

1 **Analysis and Correction of Inappropriate Image Duplication:**

2 ***The Molecular and Cellular Biology Experience***

3

4 Elisabeth M. Bik<sup>1</sup>, Ferric C. Fang<sup>2</sup>, Amy L. Kullas<sup>3</sup>, Roger J. Davis<sup>4</sup> and Arturo

5 Casadevall<sup>5</sup>

6 <sup>1</sup>uBiome Inc., San Francisco, CA

7 <sup>2</sup>Former Editor-in-Chief, *Infection and Immunity*. Departments of Laboratory Medicine  
8 and Microbiology, University of Washington, Seattle, WA.

9 <sup>3</sup>Amy L. Kullas, Publishing Ethics Manager, Journals, American Society for Microbiology  
10 (ASM), Washington, DC, USA

11 <sup>4</sup>Former Editor-in-Chief, *Molecular and Cellular Biology*, Howard Hughes Medical  
12 Institute and Program in Molecular Medicine, University of Massachusetts Medical  
13 School, Worcester, MA

14 <sup>5</sup>Editor-in-Chief, *mBio*. Department of Molecular Microbiology and Immunology, Johns  
15 Hopkins School of Public Health, Baltimore, MD

16

17

18

19 **Abstract**

20 The present study analyzed 960 papers published in *Molecular and Cellular Biology* (MCB) from  
21 2009-2016 and found 59 (6.1%) to contain inappropriately duplicated images. The 59 instances  
22 of inappropriate image duplication led to 42 corrections, 5 retractions and 12 instances in which  
23 no action was taken. Our experience suggests that the majority of inappropriate image  
24 duplications result from errors during figure preparation that can be remedied by correction.  
25 Nevertheless, ~10% of papers with inappropriate image duplications in MCB were retracted. If  
26 this proportion is representative, then as many as 35,000 papers in the literature are candidates  
27 for retraction due to image duplication. The resolution of inappropriate image duplication  
28 concerns after publication required an average of 6 h of journal staff time per published paper.  
29 MCB instituted a pilot program to screen images of accepted papers prior to publication that  
30 identified 12 manuscripts (14.5% out of 83) with image concerns in two months. The screening  
31 and correction of papers before publication required an average of 30 min of staff time per  
32 problematic paper. Image screening can identify papers with problematic images prior to  
33 publication, reduces post-publication problems and requires significantly less staff time than the  
34 correction of problems after publication.

35

## 36 **Introduction**

37 Recently we reported an analysis of 20,000 papers from 40 biomedical journals, published over  
38 a period of 20 years, in which approximately 1 in 25 papers contained at least one  
39 inappropriately duplicated image (1). The frequent occurrence of inappropriate image  
40 duplication in published papers is a major concern, because it reduces the integrity and  
41 credibility of the biomedical literature. At one end of the spectrum, inappropriate image  
42 duplications caused by simple errors in constructing figures raise concerns about the attention  
43 given to the preparation and analysis of data, while at the other end of the spectrum, problems  
44 resulting from deliberate image manipulation and fabrication indicate misconduct. Increased  
45 awareness of such image duplications has resulted from post-publication peer review websites  
46 such as PubPeer and discussions on social media (2). Whereas simple errors found in  
47 published studies can be addressed by a correction, deliberate image manipulation or  
48 fabrication can lead to retraction of a paper (3).

49  
50 Inappropriate image duplications undermine the quality of the literature and can necessitate a  
51 considerable investment of time and resources by authors and journals when discovered after  
52 publication of a scientific paper. However, we presently lack information on the causes for the  
53 inappropriate image duplications, since neither cause nor intent can be reliably inferred from  
54 inspecting images in published articles. We categorized inappropriate image duplications as  
55 simple duplications (category 1), shifted duplications (category 2) or duplications with alterations  
56 (category 3), with category 1 most likely to result from honest error, while categories 2 and 3  
57 have an increased likelihood of resulting from outright falsification or fabrication. A follow-up  
58 analysis of a subset of these papers found that several variables including academic culture,  
59 peer control, cash-based publication incentives and national misconduct policies were  
60 significantly associated with duplications in categories 2 and 3, suggesting that these variables

61 might affect scientific integrity (4). In the present study, we sought to determine whether an  
62 investment by a journal to scan images in accepted manuscripts prior to publication could  
63 resolve image concerns in less time that was required to address these issues after publication.

