increased
250 markedly in 2003 and has remained high in subsequent years. This coincides with the observed
251 increase in retracted publications (Fang et al., 2012) and provides empirical evidence that the
252 increased prevalence of problematic data is not simply a result of increased detection, as has
253 been suggested (Fanelli et al., 2015). Although the causes of the increased frequency of image
254 duplication since 2003 are not known, we have considered several possible explanations. First,
255 older papers often contain figures with lower resolution, which may have obscured evidence of
256 manipulation. Second, the widespread availability and usage of digital image modification
257 software in recent years may have provided greater opportunity for both error and intentional
258 manipulation. Third, the increasing tendency for images to be directly prepared by authors
259 instead of by professional photographers working in consultation with authors has removed a
260 potential mechanism of quality control. One possible mechanism to reduce errors at the

261 laboratory level would be to involve multiple individuals in the preparation of figures for
262 publication. The lack of correlation between author number and the frequency of image
263 duplication suggests that the roles of most authors are compartmentalized or diluted, such that
264 errors or misconduct are not readily detected. A fourth consideration is that increasing
265 competition and career-related pressures may be encouraging careless or dishonest research
266 practices (Anderson et al., 2007).

267 A large variation in the prevalence of papers containing inappropriately duplicated
268 images was observed among journals ranging from the *Journal of Cell Biology* (0.3%) to the
269 *International Journal of Oncology* (12.4%), a difference of more than 40-fold. The differences
270 among journals are important because, as noted above, these suggest that journal editorial
271 policies can have a substantial impact on this problem. The inverse correlation between the
272 prevalence of problematic papers and journal impact factor contrasts with the positive
273 correlation observed for research misconduct resulting in retraction (Fang and Casadevall,
274 2011; Fang et al., 2012; Cokol et al., 2007; Liu, 2006). Although the association was weak, this
275 may suggest that higher impact journals are better able to detect anomalous images prior to
276 publication. Alternatively, authors submitting to such journals may be more careful with figure
277 preparation. Nevertheless, we note that even the most highly selective journals contain some
278 papers with figures of concern.

279 China and the United States were responsible for the majority of papers containing
280 inappropriately duplicated images, which is not surprising given the large research output of
281 these countries. However it is noteworthy that the proportion of *PLOS ONE* papers from China
282 and India that were found to contain problematic images was higher than would be predicted

283 from their overall representation, whereas the opposite was true for papers from the US, UK,
284 Germany, Japan and Australia. This suggests that ongoing efforts at scientific ethics reform in
285 China and India should pay particular attention to issues relating to figure preparation (Yidong
286 and Xin, 2006; Patnaik, 2015). The analysis of geographic origin was limited to papers published
287 in *PLOS ONE*, because this journal offered an online tool to search for this information. The
288 geographic distribution of papers with problematic images may be different in other journals.

289 In nearly 40% of the instances in which a problematic paper was identified, screening of
290 other papers from the same authors revealed additional problematic papers in the literature.
291 This suggests that image duplication results from systematic problems in figure preparation by
292 individual researchers, which tend to recur.

293 Our findings suggest that as many as 1 out of every 25 published papers containing
294 Western blot or other photographic images could contain data anomalies. This is likely to be an
295 underestimate of the extent of problematic data in the literature for several reasons. First, only
296 image data were analyzed, thus errors or manipulation involving numerical data in graphs or
297 tables would not have been detected. Second, only duplicated images within the same paper
298 were examined, thus the reuse of images in other papers by the same author(s) would not have
299 been detected. Third, since problematic images were detected by visual inspection, the false-
300 negative rate could not be determined; we readily acknowledge that our screen may have
301 missed problematic papers. It should be noted that our findings contrast with a recent small
302 study by Oksvold that examined 120 papers from 3 different cancer research journals, reporting
303 duplicated images in 24.2% of the papers examined (Oksvold, 2015). Many of the reported
304 image duplications in the Oksvold study involved representation of identical experiments,

305 which we do not regard as necessarily inappropriate, as this form of duplication does not alter
306 the research results. For comparison, we screened 427 papers from the same 3 journals
307 examined by Oksvold and found the average percentage of problematic papers in these
308 journals to be 6.8, which is closer to our findings for other journals. Moreover, our study
309 included more than 20,000 papers from 40 journals; in addition to more rigorous inclusion
310 criteria, we required consensus between three independent examiners for an image to be
311 classified as containing inappropriate duplication, ensuring a low false-positive rate.

312 The high prevalence of inaccurate data in the literature should be a finding of
313 tremendous concern to the scientific community, since the literature is the record of scientific
314 output upon which future research progress depends. Papers containing inaccurate data can
315 reduce the efficiency of the scientific enterprise by directing investigators to pursue false leads
316 or construct unsupportable hypotheses. Although our findings are disturbing, they also suggest
317 specific actions that can be taken to improve the literature. Increased awareness of recurring
318 problems with figure preparation, such as control band duplication, can lead to the reform of
319 laboratory procedures to detect and correct such issues prior to manuscript submission. The
320 variation among journals in the prevalence of problematic papers suggests that individual
321 journal practices, such as image screening, can reduce the prevalence of problematic images
322 (Rossner, 2002; Pulverer, 2014). The problems identified in this study provide further evidence
323 for the scientific establishment that current standards are insufficient to prevent flawed papers
324 from being published. Our findings call for the need of greater efforts to ensure the reliability
325 and integrity of the research literature.

326

327 **MATERIALS AND METHODS**

328

329 **Selection strategy**

330 A total of 20,621 papers were selected from 40 different scientific journals in the fields of
331 microbiology and immunology, cancer biology, and general biology. These journals were
332 published by 14 organizations (average of 2.9 journals per publisher, range 1-6) (Table S1). All
333 journals included in the search were indexed in PubMed, with a mean impact factor of 6.9
334 (range 1.3 - 42.4, Thomson Reuters 2013). Papers were examined if they contained the search
335 term “western blot," using the search tool provided at the journal's website. Only original
336 research papers containing figures were included; retracted papers, review papers, and
337 conference abstracts were excluded. Corrected papers were included only if the correction
338 involved issues other than the problems identified by the present analysis (e.g., incorrect grant
339 statement). From a single journal (*PLOS ONE*), 8,138 papers published in 2013 and 2014 were
340 included in the study, comprising 39.5% of the dataset. From the remaining journals, a mean of
341 320.1 (range 77 - 1070) papers per journal were included. For most of these journals, if more
342 than 50 papers were found in a given year, screening was limited to the first 40-50 papers that
343 were shown in the search field. The selected papers spanned the years 1995-2014, with most
344 papers published in 2013-2014, primarily as a result of the large contribution of papers from
345 *PLOS ONE* (Fig. 1). The large number of *PLOS ONE* papers analyzed reflects both the journal
346 format, which facilitates image analysis, and the fact that *PLOS ONE* is currently the world's
347 largest scientific journal with approximately 30,000 new articles per year
348 (https://en.wikipedia.org/wiki/PLOS_ONE).

