



很多被广泛研究的细胞系继续被Korch背后的海拉细胞淹没。  
图片来源:JOANNA B PINNEO

## “李鬼”细胞荼毒科研 科学家指出被污染细胞系使生物学遭受严重损失

“我称自己为校正器。”美国科罗拉多大学遗传学家 Christopher Korch 说。他急切想校正的是已困扰生物医学研究半个多世纪的问题——实验室细胞培养遭到污染。在过去的15年间,Korch 宣称 78 个被广泛使用的细胞系最终证实长满了其他细胞。甲状腺细胞系实际上由黑色素瘤细胞组成,前列腺组织被膀胱癌取代,而正常的子宫培养细胞被证明只是乳腺癌。这些都使人对不计其数的基础生物学和疾病研究产生了疑问。

在一年紧锣密鼓的数据收集和分析后,Korch 相信自己第一次开始量化被污染细胞系对科学事业造成的损失。“我们正在检查上万篇已发表的论文、几百万个期刊引文以及潜在的数十亿美元研究经费。”

一些看过 Korch 所撰白皮书初稿的科学家被其打动,开始推动变革。“对于 Korch 博士的分析,令我印象深刻的是(实验室细胞培养被污染)问题比我预测的更加普遍。”华盛顿大学微生物学家 Ferric Fang 表示。

Fang 是《感染与免疫》杂志的编辑。他说,自己将把 Korch 的成果呈现给发表 13 种期刊的美国微生物学会的领导层。不过,Fang 提醒道:“并非所有使用被错误识别的细胞系的研究都一定意味着努力的浪费。”一些研究可能仍会获得有用信息。

而 Korch 说,他的目标是“让大家明白该问题造成的巨大影响,引起科学家、资助机构、期刊和高校的注意”。

### 约有 20% 细胞系已被污染

Korch 在瑞典、法国、挪威和美国开展了几十年稳定但平凡的遗传学研究后,于 1995 年来到科罗拉多州奥罗拉市。很快,Korch 便被要求管理科罗拉多大学的基因测序设备,并在分析细胞系以决定其真实的基因身份方面成为全球领先的专家之一。

现年 70 岁的 Korch 利用一种被称为短串联重复分析的标准方法研究特定基因序列。该技术一直被美国联邦调查局和其他执法机构长

期使用,旨在从犯罪现场留下的血液和其他组织中鉴别基因。同时,它还被用于区分来自不同个体的细胞系。

“Christopher 绝对是《犯罪现场调查》中的科学英雄。”科罗拉多大学癌症中心的创立者 Paul Bunn 说。作为测序中心主任,Bunn 表示:“Christopher 在追求自己着迷的事业方面是完全自由的,没有任何阻拦。”

细胞系被认为保留了原始组织的特性,并且作为一种生理学的活试管,用于研究基础生物学和生物医学领域最重要的问题,比如癌症和正常组织是如何对药物作出反应的。不过,细胞系的永生性既是它的巨大优势,也是其最显著的弱点。这种特性能使实验一次又一次被重复以确认某项发现。然而,随着这些相同的细胞系在几十年里通过不同的人 and 实验室传递,它们会被其他细胞污染。闯入者的生长速度会超过原始细胞,并最终完全替代正宗的实验室培养细胞。基于自己和其他调查者的研究,Korch 估计约有 20% 的细胞系已被污染。

研究人员大多无视或否认这个问题。上世纪 70 年代,生物学家 Walter Nelson-Rees 竭尽全力“曝光”了冒充的细胞系,并且力推开展正常的鉴定。然而,他的努力遭到者多中伤,最终迫使他完全告别科学界。之后,细胞生物学家 Roland Nardone 担当起斗争的重任,强烈谴责期刊和资助机构没有要求检验细胞系以验证其身份的做法。不过,他也几乎没有取得任何进展。

### 量化损失

如今,Korch 加入了这场争论。2012 年,他参加了一个由 20 位科学家组成、旨在清理细胞系相关文献的全球志愿者组织——国际细胞系鉴定委员会(ICLAC)。ICLAC 由澳大利亚细胞生物学家 Amanda Capes-Davis 担任主席,其成员通过电话会议碰面,并且发表关于细胞培养实践的文章,以将污染风险最小化。同时,他们在 PubMed Commons 网站为冒充的细胞系添加注释。这是一个允许科学家在任何一条被美国国立医学图书馆数据库(PubMed)收录的摘要下方添