64  
65 The mission of the journals published by the American Society for Microbiology (ASM) is to  
66 publish high-quality scientific articles that have been rigorously peer reviewed by experts and  
67 evaluated by academic editors (5). In 2013, the ASM journal *Molecular and Cellular Biology*  
68 (MCB) instituted a program to analyze the figures in all accepted manuscripts before publication  
69 (6), modeled after a similar program used by the *Journal of Cell Biology* (7, 8). In this study, we  
70 applied the approach used previously (1) to published papers in the journal MCB, and followed  
71 up the findings with a process that included contacting the authors of the papers. Consequently,  
72 we are now able to provide information as to how inappropriate image duplications occur. In  
73 addition, a set of manuscripts accepted for publication in MCB was inspected prior to publication  
74 for spliced, beautified, or duplicated images. For both sets of papers, the time and effort spent  
75 on following up on these papers was recorded. The results provide new insights into the  
76 prevalence, scope and seriousness of the problem of inappropriate image duplication in the  
77 biomedical literature.

## 78 **Methods**

79 Published papers set. Papers published in 2009-2016 in MCB were inspected visually for  
80 inappropriate image duplication. For each year, issues 1-12 (January-June) were selected, and  
81 the first 10 papers in each issue containing photographic images were screened. Thus, 120  
82 papers were inspected per publication year, resulting in a total of 960 papers screened. Since  
83 almost all MCB papers contain photographic images, no specific search term was used, but  
84 papers were only counted if they contained photographic images.

85

86 Image inspection. Published papers were scanned using the same procedure used in our prior  
87 study (1). Briefly, one person (EMB) scanned published papers by eye for image duplications in  
88 any photographic images or FACS plots. Problematic images were also inspected by two  
89 additional authors (AC and FCF). Such duplicated images fell into three general categories:  
90 simple duplications, duplications with repositioning, and duplications with alterations (1). As in  
91 the previous study (1), cuts and beautifications were not scored as problematic. EMB was not  
92 aware of the year in which MCB started increased screening (see below) for image problems  
93 while she screened journals. The image allegations were confirmed using ORI forensic  
94 software by the MCB Production Department. Decisions as to whether to pursue the allegations  
95 by contacting authors were based on this analysis. Each published paper containing suspected  
96 image duplication problems was reported to the Editor-in-Chief of MCB. The EIC then  
97 requested clarification from the corresponding author(s) regarding concerns with the figure  
98 using the category classification described above. The EIC followed up on all concerns from  
99 2010 and on potential concerns in Categories 2 & 3. Category 1 concerns were handled by  
100 ASM staff.

101  
102 Prospective screening of manuscripts before publication. Starting in January 2013, all MCB  
103 manuscripts accepted for publication were screened for image duplications and other problems,  
104 including undisclosed cuts and beautifications (which were not counted in the screen of the  
105 published papers described above). For this study, the time to inspect these figures in  
106 manuscripts accepted from January 14, 2013 to March 21, 2013 was recorded. In the case of  
107 image problems, the authors were contacted and asked to explain and/or remake the figure.  
108 Corrections and retractions followed COPE guidelines  
109 (<https://publicationethics.org/resources/guidelines>).

## 110 **Results**

111

112 Inappropriate duplications in MCB published papers. A set of 960 papers published in MCB  
113 between 2009 and 2016, including 120 randomly selected papers per year, was screened for  
114 image duplication. Of these, 59 (6.1%) papers were found to contain inappropriately duplicated  
115 images. The distribution of these showed a decline since 2013, when the screening of accepted  
116 manuscripts was introduced (Figure 1). From 2009-2012, the average percentage of image  
117 duplication was 7.08%, while after the introduction of screening accepted manuscripts in 2013,  
118 the percentage was 3.96%, a significant decrease (t test;  $p < 0.01$ ).