349

350 **Visual screening**

351 All papers were screened by examining images at the publisher's website. Although papers
352 were selected using the search term “western blot," all types of photographic image were
353 examined, including protein and nucleic acid electrophoretic gels and blots, histology or
354 microscopy images, and photos of culture plates, plants, and animals. Fluorescence-activated
355 cell sorting (FACS) plots were included as well since these, like photographic images,
356 purportedly represent raw data. Figure panels containing line art such as bar graphs or line
357 graphs were not included in the study. Images within the same paper were visually inspected
358 for inappropriate duplications, repositioning, or possible manipulation (e.g., duplications of
359 bands within the same blot). All papers were initially screened by one of the authors (EMB). If
360 a possible problematic image or set of images was detected, figures were further examined for
361 evidence of image duplication or manipulation using the Adjust Color tool in Preview software
362 on an Apple iMac computer. No additional special imaging software was used. Supplementary
363 figures were not part of the initial search but were examined in papers in which problems were
364 found in images in the primary manuscript. All figures found by the screening author (EMB) to
365 contain possible duplications were independently reviewed by the two co-authors (FCF and AC).
366 Consensus had to be reached among all three authors for a paper to be considered to contain
367 unequivocal evidence of inappropriate figure duplication. Consensus among all three authors
368 was reached in 90.4% of the papers selected during primary screening.

369

370 **Statistical analysis**

371 The R software package was used to plot and analyze data. The stat_smooth method
372 implemented in the R ggplot2 library was applied to find the best fit for a linear regression
373 model on semi-log transformed data examining the relationship between the 2013 Thomson
374 Reuters impact factor and percentage of papers with inappropriately duplicated images for the
375 40 journals included in this study. Pearson's correlation and Pearson's Chi-square tests with
376 Yates' continuity correction were performed in basic R. R code is available as Supplemental
377 Data S1.
378

379 **SUPPLEMENTAL MATERIAL**

380 1. Supplemental Table S1: The 40 journals screened in this study

381 2. Supplemental Data S1: R code

382 **REFERENCES**

383

384 Alberts, B., R.J. Cicerone, S.E. Fienberg, A. Kamb, M. McNutt, R.M. Nerem, R. Schekman, R.
385 Shiffrin, V. Stodden, S. Suresh, M.T. Zuber, B.K. Pope, and K.H. Jamieson. 2015.
386 SCIENTIFIC INTEGRITY. Self-correction in science at work. *Science*. 348:1420–1422.
387 doi:10.1126/science.aab3847.

388 Anderson, M.S., E.A. Ronning, R. De Vries, and B.C. Martinson. 2007. The perverse effects of
389 competition on scientists' work and relationships. *Sci Eng Ethics*. 13:437–461.
390 doi:10.1007/s11948-007-9042-5.

391 Begley, C.G., and L.M. Ellis. 2012. Drug development: Raise standards for preclinical cancer
392 research. *Nature*. 483:531–533. doi:10.1038/483531a.

393 Bowen, A., and A. Casadevall. 2015. Increasing disparities between resource inputs and
394 outcomes, as measured by certain health deliverables, in biomedical research. *Proc Natl*
395 *Acad Sci U S A*. doi:10.1073/pnas.1504955112.

396 Casadevall, A., R.G. Steen, and F.C. Fang. 2014. Sources of error in the retracted scientific
397 literature. *FASEB J*. 28:3847–3855. doi:10.1096/fj.14-256735.

398 Cokol, M., I. Iossifov, R. Rodriguez-Esteban, and A. Rzhetsky. 2007. How many scientific papers
399 should be retracted? *EMBO Rep*. 8:422–423. doi:10.1038/sj.embor.7400970.

400 Collins, F.S., and L.A. Tabak. 2014. Policy: NIH plans to enhance reproducibility. *Nature*.
401 505:612–613.

402 Fanelli, D. 2009. How many scientists fabricate and falsify research? A systematic review and
403 meta-analysis of survey data. *PLoS ONE*. 4:e5738. doi:10.1371/journal.pone.0005738.

404 Fanelli, D., R. Costas, and V. Larivière. 2015. Misconduct Policies, Academic Culture and Career
405 Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. *PLoS ONE*.
406 10:e0127556. doi:10.1371/journal.pone.0127556.
407 Fang, F.C., and A. Casadevall. 2011. Retracted science and the retraction index. *Infect Immun*.
408 79:3855–3859. doi:10.1128/IAI.05661-11.
409 Fang, F.C., R.G. Steen, and A. Casadevall. 2012. Misconduct accounts for the majority of
410 retracted scientific publications. *Proc Natl Acad Sci U S A*. 109:17028–17033.
411 doi:10.1073/pnas.1212247109.
412 Flaherty, D.K. 2011. The vaccine-autism connection: a public health crisis caused by unethical
413 medical practices and fraudulent science. *Ann Pharmacother*. 45:1302–1304.
414 doi:10.1345/aph.1Q318.
415 Friedman-Levi, Y., M. Mizrahi, K. Frid, O. Binyamin, and R. Gabizon. 2013. PrP(ST), a soluble,
416 protease resistant and truncated PrP form features in the pathogenesis of a genetic
417 prion disease. *PLoS ONE*. 8:e69583. doi:10.1371/journal.pone.0069583.
418 Friedman-Levi, Y., M. Mizrahi, K. Frid, O. Binyamin, and R. Gabizon. 2015. Correction: PrPST, a
419 Soluble, Protease Resistant and Truncated PrP Form Features in the Pathogenesis of a
420 Genetic Prion Disease. *PLoS ONE*. 10:e0133911. doi:10.1371/journal.pone.0133911.
421 Geison, G.L. 1996. The Private Science of Louis Pasteur. illustrated, revised. Princeton University
422 Press.
423 Ioannidis, J.P. 2005. Why most published research findings are false. *PLoS Med*. 2:e124.
424 doi:10.1371/journal.pmed.0020124.
425 Liu, J., P. Xu, C. Collins, H. Liu, J. Zhang, J.P. Keblesh, and H. Xiong. 2013. HIV-1 Tat protein