加评论的在线工具。ICLAC 成员还试图为细胞系“验明正身”的期刊和科学家提供帮助。

自 2013 年起,Korch 还开始量化来自被错误识别的细胞系的损失。他的第一个案例研究是 HEp-2 和 INT 407。前者被认为源于 1955 年的一个喉鳞状细胞癌样本,后者则在 1957 年从两个月胚胎的空肠和回肠培养而来。1967 年,遗传学家 Stanley Gartler 发现这两种细胞系都是生物学领域研究最多的肿瘤细胞系——海拉细胞。当它们第一次被培养出来时,就已经污染并且替代了原始细胞。不过,HEp-2 一直被广泛用于研究喉癌,而 INT 407 被普遍用作正常肠道细胞的模型。

在将实验室时间缩短到一周两天后,明年大多数时候 Korch 将在家中度过,收集并量化堆积如山的数据,以估量上述两个被污染细胞系在科学文献中留下的“足迹”。他发现,每个细胞系都拥有多个别名。一旦掌握了它们所有的迭代,Korch 便在 PubMed、谷歌学术以及很多来自诸如美国微生物学会、爱思唯尔等出版商的期刊数据库中寻找。他发现,HEp-2 打着喉癌的伪装旗号,已在 5700 多篇已发表的论文中使用。INT 407 则以正常肠道细胞的欺骗身份被用于 1300 多篇已发表的论文。

不过,Fang 争论说,在一些情况下身份被误认并不会对研究成果造成损害。“第一篇描述沙门氏菌如何侵入宿主细胞的论文就是利用 INT 407 细胞产生的。”Fang 表示,INT 407 被认为是细菌所攻击的肠道细胞的理想模型。“这是沙门氏菌领域最重要的论文之一,已被引用了 600 多次。”尽管它们属于海拉细胞,而不是肠道细胞,Fang 看不出任何质疑上述论文结果的理由。

### 唤醒科学界

在其他案例中,细胞的身份至关重要。去年 10 月,《当代微生物学》杂志发表了一项使另一种食源性病原体——李斯特菌入侵细胞的基因研究。该研究比较了李斯特菌如何入侵 3 种明显不同的细胞系:INT 407、HEp-2 和海拉。事实上,该研究对单一的海拉细胞系——海拉

检测了 3 次。

“我们使用了 3 种不同的细胞系,希望找到在很多组织中都能发挥作用的通用机制。”来自波兰华沙大学的该项研究参与人员 Radoslaw Stachowiak 说,“我已经意识到一些细胞系存在问题,尤其是 INT 407,但并不清楚问题的严重性。”他继续重复很多科学家曾经历过的承认某个细胞系是虚假的困难时期。“我不得不承认,考虑到不同的外观甚至是生长条件,很难接受所有这些细胞系都是海拉的事实。因为结果显示我们只是在海拉细胞系中进行了检测,我们并不清楚这种机制是否仅限于海拉。”Stachowiak 已向《当代微生物学》取得联系,要求在论文中添加更正。

不过,Korch 发现,即使在科学家知道自己喜爱的细胞系被污染后,他们可能仍然继续研究它。一个典型的例子是被用作血管细胞理想模型的 ECV304,因为其被认为源自脐静脉样品。事实上,由来自德国生物样本库 DSMZ 的遗传学家 Wilhelm Dirks 领导的团队在 1999 年发现,ECV304 是膀胱癌。从那以后,已有 1000 余篇关于 ECV 的论文被发表。2011 年,两位研究人员在一封写给《生物化学杂志》的信中解释了为何他们坚持使用该细胞系。来自佐治亚医学院的 Wen-Cheng Xiong 及其同事承认对 ECV 的辨认出现了错误,但同时表示,膀胱癌细胞具有诸如血管等内皮组织的特性。为此,他们得出如下结论:在缺少理想模型的情况下,我们相信 ECV304 细胞在内皮功能的研究中仍然有用。

Korch 关于成本的最新研究能否将科学界从半个世纪的惰性中唤醒?“美国国立卫生研究院(NIH)的科学家都赞成这是一个严重的问题,而且真正解决该问题的时候到了。”评审过 Korch 分析论文的 NIH 下属国立综合医学研究所所长 Jon Lorsch 表示。

最终,所有这些努力都让 Korch 保持着乐观态度。“事实上,我看见‘水闸’开始打开。”他说,“每个地方的科学家都开始对再理性提出要求。我希望自己的工作是在正确方向上的一个额外动力。我们所有人都希望文献的金字塔能牢固地建立在坚实的基础上。”(宗华)

