

## 关于生命科学论文方法学部分的复制问题（一）

## 对 8 份全球高影响力期刊行为的调查\*

文 / 谈旭翡 张月红

**【摘要】** 本调查以生命科学领域高影响力期刊所刊发的论文在方法部分是否存在复制抄袭现象为研究目的。调查选择了该行业 8 种高影响力（或顶尖）期刊，选择该领域常用的几种经典实验方法中的“蛋白质印迹法（Western blot）”为调查项，各随机选取 10 篇文章进行等比检测分析。最后得出基本结论，即顶尖期刊较为自律，在方法部分的重复率较低，即使有也常见引用标记。同时发现一些顶尖期刊对方法部分的重复问题的解决也颇有见地。

**【关键词】** 高影响力期刊 蛋白质印迹法 CrossCheck 抄袭 自我抄袭

进入 21 世纪，生命科学在整个科学技术领域中的重要性更加凸显。对生命科学研究论文中的抄袭，尤其是方法学上的复制（或抄袭）现象，引起了学者与出版编辑们的关注与思考。众所周知，科学与技术的研究可分为基础研究、应用研究和实验发展。一般认为，实验发展在方法学方面更偏重创新，而基础与应用研究中，新手们更偏爱对业已成熟的经典方法或前实验过程进行模仿或照抄，如此这般，不可避免在论文的方法描述部分有复制抄袭的痕迹。但是方法描述中的重复或复制是否属于抄袭？目前还没有明确定论，学术界也没有更好的防范措施和解决办法。因此，本研究将该领域的 8 种顶尖期刊作为样本，通过抽样调查与分析，考察在方法描述这一选项中优秀期刊的表现行为，探究其刊发的学术论文在方法描述部分中是否存有复制或抄袭现象，以及对于这一棘手问题的解决思路，力求对该领域的出版编辑与作者有所借鉴与启发。

## 一、方法

### 1. 期刊选择

以汤姆森路透科技信息所发布的 2011 年最新 JCR 影响因子排名为依据，在医药卫生期刊和生命科学期刊中选择高影响力且开放性较大的 8 种期刊，它们是 *Cell*（《细胞》），*Journal of the American*

*Medical Association*（《美国医学协会杂志》），*the Lancet*（《柳叶刀》），*Nature*（《自然》），*Nature Biotechnology*（《自然生物技术》），*the New England Journal of Medicine*（《新英格兰医学杂志》），*PLoS Biology*（《PLoS 生物学》）和 *Science*（《科学》）。样本选取时间段截止于 2012 年 7 月 11 日。

### 2. 文献的选取与分类

从 PubMed 全文数据库的上述 8 种期刊中，选定蛋白质印迹法（Western blot）<sup>[1]</sup> 字样，通过检索搜索到全部文章。然后将每一种期刊最近发表的、可以全文下载的文章样本各选出 10 篇，作为考查样本。将检索文章中涉及蛋白质印迹法部分，分为三种表达类型：①描述型，在正文的方法学部分详细或比较详细地对蛋白质印迹法方法进行描述；②引用型，在正文的方法学部分将蛋白质印迹法通过完全引用已发表文献，或引用加简单的描述；③附件型，在正文中不出现蛋白质印迹法方法的具体描述，而是在附件中详述。

对考察样本中的描述型文献，用 CrossCheck 对其描述部分内容进行检测分析。对附件型文献，下载附件后用 CrossCheck 对其附件内容进行检测分析。

## 二、结果

### 1. 来源文献的概况

选取的 8 种高影响力期刊的正文选样中，方法

学部分含有“蛋白质印迹法”的文章共有 512 篇。其中, *Cell* (《细胞》) 中有 140 篇相关文献, 出版年份为 1983 ~ 2007; *JAMA [the journal of the American Medical Association* (《美国医学协会杂志》)) 中有 47 篇相关文献, 出版年份为 1985 ~ 2010; *Nature* (《自然》) 中有 109 篇相关文献, 出版年份为 1986 ~ 2011; *Nature Biotechnology* (《自然生物技术》) 中有 44 篇相关文献, 出版年份为 1996 ~ 2011; *PLoS Biology* (《PLoS 生物学》) 中有 33 篇相关文献, 出版年份为 2003 ~ 2012; *Science* (《科学》) 中有 36 篇相关文献, 出版年份为 1984 ~ 2009; *The Lancet* (《柳叶刀》) 中有 72 篇相关文献, 出版年份为 1985 ~ 2011; *The New England Journal of Medicine* (《新英格兰医学杂志》) 中有 31 篇相关文献, 出版年份为 1985 ~ 2011。

