

# 基因测序“黑科技”给生命来个完整的“数字化解读”

本报记者 张佳星

“测序技术抓住了生命科学技术的龙头,基因组测序成本从最初的3000万美元下降到1000美元甚至更低,使得每个人都能为自己进行基因组测序,这将加速生命数字化的进程。”6月21日,2019测序技术和应用高峰论坛上,中国科学院院士杨焕明表示,在高通量测序技术自主化的基础上,不断有新的“黑科技”跟进,这使得人们将获得更加准确的基因序列,作为生命数字化的“蓝本”。

## 吐口唾液就能测序,但基因组序列不完整

吐口唾液邮寄,一个人的基因组序列就会得到分析。

“目前从全世界范围来讲,大概有接近2000万人有自己的基因信息。”华大智造首席运营官蒋慧表示,随着这些年测序技术的普及,测序的门槛越来越低,人们越来越容易对自己的基因进行部分测序。

但并不是所有的测序都是获得完整的基因组序列。蒋慧介绍,大部分此类的测序是对特定区域的测序,而不是整个基因组序列的测序。

如果将一个人的基因组比喻成一座奇峰峻岭的高山,可以理解为,有些测序只是拍摄了“一线天”“巨石阵”“黑龙潭”这些有代表性的景点;更完整的是,有些测序是把所有的景点拍下来,被称为“功能基因组”;而完整的基因组测序才是把高山全部拍下来,包括像荒草一样丛生的“沙漠基因”或者“无功能基因”等。

## 从头组装,完成整个基因链条的拼接

数字化生命实际上需要一个高清晰、完整的基因组作为“蓝本”。这意味着需要高质量地从头组装基因组。

“目前最大的难点是获得尽可能长的序列进行组装。”蒋慧说。如果单次测序读长变长,那么获得从头组装基因组的难度就会降低。

人类基因组长度为30亿个碱基对,而现在单次测序的读长仅为50-70kb(5万到7万个碱基对的长度),这犹如拼一副45000块拼图,其中却有很多重复相似的“小块”,很难拼接完整。不仅需要以参考基因组做对比,还要反复测序。“目前的行业标准是需要30倍的重复测序,以完成整个基因链条的拼接。”蒋慧说。

如果将短片段加上标签就不一样了。玩过拼图的人都知道,在成千块的拼图背面会有不同的区块标记,提示这些块在同一区域。

“以专有DNA分子共标签技术为基础的sLFR(单管长片段)技术就是基于这样的原理。”蒋慧说,华大智造的这一自主技术通过将

无疑,要实现生命数字化,需要的是把高山全部拍下来的完整的基因组测序。

但目前的测序,即便是基因组测序,仍旧不能满足生命数字化的需要。“目前无论是数据库构建,还是群体研究、疾病研究,都是采用重测序的方式,是通过与参考基因组的比对获得的,而不是从头组装。”蒋慧说,比如在炎黄基因组研究项目中,科学家找到了大概有4-5兆的区域是个体所特有的,进行测序时就是对个体区域比对,形成测序数据,这是受测序效率制约的。

可以理解为,每一座高山的形状、代表性景点都相似,就先找了最具代表性的山做个详细的“数字化解读”,描述其他山的时候,就用差异来体现。

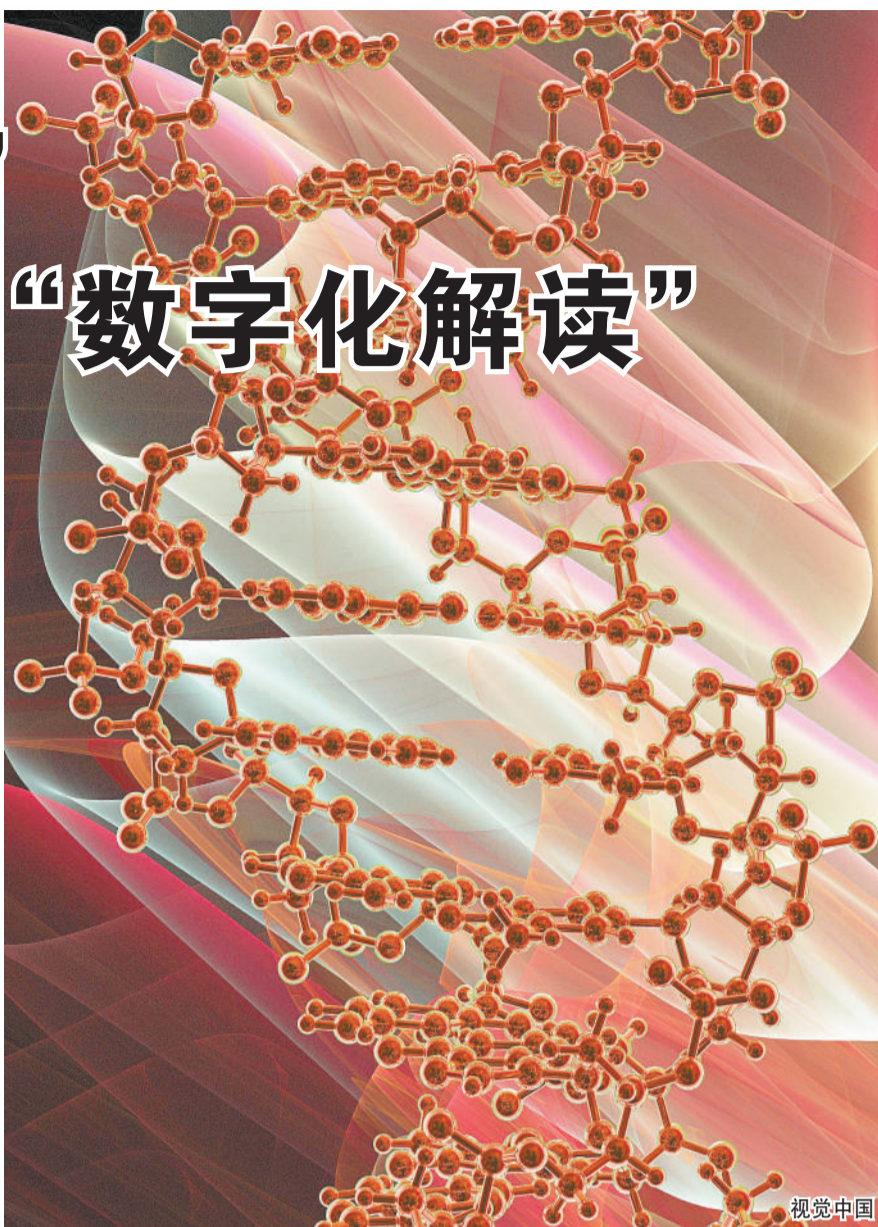
“或许每个人拥有属于自己的特点序列,但目前的测序方式难以发现这些独特部分。”蒋慧说,与参考基因组比较获得的基因组测序“密码”不会对这些未知谜题给出答案。