## 科学线人

全球科技政策新闻与解析

## 编辑辞职 抗议同行评议也能“加急”



图片来源:《科学》

自然出版集团(NPG)旗下开放获取期刊《科学报告》的一位编辑近日在推特上表示,自己已经辞职,以抗议 NPG 允许作者为提交的论文支付加急同行评议费用的决定。

“我反对建立双重制度,不再选择最好的科学成果及时发表,这会使得政策向资金充足的实验室和团队倾斜。”英国伦敦大学学院生物地理学家 Mark Maslin 说,“学术出版业正在经历革命,过程将会坎坷。但这种做法是我无法接受的。”

期刊的同行评议通常是一个缓慢的过程。毕竟,这是个匿名、自愿的工作,提供评议的科学家除了期刊编辑的感谢和为科学界作出贡献的成就感外没有其他报酬。但一些期刊标出价格,作者能够通过选择经由快速通道提交论文,以加快同行评议过程。

近日,NPG 宣布正在试验名为 Rubriq 的同行评议服务,该服务由美国达拉谟的一家公司提供。如果支付给 NPG 750 美元费用,作者就能在 3 周内获得评议或退回这笔费用。NPG 拒绝公布从该公司获得多少资金。

这家公司如何如此快速地进行同行评议?“我们拥有约 100 名博士雇员。”该公司 CEO Shashi Mudunuri 说。这支小型编辑团队从全世界招募科学家充当评议者。审稿人完成一篇全文评价将获得 100 美元。

该评议本身也是流程化的,使用了一种在线“记分卡”,而非传统的评论、提问和建议的方式。该公司还为作者提供直接服务,帮助他们润色论文和寻找发表刊物。

在 Maslin 的推特发布之后,NPG 迅速为使用 Rubriq 进行了辩护。该出版集团表示:“这只是一个小型范围的试点。”NPG 出版主任 Nandita Quaderi 在一篇博客中写道:“不用说,一个作者选择了快速通道,而由此获得的好处只是得到快速结果。这个服务的引入对编辑决策过程没有任何影响。”

“值得注意的是,其他一些出版社同样提供快速同行评议服务。”一位 NPG 的代表在回复《科学》杂志的邮件中说:“我们已经非常严格地评估了 Rubriq 提供的服务质量,并确信它们发布的同行评议报告比得上 NPG 的标准。”但 Maslin 表示,“深思熟虑与仔细评估无疑比速度更重要。”(张章)

## 英政府科学家反对媒体限制



英政府希望政府机构科学家在接受采访前获得许可。  
图片来源:气象办公室

近日,英国的科学传播拥护者对《公共服务守则》的改变表达了“深切关心”。该守则主要针对包括政府机构研究人员在内的公职人员。3月16日新增添的内容要求所有的媒体接触活动都必须预先得到相关主管部门的批准。

伦敦科学媒体中心(SMC)及其他两个机构担忧相关政策会让科学家失去及时回应记者的采访,从而阻碍科学传播。“他们已经非常安静了。”SMC 主任 Fiona Fox 说,“如果政策改变让他们更安静,这是件坏事。”加拿大也没有类似的媒体联络限制条件,而这些限制延迟了对科学家的采访,使得加拿大研究逐渐淡出了媒体视野。

SMC 要求英国政府重新考虑政策修订,至少豁免科学家。英国科学作家协会和科学公共关系工作者网络 Stempra 也联合签署了一封信件,并将它提交给负责相关政策修改的内阁大臣 Francis Maude。他们在信中写道:“我们相信这将对公众了解科学产生负面影响,也会影响针对重大和有争议议题的公共话语的质量。”

为了回应这些机构的担忧,Maude 表示,这些变化只是简单地为现存媒体联系“提供透明度”,并确保这不会削弱对举报者的保护。但这一回答并没有说服所有人。一个名为 FDA 的由 1.8 万名公务员组成的联盟不买账。

FDA 总干事 Dave Penman 在一份声明中指出:“无论豁免是否适合揭发内幕,不适当和没有必要的改变使得《公共服务守则》让公务员在处理或回应媒体采访时考虑再三。”FDA 发表请愿书希望改变该政策。(张章)

## 加密“云”为个性化医疗铺平道路

极致方法使安全访问敏感信息成为可能

未来医疗的梦想是掌握基因与疾病之间的关系,并据此制定个性化的医疗方案。但是科学家认识到这种“个性化”或“精准”医学存在一个问题:即如何在保证基因数据和医疗记录安全的同时,用于基于云端的大规模数据分析处理这些数据。如今,一种方兴未艾的数据加密技术表明,这一难题或已得到解决。