119

120 Investigation by ASM staff into published papers with image duplication. The 59 papers with  
121 inappropriate image duplications in MCB were investigated by contacting the corresponding  
122 authors and requesting an explanation for the apparent problem. The 59 instances of  
123 inappropriate image duplications led to 42 corrections, 5 retractions and 12 instances in which  
124 no action was taken (Table 1). The reasons for not taking action included origin from  
125 laboratories that had closed (2 papers), resolution of the issue in correspondence (4 papers),  
126 and occurrence of the event more than six years earlier (6 papers), consistent with ASM policy  
127 and Federal regulations established in 42 CFR § 93.105 for pursuing allegations of research  
128 misconduct. Of the retracted papers, one contained multiple image issues such that a correction  
129 was not an appropriate remedy, and for another retracted paper, the original and underlying  
130 data was not available, but the study was sufficiently sound to allow resubmission of a new  
131 paper for consideration, which was subsequently published.

132

133 Analysis of inappropriate image duplications. Authors who were contacted about image  
134 irregularities most frequently reported errors during assembly of the figures. The most  
135 commonly reported error was the accidental inclusion of the same blot or image twice. Other

136 commonly reported mistakes were the selection of the wrong photograph, the assembly of  
137 figure panels with mock photographs that were not properly replaced, etc.

138  
139 Time effort for published papers. For the 59 papers published with potential image duplication  
140 concerns, the ASM publication staff members recorded ~580 emails pertaining to these cases,  
141 or an average of ~10 emails per case (range 4-103). In addition, at least two phone  
142 conversations with authors took place, each approximately 1 h. The Production Editor and  
143 Assistant Production Editor handled ~800 emails in their folders regarding these corrections. In  
144 addition, for 20 papers the Editor in Chief (EIC) was involved in communications with the  
145 authors, which involved a total of 244 emails (range per paper 4-29) or an average of 12.2  
146 messages per paper. Including the EIC time would add another 61 h (~15 min x 244 emails).  
147 The exact content of these emails was not disclosed to any individuals outside of the MCB  
148 ethics panel. The breakdown of the Production Staff emails were: correspondences with staff  
149 members to keep them apprised of what had been received, discussions about wording (since  
150 each item needed individual assessment of the appropriate approach), or logistical details  
151 regarding retracted or republished papers. Correspondences with authors comprised the next  
152 largest category (less than half the amount of staff correspondence), followed by  
153 correspondences with the EIC. Correspondences with the printer was the smallest category.  
154 Hence, the problem of inappropriate image duplication after publication imposed a large time  
155 burden on the journal, with an average of 6 h of combined staff time (1400 emails estimated to  
156 take 15 min each to write and follow up per 59 papers) spent to investigate and follow-up each  
157 paper.

158  
159 Screening of manuscripts prior to publication. Analyzing the papers with inappropriately  
160 duplicated images as a function of time revealed a decline in incidence beginning in 2013, which  
161 coincided with a change in the editorial process to include pre-publication screening for image

162 problems (Figure 1). During a period of 2 months in the beginning of 2013, 83 papers were  
163 accepted with 452 images inspected. In this recording period, 12 papers (14.5%) were detected  
164 in which an image concern (duplication or undisclosed cuts) was identified. The percentage of  
165 papers flagged during pre-publication screening was higher than the frequency of duplicated  
166 images detected in published papers, because beautification or undisclosed cuts were flagged  
167 as well. Prior to this time, no manuscript was rejected by MCB because of image duplication,  
168 but starting in 2013, after the introduction of pre-publication screening, the percentage of  
169 manuscripts rejected for image problems steadily increased (Figure 2).

170  
171 Outcome of pre-publication screening of manuscripts. During the recording period in 2013, 84  
172 manuscripts were screened, and 12 manuscripts were flagged for containing duplications or  
173 other irregularities. The authors of each manuscript were contacted for follow-up by the handling  
174 editor. In 11 cases, the problem could be corrected by the submission of a new version of the  
175 figure, while in 1 instance, the authors provided the original data to show that the figure did not  
176 misrepresent the original data.