426 increases microglial outward K(+) current and resultant neurotoxic activity. *PLoS ONE*.
427 8:e64904. doi:10.1371/journal.pone.0064904.
428 Liu, S.V. 2006. im1.biz/TopRetraction.pdf. *Scientific Ethics*. 1:91–93.
429 Liu, Y., Y. Liu, C. Sun, L. Gan, L. Zhang, A. Mao, Y. Du, R. Zhou, and H. Zhang. 2014. Carbon ion
430 radiation inhibits glioma and endothelial cell migration induced by secreted VEGF. *PLoS*
431 *ONE*. 9:e98448. doi:10.1371/journal.pone.0098448.
432 Liu, Y., Y. Liu, C. Sun, L. Gan, L. Zhang, A. Mao, Y. Du, R. Zhou, and H. Zhang. 2015. Correction:
433 Carbon Ion Radiation Inhibits Glioma and Endothelial Cell Migration Induced by Secreted
434 VEGF. *PLoS ONE*. 10:e0135508. doi:10.1371/journal.pone.0135508.
435 Lüscher, T.F. 2013. The codex of science: honesty, precision, and truth--and its violations. *Eur*
436 *Heart J*. 34:1018–1023. doi:10.1093/eurheartj/eht063.
437 Oksvold, M.P. 2015. Incidence of Data Duplications in a Randomly Selected Pool of Life Science
438 Publications. *Sci Eng Ethics*. doi:10.1007/s11948-015-9668-7.
439 Open Science Collaboration. 2015. PSYCHOLOGY. Estimating the reproducibility of psychological
440 science. *Science*. 349:aac4716. doi:10.1126/science.aac4716.
441 Patnaik, P.R. 2015. Scientific Misconduct in India: Causes and Perpetuation. *Sci Eng Ethics*.
442 doi:10.1007/s11948-015-9677-6.
443 Prinz, F., T. Schlange, and K. Asadullah. 2011. Believe it or not: how much can we rely on
444 published data on potential drug targets? *Nat Rev Drug Discov*. 10:712.
445 doi:10.1038/nrd3439-c1.
446 Pulloor, N.K., S. Nair, K. McCaffrey, A.D. Kostic, P. Bist, J.D. Weaver, A.M. Riley, R. Tyagi, P.D.
447 Uchil, J.D. York, S.H. Snyder, A. García-Sastre, B.V. Potter, R. Lin, S.B. Shears, R.J. Xavier,

448 and M.N. Krishnan. 2014. Human genome-wide RNAi screen identifies an essential role
449 for inositol pyrophosphates in Type-I interferon response. *PLoS Pathog*. 10:e1003981.
450 doi:10.1371/journal.ppat.1003981.
451 Pulverer, B. 2014. EMBO Press - a new way to publish. *EMBO J*. 33:1–2.
452 doi:10.1002/embj.201387566.
453 Retraction for Xu et al., Reactive oxygen species-triggered trophoblast apoptosis is initiated by
454 endoplasmic reticulum stress via activation of caspase-12, CHOP, and the JNK pathway
455 in Toxoplasma gondii infection in mice. 2015. *Infect Immun*. 83:1735.
456 doi:10.1128/IAI.00118-15.
457 Rossner, M. 2002. Figure manipulation: assessing what is acceptable. *J Cell Biol*. 158:1151–
458 1151. doi:10.1083/jcb.200209084.
459 Rossner, M., and K.M. Yamada. 2004. What's in a picture? The temptation of image
460 manipulation. *J Cell Biol*. 166:11–15. doi:10.1083/jcb.200406019.
461 Steen, R.G., A. Casadevall, and F.C. Fang. 2013. Why has the number of scientific retractions
462 increased? *PLoS ONE*. 8:e68397. doi:10.1371/journal.pone.0068397.
463 Steneck, N.H. 2006. Fostering integrity in research: definitions, current knowledge, and future
464 directions. *Sci Eng Ethics*. 12:53–74. doi:10.1007/s11948-006-0006-y.
465 The PLOS ONE Staff. 2014. Correction: HIV-1 Tat Protein Increases Microglial Outward K+
466 Current and Resultant Neurotoxic Activity. *PLoS ONE*. 9:e109218.
467 doi:10.1371/journal.pone.0109218.
468 The PLOS Pathogens Staff. 2014. Correction: Human Genome-Wide RNAi Screen Identifies an
469 Essential Role for Inositol Pyrophosphates in Type-I Interferon Response. *PLoS Pathog*.

470 10:e1004519. doi:10.1371/journal.ppat.1004519.

471 Xu, X., T. Liu, A. Zhang, X. Huo, Q. Luo, Z. Chen, L. Yu, Q. Li, L. Liu, Z.R. Lun, and J. Shen. 2012.
472 Reactive oxygen species-triggered trophoblast apoptosis is initiated by endoplasmic
473 reticulum stress via activation of caspase-12, CHOP, and the JNK pathway in Toxoplasma
474 gondii infection in mice. *Infect Immun*. 80:2121–2132. doi:10.1128/IAI.06295-11.

475 Yidong, G., and H. Xin. 2006. Research ethics. China's science ministry fires a barrage of
476 measures at misconduct. *Science*. 312:1728–1729. doi:10.1126/science.312.5781.1728a.

477 Zou, X., B.S. Sorenson, K.F. Ross, and M.C. Herzberg. 2013. Augmentation of epithelial
478 resistance to invading bacteria by using mRNA transfections. *Infect Immun*. 81:3975–
479 3983. doi:10.1128/IAI.00539-13.

480 Zou, X., B.S. Sorenson, K.F. Ross, and M.C. Herzberg. 2015. Correction for zou et Al.,
481 augmentation of epithelial resistance to invading bacteria by using mRNA transfections.
482 *Infect Immun*. 83:1226–1227. doi:10.1128/IAI.02991-14.