在美国加州大学圣迭戈分校(UCSD)近日举行的一次研讨会上,密码学专家分析了用于测试的基因数据。他们利用较小的数据组以及一种叫作“同态加密”的方法,在 10 分钟左右就可以找到与疾病相关联的基因变异。而在现实中,计算机经常会在数小时内陷入停顿,例如有在相当于全基因组的数十万分之一的基因片段中找出与某种疾病相关联的基因变异。“这是一项很有前途的研究成果。”UCSD 计算机学家、帮助组织此次研讨会的小肖倩 Jiang 说,“但在扩展其用途方面依然存在许多挑战。”

医生和研究人员认为,为了了解基因如何影响疾病,就需要获得采集自数千个人的基因和健康数据。他们已经开始计划其中的一些项目,如美国总统奥巴马的“精准医学计划”以及英国的“10 万人基因工程”等。类似的研究项目需要具备云计算网络处理能力,但近年来的网络安全问题却表明,把如此海量、敏感的信息放在网络空间上存在极大风险。例如美国国立卫生研究院的基因型和表型数据库(dbGaP)中的基因和医学数据便因为安全问题,而禁止使用这些数据把数据储存在联网的计算机上。

而同态加密则让这一问题得到解决,其方



基因测序数据云处理有望加速发现与疾病相关联的基因突变。  
图片来源:David Paul Morris

法是通过让科学家仅在云端存储在数学上打乱的(或称作为加密)形式的数据。这种方法包括在一台本地电脑上对数据进行加密,然后把打乱的数据上传到云端。已加密数据可以在云端进行处理,然后生成的加密结果被返回至本地计算机,最终由本地计算机解码答案。如果有数据窃取者想在此过程的任何一点上截取加密数据,那么其中的所有数据都将保持安全。

“如果我们能证明,这些技术确实可以发挥作用,那么它就可以为通过计算机运算及储存这些需要保护个人隐私的高通量数据提供更安全的保证。”UCSD 计算机学家及此次会议组织者之一的 Lucila Ohno-Machado 说。

同态数据加密首次在 1978 年提出,它与其他加密方法的不同之处在于:它可以交由云端

操作打乱这些数据,而从本质上说,云端永远不会“看到”它所运作的数据库。而且,和其他加密方法不同,当它进行加密数据运算时,其运算出的结果和正常情况没有出入。

然而,在 2009 年纽约州 IBM 托马斯·沃森研究中心的加密专家 Craig Gentry 证明这种方法可以对同态加密数据执行任何一种运算之前,该方法一直仅停留在理论概念层面。Gentry 通过把每个数据点转化成一段加密信息,或称密码文本,而加密后的信息比初始数据更加庞大、复杂。一个未加密的数据可能会被加密成拥有数百万字节的密码文本——其大小相当于一张数字化照片的尺寸。这种方法是一个进步,但由于需要对其解密,相关运算却可能达到 14 个数量级。因此 Gentry 的方法就其特点来说

解决上述难题提供了一种可能性,但其运算方面的复杂程度却让其难以走向现实应用。

自那时起,密码学专家就在开发新的系统解决这些问题,例如通过对许多数据片段一起加密,这样可以使它们同时得到处理;或是通过把真实数据直接加密为单个加密文本,而不是像起初那样把它们加密为多个加密文本。这些创新比原来对每个片段的数据分别进行加密及执行运算所花费的时间和内存都更少。来自 IBM 研究中心的 Shai Halevi 说,现在的同态加密运算速度比 2009 年的运算速度快 15 万倍。“同样的一项运算,2012 年可能要花费一天半,但现在我们 5 分钟就能搞定。”他说,“现在是时候提出,这样的速度是否足以达到应用阶段了。”

尽管如此,依然有人对这种方法持怀疑态度。马里兰州贝塞斯达国家生物技术信息中心收藏部主管 Steven Sherry 表示,即便同态加密可以起作用,也不足以保护研究人员电脑中的数据。但是一些密码学专家已经开始在生物医药领域配置有限的同态密码系统。例如,瑞士一些艾滋病病毒研究管理人员和一家生物医学库就在利用这种方法:他们仅对一小部分相对简单的运算采取这种方法,因此并不介意当前的系统并不适宜快速地进行复杂运算。

尽管如此,也有人担心这些基因数据存在被泄露的潜在风险。设计该系统的瑞士联邦理工大学 Jean-Pierre Hubaux 说:“科学家确实担心不久的将来可能会出现一些丑闻,但他们却因为没有事先采取恰当的防范措施遭受批评。”(红枫)