177  
178 Time effort for manuscripts. When image screening was first instituted by MCB in 2013, time  
179 records were maintained for approximately two months to ascertain the time cost of this  
180 procedure. The total time required to inspect all images in the 84 manuscripts screened during  
181 this period was 687 minutes (8.3 minutes per screened paper). The total time required for  
182 reporting and following up of ethical concerns found in 12 papers was 375 minutes of ASM staff  
183 time, not counting the time devoted by the EIC to addressing these problems. Thus, the time  
184 effort of ASM staff/editors translated into 31.3 minutes per manuscript.

## 185 **Discussion**



186 Here we report the first detailed investigation of inappropriate image duplications in biomedical  
187 research papers and a systematic process for their correction. By focusing on one journal  
188 within the ASM journals portfolio, we were able to determine the outcome of image concerns.  
189 The most reassuring outcome of our findings is that the majority of inappropriate image  
190 duplications resulted from errors during figure construction that could be easily corrected by the  
191 authors. The finding that 5.5% of MCB articles had inappropriate image duplications is a  
192 percentage consistent with prior findings involving over 40 journals (1). This confirmation is  
193 noteworthy because the approach used in the current study differs from prior work in that it  
194 focused on a single journal with a 120-paper sample for each of six publication years. Of  
195 concern is that approximately 10% of the papers containing problematic images required  
196 retractions after the adjudication process, due to apparent misconduct, an inadequate author  
197 response, or errors too numerous for an authors' correction. Other efforts to investigate causes  
198 of inappropriate image duplication for papers published at two other American Society of  
199 Microbiology journals, *Journal of Virology* and *Infection and Immunity*, including some from a  
200 prior study (1), produced retraction rates ranging from 2.9 (1 of 35) to 21% (4 of 19),  
201 respectively, which yields an average of  $10.6 \pm 8.1\%$  for the three journals.

202

203 Research misconduct has always existed, but this topic has been of increasing concern in  
204 recent years in view of several high profile scandals, a perceived reproducibility crisis and an  
205 epidemic of retracted papers, most of which are due to misconduct (9). The actual number of  
206 compromised papers in the extant literature is unknown, but our observations permit some  
207 estimates. Although extrapolation from three American Society of Microbiology journals to the  
208 general biomedical literature must be made with caution, our study allows a rough estimate of  
209 the number of seriously compromised papers in print. Based on the average percent retraction  
210 from the three journals, the 95% confidence interval ranges from 1.5-19.8%. If 3.8% of the  
211 8,778,928 biomedical publications indexed in PubMed from 2009-2016

212 (<http://dan.corlan.net/medline-trend.html>) contain a problematic image (1), and 10.6% (CI 1.5-  
213 19.8%) of that group contain images of sufficient concern to warrant retraction, then we can  
214 estimate that approximately 35,000 (CI 6,584-86,911) papers are candidates for retraction due  
215 to image duplication. These numbers are almost certainly an overestimate since not all papers  
216 in the literature have images of the type studied here. On the other hand, we only screened for  
217 visible duplications, and papers might contain additional problems in graphs, tables, or other  
218 datasets that are less easy to find, suggesting that this could also be an underestimate.  
219 Whatever the actual number, it is clear that the number of compromised papers in the literature  
220 is significant. The continued presence of compromised papers in the literature could exert  
221 pernicious effects on the progress of science by misleading investigators in their fields.  
222 Nevertheless, even the most liberal estimates the total number of papers that are candidates for  
223 retraction represent a very small percentage of the literature. Our findings are consistent with  
224 other studies reporting that a significant number of papers in the literature have problems  
225 associated with misconduct (10, 11).

226  
227 Our study also documents the potential value of increased journal vigilance for reducing  
228 inappropriate image duplications in published papers. A significant reduction in the number of  
229 inappropriate image identified in MCB papers was observed after initiation of dedicated image  
230 inspections by the journal in 2013 (6). Increased vigilance reduces problematic images by  
231 identifying and correcting errors before publication and by heightening awareness among  
232 authors to prevent such problems. However, such efforts come at considerable time and  
233 financial costs to the journal. The time invested in inspecting manuscripts pre-publication was  
234 approximately 8.3 minutes per paper, and the identification of a problematic image resulted in  
235 additional time investment in communicating with authors and deciding if a problem raised an  
236 ethical concern. Additional costs to science include the time taken by the authors to correct  
237 figures and the delays in publication. However, these costs may be significantly lower than the