483

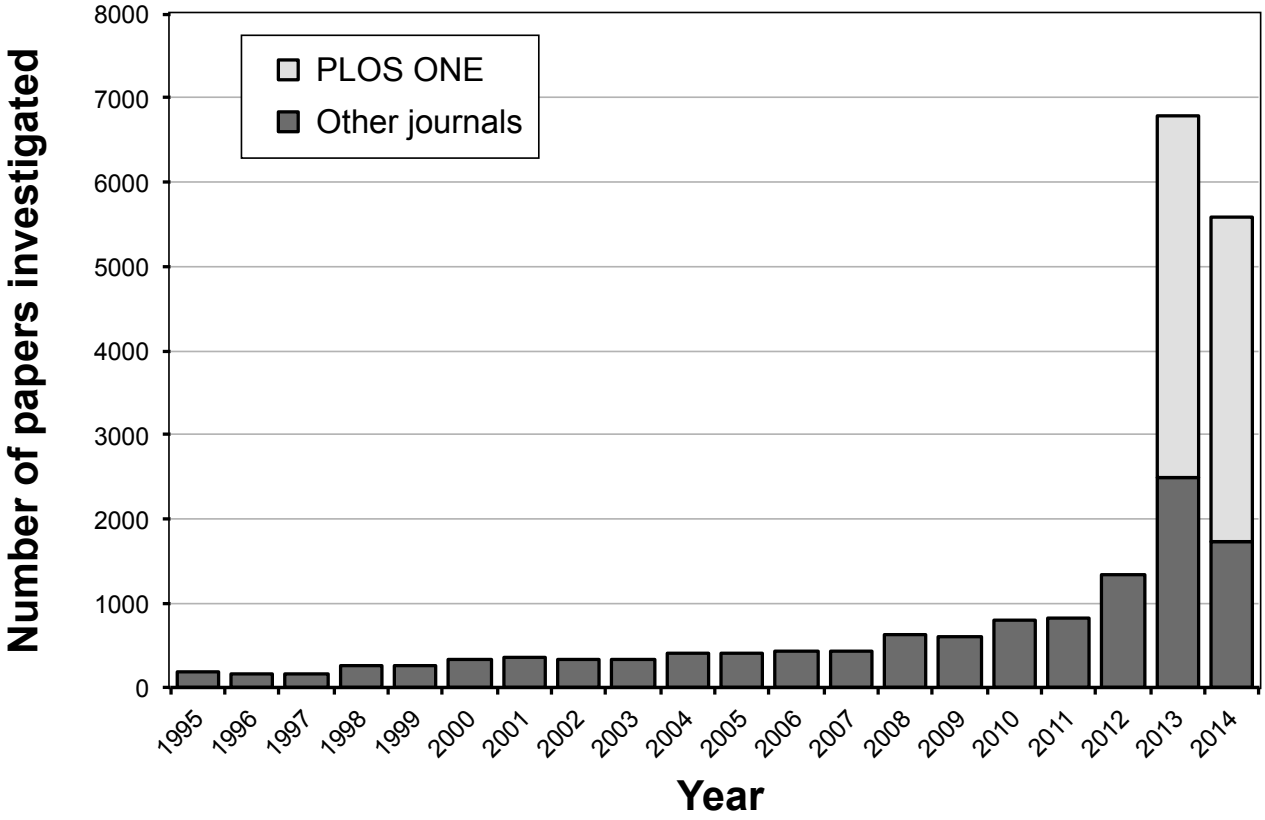
484

485

486

487 **FIGURES and LEGENDS**

488



489

490

491 **Figure 1. Publications investigated in this study by year of publication.** The majority of

492 screened papers were published in 2013 and 2014 due to the large proportion of *PLOS ONE*

493 papers (39.5%) in the dataset. The lowest number of papers (n=151) screened in this study was

494 published in 1996.

495

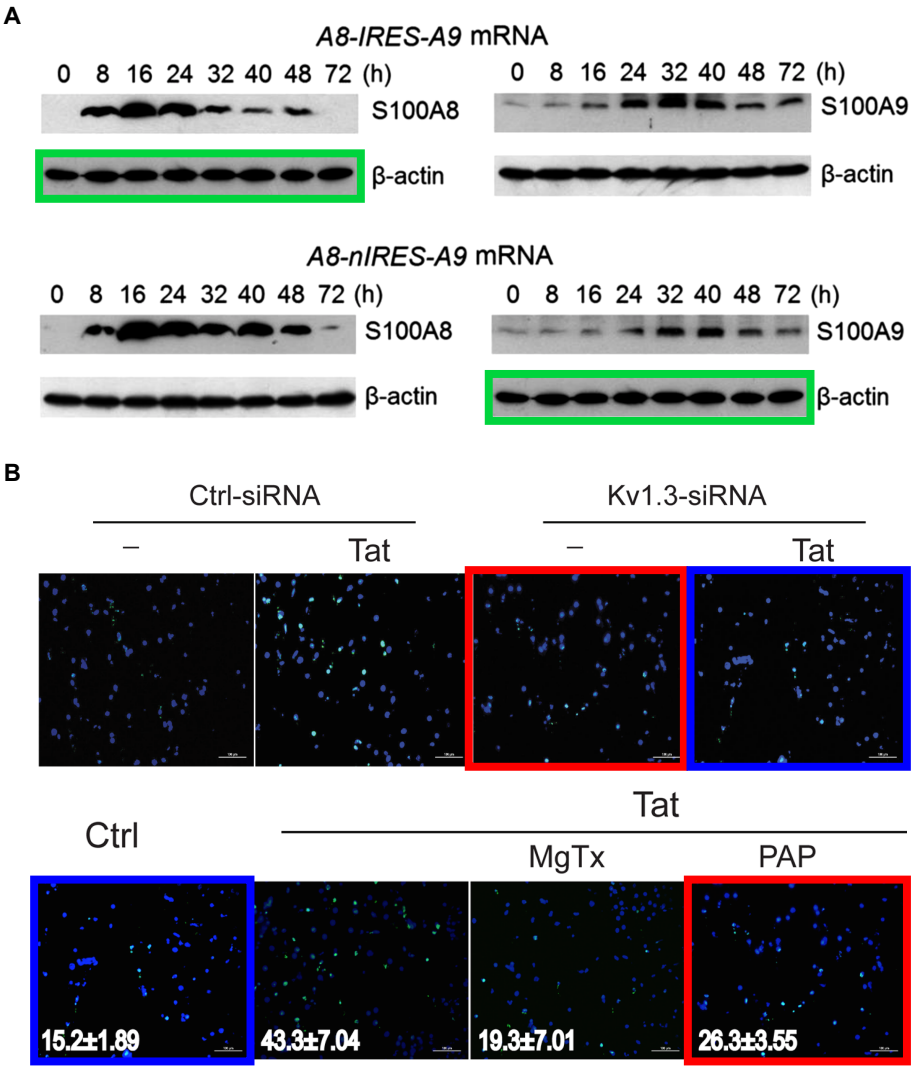
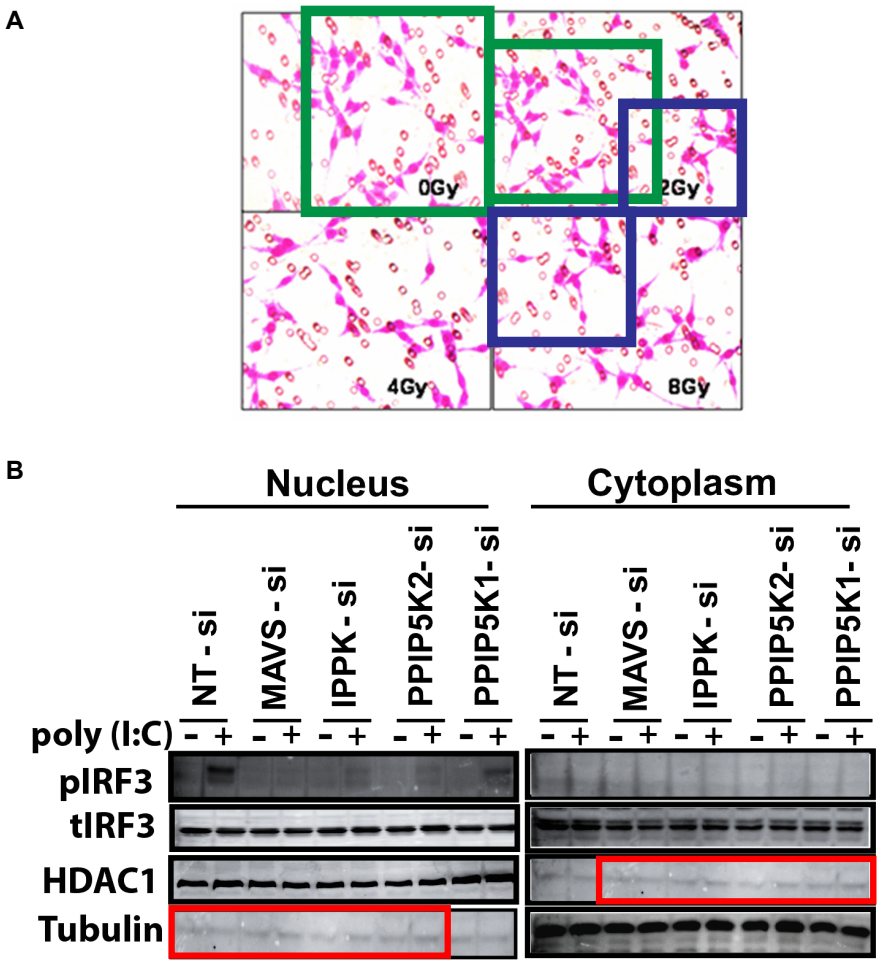