### 2. 考察文献的分类

在上述选取的来源文献中, 将各刊近年来发表的、可以全文下载的文章共有 80 篇, 根据三种表达类型对其分类。由图 1 可知, 描述型是文献中实验方法比较常用的表达方法。从 2002 年之后, 逐渐呈现附件型的文献表达方式, 例如《科学》杂志。

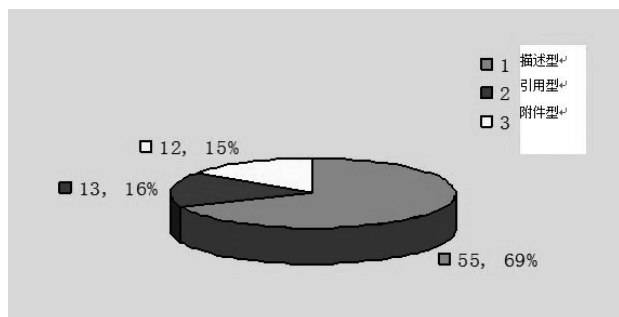


图1 80篇文本中蛋白质印迹法方法表达部分的类型分布

### 3. CrossCheck 分析结果

#### (1) 描述型中蛋白质印迹法内容

在 80 篇文献中, 描述型的有 55 篇。通过对这

表1 描述型文献的CrossCheck 分析结果

期刊名	文献数	存在段落式相似的文献数	相似段落		相似段落的情况说明
			有引用	无引用	
《细胞》	8	0			
《美国医学协会杂志》	9	1	1		相似段落有大约 60 个单词与已经发表的文献匹配 (相似的段落里没有引用该匹配文献, 但全文其他地方有引用到该文献)
《自然》	8	0			
《自然生物技术》	8	1		1	相似段落有大约 45 个单词与已经发表的文献匹配
《PLoS 生物学》	7	0			
《科学》	1	0			
《柳叶刀》	7	1		1	相似段落有大约 45 个单词与已经发表的文献匹配
《新英格兰医学杂志》	7	0			

55 篇文献内容的分析, 得到的结果见表 1。有段落式相似 (整段或者一个段落的大部分相似, 具体字数详见表 1) 的文献共 3 篇, 《柳叶刀》中有 1 篇,<sup>[2]</sup> 《自然生物技术》中有 1 篇,<sup>[3]</sup> 《美国医学协会杂志》中有 1 篇。<sup>[4]</sup> 其中, 《美国医学协会杂志》中的文献, 在方法部分相似的段落里虽然没有引用该匹配文献, 但全文其他地方有引用到该文献的标识。《自然生物技术》和《柳叶刀》上有 2 篇, 虽然全文没有引用相关文献, 但与该文中有 1 名至多名作者是相似文献的原作者。除了上述 3 例, 其他文献均没有发现段落式的相似。

#### (2) 附件型中蛋白质印迹法内容

表2 选取的文献中蛋白质印迹法方法部分不同表达类型的情况分析

期刊名	总文献数	各表达类型的文献数		
		描述型	引用型	附件型
《细胞》	10	8	2	0
《美国医学协会杂志》	10	9	1	0
《自然》	10	8	2	0
《自然生物技术》	10	8	2	0
《PLoS 生物学》	10	7	3	0
《科学》	10	1	0	9
《柳叶刀》	10	7	2	1
《新英格兰医学杂志》	10	7	1	2

从表 2 可知, 近年来, 附件型表达方式在《科学》《柳叶刀》与《新英格兰医学杂志》中均有出现, 其中尤以《科学》较为常见。以上 12 篇的附件文献均经 CrossCheck 检查, 所得的相似度区间分布图见

图2, 其中在《柳叶刀》和《新英格兰医学杂志》的2篇文献中, 附件的相似度在20%以下, 属于较正常的相似度值。而《科学》杂志中, 有3篇文献<sup>[5,6,7]</sup>的附件相似度值在30%以上(容易出现段落式相似的相似度值)。<sup>[8]</sup>我们对这3篇的相似度报告做了进一步的分析, 研究发现附件内容上都存在段落式的相似, 其中一篇有引用, 但引用的文献并不是CrossCheck检索到的文献(而CrossCheck检索到的文献却引用出现在正文的其他部分)。另外2篇, 没有引用相似文献, 但与该文献有1名至多名的相同研究者。