238 overall cost associated with discovery of image duplication after publication, which triggers an  
239 investigation by the journal that consumes considerable time, as is evident from the average of  
240 10 emails per case, to outcomes including publication of corrections and retractions. In our  
241 analysis, we found that following up on problematic images before publication costs about 30  
242 min per problematic paper, whereas the time spent to follow up similar issues after publication,  
243 not including EIC time, was 6 h per paper, which is twelve times greater. Hence, even though  
244 the majority of inappropriate image duplications result from simple errors in assembling figures,  
245 their occurrence once identified imposes considerable costs to journals and authors, and by  
246 extension, to the scientific enterprise. Identifying image problems before publication, even  
247 though this requires additional time for journal staff, might save journals time in the end by  
248 preventing problematic images from appearing in published papers. In addition, identifying  
249 potential problems before publication protects authors' reputations and prevents the collateral  
250 damage to the reputations of all authors of a retracted paper (12).

251  
252 Peer review is a cornerstone of science (13, 14), which is primarily designed to look for  
253 fundamental errors in experimental setup and data analysis. Most peer reviewers do not have  
254 the expertise to analyze papers for scientific misconduct. Consequently, the responsibility of  
255 screening for plagiarism, falsification, fabrication, and other forms of science misconduct often  
256 lies with editors (15). Although sloppiness and misconduct have always existed in science, the  
257 problem may be becoming more acute because of advances associated with the information  
258 revolution. The ability to cut-and-paste text or images combined with availability of software to  
259 manipulate and generate photographic images gives authors powerful tools that can be  
260 misused. Our prior study noted that the problem of inappropriate image duplications was largely  
261 a 21<sup>st</sup> century phenomenon temporally associated with the proliferation of software for image  
262 construction (1). However, the information revolution has also provided tools to reduce error and  
263 abuse. Some publishers, including ASM, already perform routine screening of manuscripts

264 using plagiarism-detection software. Combined with manual curation and supervision, these  
265 tools work reasonably well (11, 16). However, identifying image duplication of the types reported  
266 here and in our prior study (1) is more challenging and dependent on individuals capable of  
267 spotting suspicious patterns. We noted that the pre-screening process at MCB is quite good at  
268 picking up spliced images but poor at finding image duplications of the type reported in this  
269 study. Hence, without routine screening by individuals who are gifted at identifying image  
270 duplications and modifications, it is likely that the type of image problems identified here will  
271 continue (1). Although detecting image problems is difficult, the recent development of  
272 improved software tools appears promising (17).

273  
274 The finding that most inappropriate image duplications result from sloppiness and error during  
275 figure construction but impose large costs to authors and journals for their correction indicates  
276 that greater efforts to prevent such errors should be instituted by research laboratories. Given  
277 that most figures are currently constructed by authors themselves, it may be possible to reduce  
278 the prevalence of image problems by asking others in the laboratory that are not directly  
279 involved with the current research to participate in figure construction or review. Prior to the  
280 availability of image editing software, figures for research papers were usually made by  
281 individuals who specialized in this activity and were not involved in data collection. We note that  
282 in our previous study we found no instances of inappropriate image duplication prior to 1997 (1).  
283 We hypothesize that prior to the availability of software that allowed authors to construct their  
284 own figures, the discussions between photographers or illustrators and authors combined with  
285 the separation of data generation from figure preparation reduced the likelihood of these types  
286 of problems.

287  
288 In addition, providing clear guidelines for the preparation of photographic images as part of a  
289 journal's instructions for authors is helpful. For example, instructions might include rules about

290 how to disclose cuts in Western blots, the requirement of each experiment to have their own  
291 control (e.g.  $\beta$ -actin or globin) protein control blots (no re-use of these blots allowed), etc.  
292 Examples of such guidelines currently exist (18). ASM maintains an ethics portal in its website  
293 with information that may be helpful to authors  
294 (<http://journals.asm.org/site/misc/ethicsportal.xhtml>)

295  
296 In summary, we confirmed our prior results by inspecting a single journal using a systematic  
297 approach and provide insights into the causes of inappropriate image duplication in research  
298 papers. The results provide both reassurance and concern regarding the state of the  
299 biomedical literature. We are reassured that the majority of duplication events result from errors  
300 that do not compromise validity of the scientific publication and are amenable to correction,  
301 notwithstanding the cost of considerable time investment on the part of the journal staff, editors  
302 and authors. However, of concern is the significant minority of papers with inappropriate image  
303 duplications that result in retractions, suggesting that the current biomedical research literature  
304 contains many such publications that warrant retraction. At the very least, our findings suggest  
305 the need for both authors and journals to redouble their efforts prevent inappropriate image  
306 duplications.