Figure 2. Examples of simple duplication (Category I). A. The beta-actin control panel in the top left is identical to the panel in the bottom right (shown by green boxes), although each panel represents a different experimental condition (Zou et al., 2013); corrected in: (Zou et al., 2015). Figure reproduced with permission from the publisher. **B.** The panels shown here were derived from two different figures within the same paper. Two of the top panels appear identical to two of the bottom panels, but they represent different experimental condition (shown with red and blue boxes) (Liu et al., 2013); corrected in: (The PLOS ONE Staff, 2014). Figure reproduced under the Creative Commons (CC BY) license. All duplications might have been caused by honest errors during assembly of the figures.



506

507 **Figure 3. Examples of duplication with repositioning (Category II). A.** Although the panels

508 represent four different experimental conditions, three of the four panels appear to show a

509 region of overlap (shown with green and blue boxes), suggesting that these photographs were

510 actually obtained from the same specimen (Liu et al., 2014); corrected in: (Liu et al., 2015)**. B.**

511 Western blot panels "Nucleus-Protein D" and "Cytoplasm Protein C" purportedly depict

512 different proteins and cellular fractions, but the blots appear very similar albeit shifted by two

513 lanes (shown with red boxes) (Pulloor et al., 2014); corrected in: (The PLOS Pathogens Staff,

514 2014). Both figures reproduced under the Creative Commons (CC BY) license.

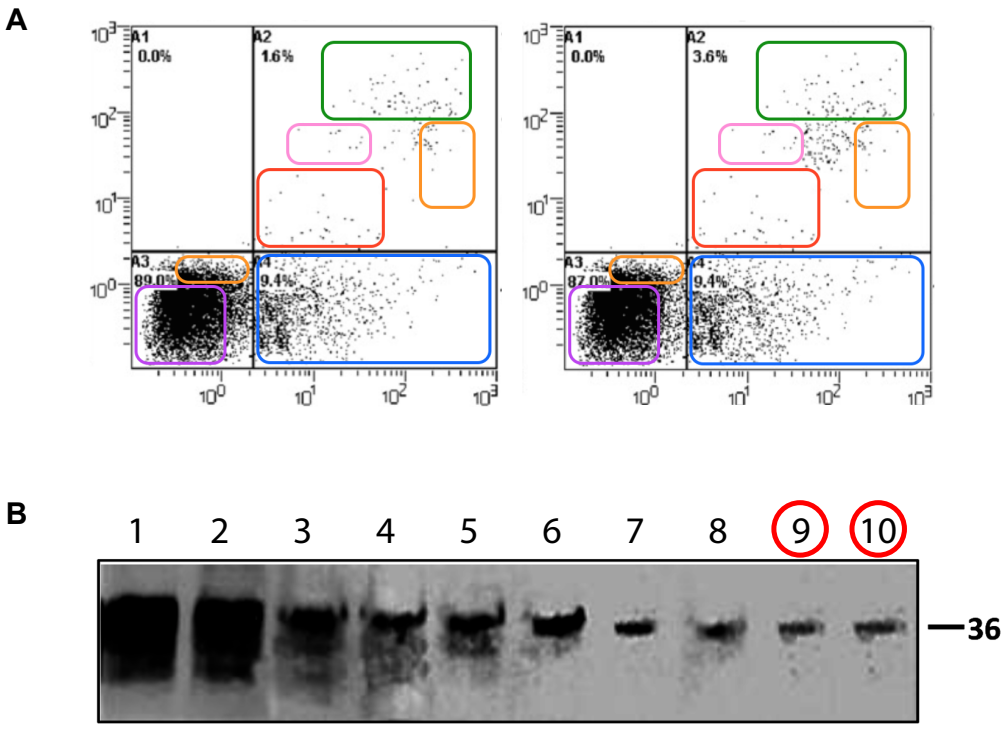


Figure 4. Examples of duplication with alteration (Category III). A. The left and right FACS panels represent different experimental conditions and show different percentages of cell subsets, but regions of identity (colored boxes) between the panels suggest that the images have been altered (Xu et al., 2012); retracted in (Retraction for Xu et al., 2015). Figure reproduced with permission from the publisher. **B.** The figure shown here displays a Western blot of 10 different protein fractions isolated from a density gradient. The figure appears to show a single blot, but the last two lanes (highlighted with red circles) appear to contain an identical band. Exposure was altered to bring out details (Friedman-Levi et al., 2013); corrected in: (Friedman-Levi et al., 2015). Figure reproduced under the Creative Commons (CC BY) license.