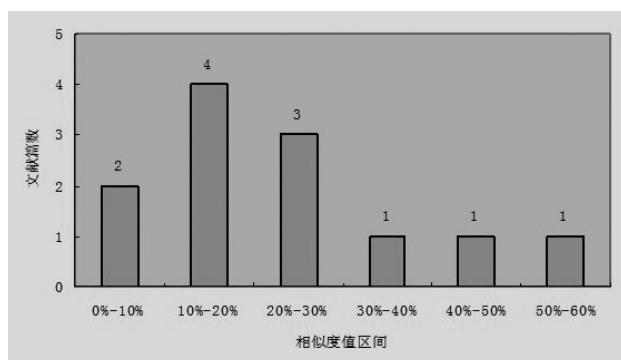


图2 附件型文献内容相似度值分布情况

### 三、讨论

针对同一蛋白质印迹法的三种表述方式, 描述型的运用较为普遍。而《科学》论文中实验方法的表述形式, 正逐渐由附件型替代描述型, 这或许是一种新的表述方式的尝试, 从而避免该学科方法部分复制抄袭的嫌疑。<sup>[9,10]</sup>

以蛋白质印迹法的方法描述为例, CrossCheck的结果显示, 极少存在抄袭现象。从仅发现的3个例子中, 1篇是相似文字较多, 但有文献引用标识; 另2篇是与相似文章有相同的研究者(或可以归于自我抄袭一类)。在附件型的方法表达中, CrossCheck的相似度比在正文中的高, 特别是《科学》杂志, 9篇入选的文献中有3篇附件的相似度值在30%以上。总之, 这些高影响力的期刊, 在研究方法的表述上, 极少存有抄袭现象。而近年来作为附件型的表述方式, 其相似度上是否可以降低要求, 或许值得探究。但不可否认的是, 附件型的表述方式, 的确减少了正文(或者说原创论文)关于某种实验方法描述的重复现象, 而附件本身在行业里并不作为原创文的一部分, 因此避免了重复发表之嫌疑。

(作者单位: 浙江大学学报英文版编辑部)

\* 本文系国际出版伦理委员会(COPE)研究基金项目(2010-12)和国家自然科学基金“重点学术期刊”资助项目(No.30824802)部分成果

#### 注释:

- [1] Burnette WN. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.*, 1981, 112(2):195-203
- [2] Phillips MJ, Azuma T, Meredith SL, Squire JA, Ackerley CA, Pluthero FG, Roberts EA, Superina RA, Levy GA, Marsden PA. Abnormalities in villin gene expression and canalicular microvillus structure in progressive cholestatic liver disease of childhood. *Lancet*, 2003, 362(9390):1112-1119. [doi:10.1016/S0140-6736(03)14467-4], 2003,
- [3] Subramanyam D, Lamouille S, Judson RL, Liu JY, Bucay N, Derynck R, Belloch R. Multiple targets of mir-302 and mir-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat. Biotechnol.*, 2011, 29(5):443-448. [doi:10.1038/nbt.1862]
- [4] Quinn TC, Kline RL, Halsey N, Hutton N, Ruff A, Butz A, Boulos R, Modlin JF. Early diagnosis of perinatal hiv infection by detection of viral-specific iga antibodies. *JAMA*, 1991, 266(24):3439-3442.
- [5] Legname G, Baskakov IV, Nguyen HO, Riesner D, Cohen FE, Dearmond SJ, Prusiner SB. Synthetic mammalian prions. *Science*, 2004, 305(5684):673-676. [doi:10.1126/science.1100195]
- [6] Launay P, Cheng H, Srivatsan S, Penner R, Fleig A, Kinet JP, Trpm4 regulates calcium oscillations after t cell activation. *Science*, 2004, 306(5700):1374-1377. [doi:10.1126/science.1098845]
- [7] Van Sickle MD, Duncan M, Kingsley PJ, Mouihate A, Urbani P, Mackie K, Stella N, Makriyannis A, Piomelli D, Davison JS, et al. Identification and functional characterization of brainstem cannabinoid cb2 receptors. *Science*, 2005, 310(5746):329-332. [doi:10.1126/science.1115740]
- [8] Kirsty Meddings. CrossCheck Plagiarism Screening: Understanding the Similarity Score. Available from <http://www.ithenticate.com/plagiarism-checker-blog/bid/63534/CrossCheck-Plagiarism-Screening-Understanding-the-Similarity-Score>, 2011
- [9] Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC. Sir2-dependent activation of acetyl-coa synthetase by deacetylation of active lysine. *Science*, 2002, 298(5602):2390-2392. [doi:10.1126/science.1077650]
- [10] Gruber N, Keeling CD, Bates NR. Interannual variability in the North Atlantic ocean carbon sink. *Science*, 2002, 298(5602):2374-2378. [doi:10.1126/science.1077077]