307

### 308 **Conflict of interest statement**

309 Elisabeth Bik is an employee of uBiome, but conducted this work outside of working hours.  
310 uBiome did not sponsor this research. AC, FCF and RJD are either current or former editors of  
311 the ASM journals *mBio*, *Infection and Immunity*, and *Molecular and Cellular Biology*,  
312 respectively. AK is the publishing ethics manager in the ASM journals department and  
313 participated in this research in a retrospective capacity. This paper reports on an effort by ASM  
314 journals to review the integrity of figures in a subset of published manuscripts. That effort was

315 not initially intended as a study but rather as due diligence in maintaining the integrity of the  
316 scientific record. However, since the results of this effort provided important information that  
317 could inform future efforts at improving the reliability of the literature, a decision was made to  
318 present the data in a publication. The views expressed in this article do not necessarily reflect  
319 the views of this journal or the ASM.

320

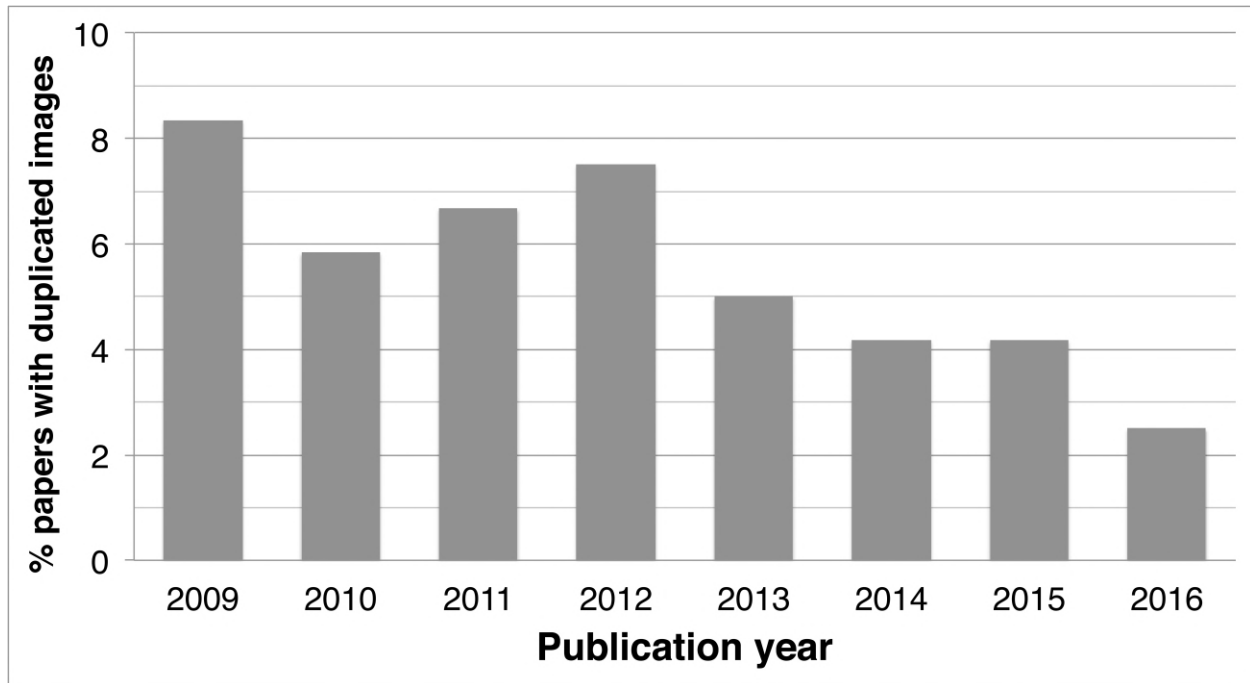
321

## 322 REFERENCES

- 323 1. Bik EM, Casadevall A, & Fang FC (2016) The Prevalence of Inappropriate Image  
324 Duplication in Biomedical Research Publications. *MBio* 7(3).
- 325 2. Knoepfler P (2015) Reviewing post-publication peer review. *Trends Genet*  
326 31(5):221-223.
- 327 3. Fang FC & Casadevall A (2011) Retracted science and the retraction index.  
328 *Infect. Immun* 79(10):3855-3859.
- 329 4. Fanelli D, Costas R, Fang FC, Casadevall A, & Bik EM (2018) Testing  
330 Hypotheses on Risk Factors for Scientific Misconduct via Matched-Control  
331 Analysis of Papers Containing Problematic Image Duplications. *Science and*  
332 *engineering ethics*.
- 333 5. Casadevall A, *et al.* (2016) ASM Journals Eliminate Impact Factor Information  
334 from Journal Websites. *MBio* 7(4).
- 335 6. Kullas AL & Davis RJ (2017) Setting the (Scientific) Record Straight: Molecular  