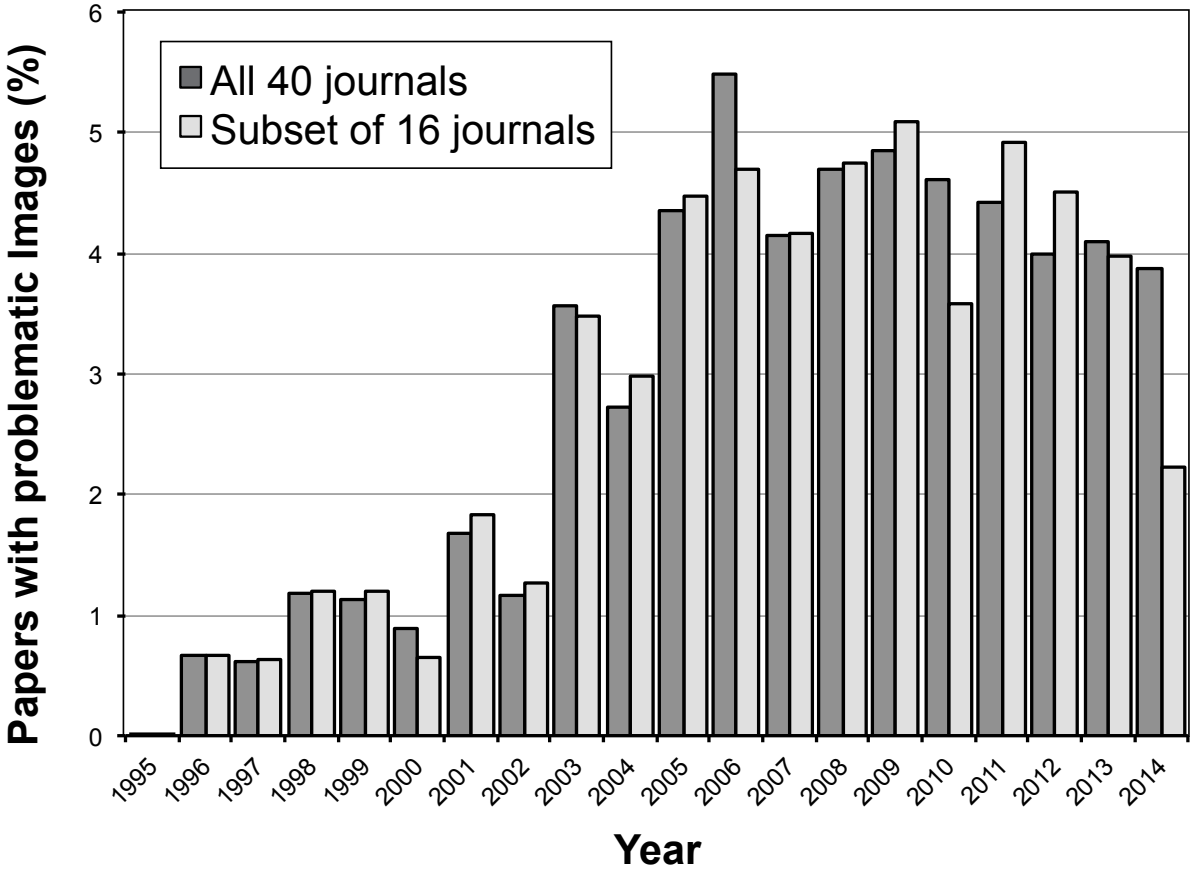
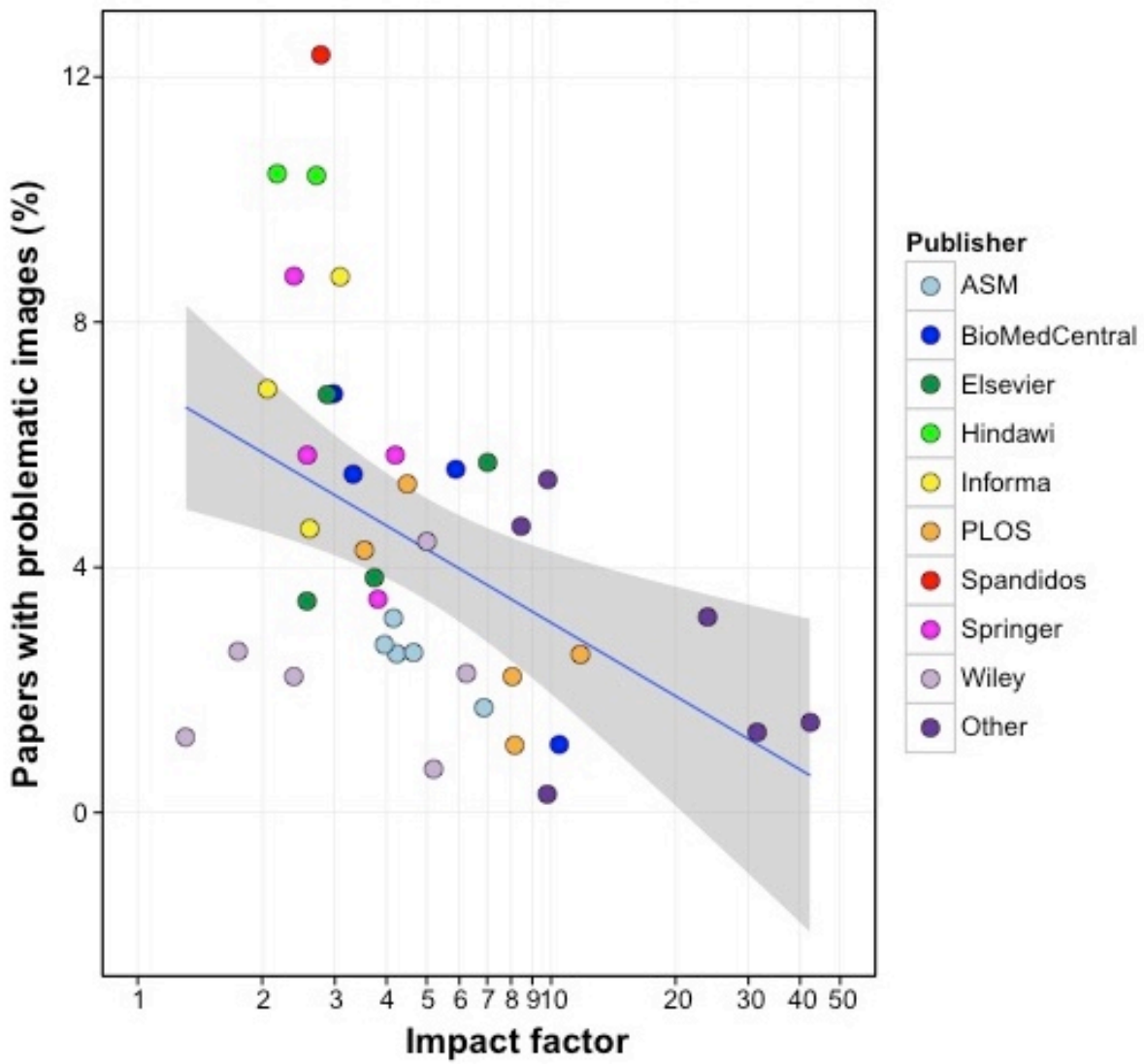


Figure 5. Percentage of papers containing inappropriate image duplications by year of publication. No papers with duplications were found in 1995. The dark gray bars show the data for all 40 journals. The light gray colored bars show a subset of 16 journals for which papers spanning the complete timespan of 20 years were scanned. The total numbers of papers screened in each year are shown in Figure 1.