336 and Cellular Biology Responds to Postpublication Review. *Molecular and cellular*  
337 *biology* 37(11).
- 338 7. Rossner M (2002) Figure manipulation. *assessing what is acceptable*  
339 158(7):1151-1151.
- 340 8. Yamada KM & Hall A (2015) Reproducibility and cell biology. *The Journal of cell*  
341 *biology* 209(2):191-193.
- 342 9. Fang FC, Steen RG, & Casadevall A (2012) Misconduct accounts for the majority  
343 of retracted scientific publications. *Proc. Natl. Acad. Sci. U. S. A* 109(42):17028-  
344 17033.
- 345 10. Looi LM, Wong LX, & Koh CC (2015) Scientific misconduct encountered by  
346 APAME journals: an online survey. *The Malaysian journal of pathology*  
347 37(3):213-218.
- 348 11. Taylor DB (2017) JOURNAL CLUB: Plagiarism in Manuscripts Submitted to the  
349 AJR: Development of an Optimal Screening Algorithm and Management  
350 Pathways. *AJR. American journal of roentgenology* 208(4):712-720.
- 351 12. Bonetta L (2006) The aftermath of scientific fraud. *Cell* 124(5):873-875.
- 352 13. Glonti K, *et al.* (2017) A scoping review protocol on the roles and tasks of peer  
353 reviewers in the manuscript review process in biomedical journals. *BMJ open*  
354 7(10):e017468.
- 355 14. Kronick DA (1990) Peer review in 18th-century scientific journalism. *Jama*  
356 263(10):1321-1322.
- 357 15. Babalola O, Grant-Kels JM, & Parish LC (2012) Ethical dilemmas in journal  
358 publication. *Clinics in dermatology* 30(2):231-236.
- 359 16. Lykkesfeldt J (2016) Strategies for Using Plagiarism Software in the Screening of  
360 Incoming Journal Manuscripts: Recommendations Based on a Recent Literature  
361 Survey. *Basic & clinical pharmacology & toxicology* 119(2):161-164.

- 362 17. Butler D (2018) Researchers have finally created a tool to spot duplicated images  
363 across thousands of papers. *Nature* 555(7694):18.
- 364 18. Collins S, Gemayel R, & Chenette EJ (2017) Avoiding common pitfalls of  
365 manuscript and figure preparation. *Febs j* 284(9):1262-1266.
- 366
- 367

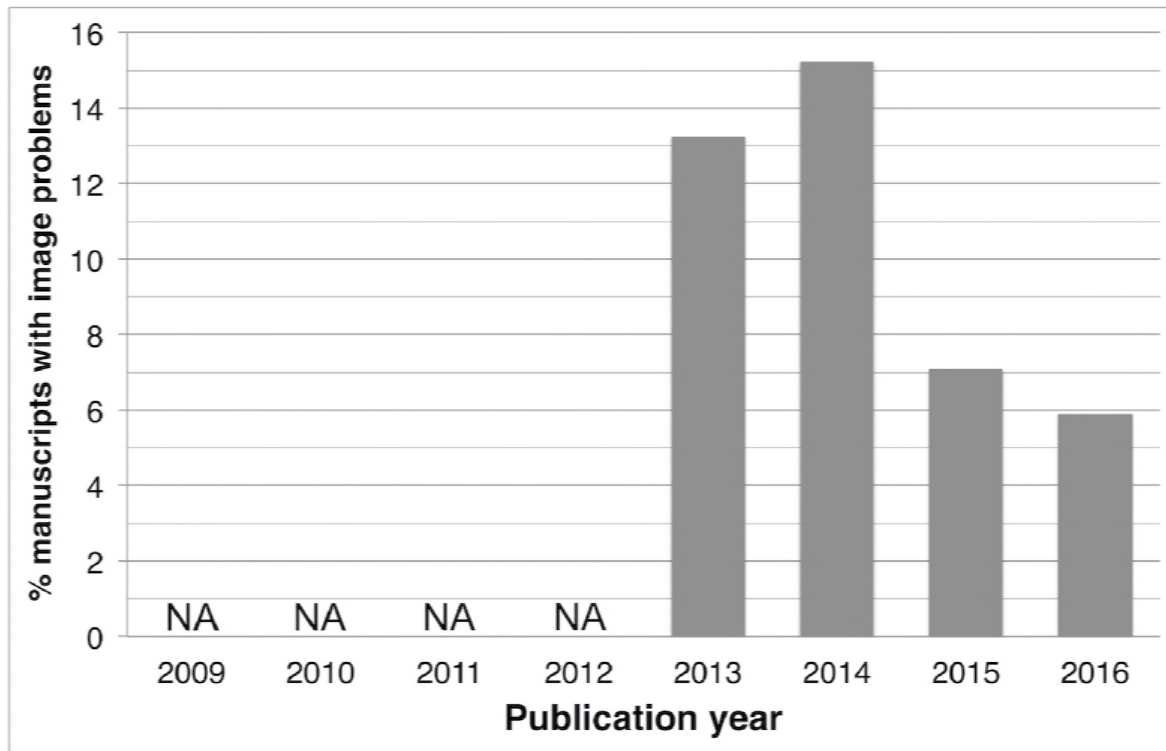




368

369 **Figure 1.** Percentage of papers published in ASM's *Molecular and Cellular Biology*  
370 containing duplicated images. Inspection of manuscripts pre-publication started in 2013.

371



372

373 **Figure 2.** Percentage of accepted MCB manuscripts that were found to have problematic  
374 images, 2013-2016. No screening for problematic images was done before 2013 and NA means  
375 non-applicable.

376

377

378 **Table 1.** Summary of results and comparison of image duplication problems in  
 379 published MCB papers and accepted MCB manuscripts.

Action	Explanation	Papers with Image duplication problems	
		Post publication (59, 6.1%) <sup>1</sup>	Pre-publication (12, 14.5%) <sup>2</sup>
None	Duplication could not be confirmed or not a strong case	3 (5.1%)	0
	Duplication was legitimate	1 (1.7%)	0 (0%)
	Lab had closed after submission of paper, not pursued	2 (3.3%)	0 (0%)
	Older than 6 years	4 (6.8%)	0 (0%)
	Authors did not reply	2 (3.3%)	0 (0%)
	Authors provided original blot showing no duplication	1 (1.7%)	0 (0%)
	Correction	Simple duplication during figure assembly	40 (67.8%)
Error in figure assembly		1 (1.7%)	1 <sup>4</sup>
Retraction (or rejection of manuscript)	Too many errors for simple correction	3 (5.1%)	0
	Intention to mislead suspected	2 (3.3%)	0
Staff effort	Emails sent to resolve problems	1400-1600 <sup>3</sup>	N.A. <sup>5</sup>
	Average time spent per paper	6 h	1.5 h

380

381 <sup>1</sup> Published papers (n=960), 2009-2016

382 <sup>2</sup> Accepted papers (n=83), 2 months in 2013

383 <sup>3</sup> Email estimate includes EIC correspondence.

384 <sup>4</sup> Analysis revealed potential case of non-uniform enhancements

385 <sup>5</sup> N.A. means not available