534

535 **Figure 6. Correlation between journal impact factor and percentage of papers with image**
536 **duplication.** Only papers from 2005-2014 (n = 17,816) were included in this analysis. Each data
537 point represents a journal included in this study (n=40), with data points color-coded according
538 to their publisher (n=14; journals published by AAAS, Nature, Cell Press, the National Academy
539 of Sciences, and the Rockefeller University Press are grouped under "Other"). The x-axis is
540 shown on a logarithmic scale due to the small number of journals with a high impact factor
541 included in this study. The blue line shows a linear regression model. The grey zone depicts the
542 95% confidence interval.

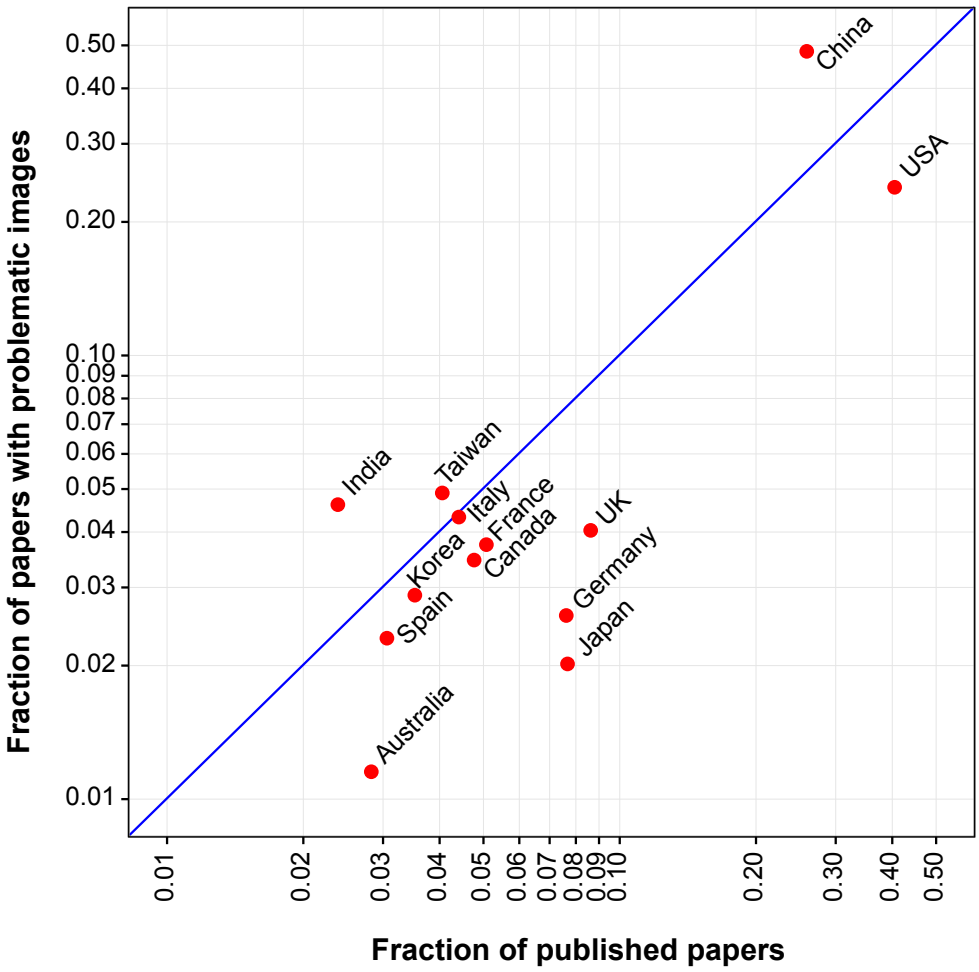


Figure 7. Proportion of papers with image duplications by country. The proportion of papers affiliated with specific countries submitted to *PLOS ONE* during a 16-month period in the years 2013 and 2014 (n = 8,138) is plotted versus the proportion of papers containing inappropriate image duplication affiliated with specific countries (n = 348). Each data point represents a country for which 100 or more papers were screened. Some papers were affiliated with more than one country. The blue line represents the line where data points are expected to fall if problematic papers are distributed as expected according to their representation in the journal. Countries plotted above the blue line had a higher-than-expected proportion of problematic papers; countries plotted below the line had a lower-than-expected ratio.

553 **SUPPLEMENTARY TABLE**

554

555 **Table S1. The 40 journals screened in this study.** Table includes publisher, impact factor

556 (Thomson Reuters 2013), number of papers containing the term "western blot" (WB) screened

557 per year, and number of papers with inappropriate image duplication (IDs) found in that year.

Bik et al. Supplementary Table 1

Journal Title	Publisher	Impact Factor 2013	Screened	Papers with ID	% ID	2014		2013		2012		2011		2010		2009		2008		2007		2006		2005		2004		2003		2002		2001		2000		1999		1998		1997		1996		1995	
						WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID
PLOS ONE	PLOS	3.53	**8138**	**348**	**4.28**	3846	167	4292	181	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS Pathogens	PLOS	8.06	**406**	**9**	**2.22**	200	5	206	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS Genetics	PLOS	8.17	**362**	**4**	**1.10**	164	2	198	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS Biology	PLOS	11.77	**233**	**6**	**2.58**	36	2	42	0	47	1	47	1	61	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS NTD	PLOS	4.49	**317**	**17**	**5.36**	72	3	79	4	53	3	37	3	42	3	17	0	15	1	2	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Journal of Clinical Microbiology	ASM	4.23	**595**	**11**	**1.85**	13	0	10	0	12	0	15	0	19	0	18	1	20	0	11	**0**	31	2	44	2	55	3	45	2	55	0	57	0	49	0	58	0	46	1	13	0	8	0	16	0
Applied and Environmental Microbiology	ASM	3.95	**292**	**8**	**2.74**	56	0	46	1	65	0	60	3	65	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
mBio	ASM	6.88	**175**	**3**	**1.71**	64	0	53	1	30	1	22	1	6	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infection and Immunity	ASM	4.16	**1070**	**30**	**2.80**	120	2	95	5	121	6	40	2	51	2	51	2	51	0	53	0	52	1	58	2	52	1	42	3	47	0	45	1	49	1	51	1	47	0	11	0	14	1	19	0
Journal of Virology	ASM	4.65	**421**	**11**	**2.61**	NA	NA	421	11	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
International journal of cancer	Wiley	5.01	**226**	**10**	**4.42**	40	0	186	10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Clinical Microbiology and Infection	Wiley	5.20	**199**	**1**	**0.50**	9	0	6	0	10	1	11	0	14	0	23	0	17	0	14	0	21	0	15	0	15	0	7	0	5	0	11	0	6	0	3	0	5	0	4	0	1	0	2	0
Journal of Applied Microbiology	Wiley	2.39	**200**	**3**	**1.50**	15	0	19	0	8	0	14	0	11	1	12	0	19	0	23	1	6	1	8	0	9	0	7	0	11	0	14	0	6	0	9	0	5	0	4	0	NA	NA	NA	NA
Environmental Microbiology	Wiley	6.24	**189**	**5**	**2.65**	49	1	24	0	19	2	18	0	22	0	17	1	16	0	5	0	3	0	3	0	2	0	2	0	4	0	0	0	3	1	2	0	NA	NA	NA	NA	NA	NA	NA	NA
Microbiology and Immunology	Wiley	1.31	**358**	**3**	**0.84**	19	0	29	0	23	0	22	0	19	1	17	0	19	0	6	0	3	0	5	1	7	1	7	0	18	0	18	0	35	0	23	0	27	0	27	0	18	0	16	0
Letters in Applied Microbiology	Wiley	1.75	**123**	**2**	**1.63**	10	0	6	0	7	2	10	0	17	0	7	0	6	0	3	0	8	0	2	0	6	0	7	0	2	0	6	0	7	0	8	0	1	0	2	0	5	0	3	0
BioMed Research International	Hindawi	2.71	**77**	**8**	**10.39**	42	4	35	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Evid Based Compl Alternat Med	Hindawi	2.18	**96**	**10**	**10.42**	20	2	34	4	25	3	12	1	3	0	1	0	0	0	0	0	1	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BMC Microbiology	BMC	2.98	**340**	**23**	**6.76**	44	3	44	3	28	0	50	3	56	6	39	3	37	2	11	**0**	8	2	5	0	5	0	2	1	5	0	6	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Genome Biology	BMC	10.47	**105**	**1**	**0.95**	24	0	15	1	8	0	7	0	4	0	6	0	6	0	9	0	6	0	5	0	5	0	5	0	2	0	3	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Breast Cancer Research	BMC	5.88	**403**	**20**	**4.96**	58	4	38	2	49	5	49	0	36	2	24	0	31	1	25	2	12	2	35	2	13	0	4	0	3	0	5	0	17	0	4	0	NA	NA	NA	NA	NA	NA	NA	NA
BMC Cancer	BMC	3.32	**145**	**8**	**5.52**	NA	NA	145	8	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Diagn Microbiol Infect Dis	Elsevier	2.57	**115**	**3**	**2.61**	6	0	7	0	14	1	15	2	9	0	6	0	8	0	14	0	5	0	3	0	3	0	6	0	4	0	5	0	1	0	4	0	3	0	0	0	1	0	1	0
Lung Cancer	Elsevier	3.74	**317**	**11**	**3.47**	12	0	24	0	38	1	28	1	25	2	28	1	23	1	23	1	20	2	14	0	15	0	29	1	14	0	13	1	1	0	2	0	4	0	0	0	2	0	2	0
Cytokine	Elsevier	2.87	**464**	**28**	**6.03**	31	1	55	2	66	3	60	5	30	1	28	4	39	4	15	**1**	35	4	37	2	23	1	15	0	6	0	2	0	6	0	3	0	3	0	4	0	4	0	2	0
Journal of Autoimmunity	Elsevier	7.02	**150**	**6**	**4.00**	8	1	13	0	13	0	11	0	11	0	6	1	8	0	6	1	16	1	13	2	14	0	8	0	1	0	2	0	2	0	3	0	3	0	3	0	2	0	7	0
Appl Microbiol Biotechnol	Springer	3.81	**230**	**8**	**3.48**	78	3	91	5	61	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Breast Cancer Res Treatment	Springer	4.20	**206**	**12**	**5.83**	61	5	45	3	100	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cancer Chemotherapy and Pharmacolog	Springer	2.57	**542**	**29**	**5.35**	29	0	42	2	39	1	54	5	43	1	52	4	59	4	25	3	38	1	31	3	30	3	21	2	18	0	12	0	12	0	11	0	8	0	7	0	7	0	4	0
Molecular and Cellular Biochemistry	Springer	2.39	**800**	**43**	**5.38**	40	5	40	4	40	6	40	3	40	3	40	2	40	3	40	3	40	4	40	2	40	0	40	1	40	3	40	2	40	1	40	0	40	1	40	0	40	0	40	0
Growth Factors	Informa	3.09	**166**	**10**	**6.02**	0	0	4	0	11	0	12	1	12	1	19	1	19	2	13	2	3	1	10	1	5	1	4	0	3	0	7	0	7	0	5	0	13	0	10	0	4	0	5	0
Cancer Investigation	Informa	2.06	**220**	**13**	**5.91**	12	0	11	1	17	1	23	2	35	2	37	2	26	3	13	1	9	0	5	1	9	0	6	0	3	0	4	0	3	0	2	0	1	0	0	0	1	0	3	0
Leukemia & Lymphoma	Informa	2.61	**404**	**13**	**3.22**	54	1	26	2	39	1	26	0	29	3	17	1	20	1	23	2	16	1	9	0	12	0	12	1	20	0	23	0	15	0	10	0	20	0	9	0	11	0	13	0
International Journal of Oncology	Spandidos	2.77	**89**	**11**	**12.36**	NA	NA	51	6	38	5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Science	AAAS	31.48	**681**	**9**	**1.32**	37	0	35	1	32	0	31	0	41	0	34	1	38	1	51	1	36	1	47	0	58	1	43	1	56	1	59	1	52	0	4	0	3	0	5	0	8	0	11	0
Nature	Nature	42.35	**750**	**12**	**1.60**	50	0	50	1	50	1	50	1	50	0	50	2	50	2	50	0	50	0	25	0	25	0	25	0	25	0	25	1	25	0	25	2	25	1	25	1	25	0	50	0
Nature Oncogene	Nature	8.46	**150**	**7**	**4.67**	50	2	50	4	50	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cancer Cell	Cell Press	23.89	**188**	**6**	**3.19**	55	2	63	2	70	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Journal of Cell Biology	RU Press	9.79	**329**	**1**	**0.30**	102	0	110	0	117	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PNAS	NAS	9.81	**350**	**19**	**5.43**	50	1	50	4	50	2	50	2	50	3	50	3	50	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Sum		**20621**	**782**		**5576**	**216**	**6790**	**278**	**1350**	**54**	**814**	**36**	**801**	**37**	**599**	**29**	**617**	**29**	**435**	**18**	**419**	**23**	**414**	**18**	**403**	**11**	**337**	**12**	**342**	**4**	**357**	**6**	**336**	**3**	**267**	**3**	**254**	**3**	**164**	**1**	**151**	**1**	**194**	**0**
	% ID				**3.79**	**3.87**		**4.09**		**4.00**		**4.45**		**4.52**		**4.74**		**4.41**		**4.14**		**5.49**		**4.35**		**2.73**		**3.56**		**1.17**		**1.68**		**0.89**		**1.12**		**1.18**		**0.61**		**0.66**		**0.00**	