来源于同一DNA长片段的短读长测序片段标记上相同分子标签,能够基于高精度短读长测序获取长片段的DNA信息。

蒋慧说,读长是为了确定他山之石的位置,如果标签可以解决定位问题,将弥补短读长的弱点。基于此,从头组装两条染色体,获取二倍体测序数据也成为可能。

与此同时,华大智造将执行能得到从头测序基因组的“676”标准。基于这些数据,可以检测所有类型的结构变异,且无需与参考基因组进行比对,将大大提高基因组数据的参考性,最终帮助个体进行复杂疾病的诊断和预测。

蒋慧介绍:“我们首先会用新的技术,新的标准在全球完成1000个人的基因组测序,希望可以建出来一个模板让大家先试用一下,以在不同人群中创建高质量的参考基因组和更为完整的人类基因组多样性数据库,构建一个全新的数据集。这些数据将免费向研究人员开放。”



视觉中国

## 解读生命,别忘了体内的微生物

2018年,《自然》发表了一篇“为了生孩子把猫狗扔掉”的报道,成为科普文章的爆款。文中显示,加拿大科学家研究了770个婴儿的肠道菌群,发现家中有宠物的婴儿,其肠道菌群多样性明显高于家中没有宠物的婴儿;瑞典调研100多万儿童发现,与狗狗生活在一起的小孩子得哮喘病的风险能降低13%。

人体是一个庞大的共生体。人体皮肤表面、口腔、呼吸道、肠道生存着大量微生物,它们的数量是人体本身细胞的数十倍,编码的基因是人体基因的100倍。每个人的经历会以微生物的痕迹留在身体里,而人体的健康会与体内的菌群休戚与共。人们将特定环境中包括微生物在内的总DNA称为宏基因组。

数字化生命中,人体与微生物的作用和

关联如果得到客观地反映,甚至可能反映生命的进程。例如,老年人肠道内变形菌门、黏胶球形菌门、拟杆菌属等含量会降低。”为此,我们在与瑞典卡罗林斯卡医学院的微生物转化医学研究中心做了一个合作,启动了万人微生物研究项目,将对一万人的宏基因组进行测序。”蒋慧说,“微生物是人体内另外一套基因组,除开自己的基因组之外,我们的身体就像一个一个小社会,大家平时和环境做一些接触、吃的各种东西,人体内的病原或者微生物都会反映出来。”

依据计划,万人宏基因组测序将在3个月内完成。卡罗林斯卡医学院之前已经收集了大量的样本,华大智造将负责完成数据产出,以及产出后的数据分析和疾病关联的工作。

## DNA是本源,但还需其他参数

未来,数字化生命还有很多基础性工作要做,比如要了解基因和表型是什么关系,构建模型以研究生活习惯、疾病与特定基因表达的相关关系……

但归根结底,遗传性的DNA信息是不会再改变的,它不像其他一些因素,比如人体内的一些微生物会随着环境、饮食发生一些变化。最源头的遗传,来自于父亲和母亲的基因组不会再变化,是一个本源核心的东西。”华大集团有一个非常宏大的计划叫做

8B(80亿人的高清基因组测序),我们希望在未来通过努力让每个人都有自己的基因组数据。”蒋慧说。

“这些需要一步一步去实现,DNA是本源,是核心,但我们还会有很多其他的数据,包括蛋白、代谢等。”蒋慧说,除了基因组数据,还可以加入很多现有的医学资料,例如体检时的血液常规性筛查,可穿戴设备上的数据等,把这些所有的数据收集起来,逐步完成生命数字化的过程。

# 脂肪细胞膜蛋白:治疗前列腺癌新靶点

## 第二看台

通讯员 施为 本报记者 张晔

前列腺癌是一种泌尿系统恶性肿瘤,对中老年男性造成了巨大的危害。目前,国际公认的研究表明,通过抑制雄性激素的合成,对于治疗前列腺癌已经取得了较好效果。但是,相当一部分患者会逐渐产生抗药性并最终导致癌细胞转移。

对于产生药物抗体的患者来说,是否意味着生存的希望就此破灭呢?答案是否定的。

近日,位于苏州高新区的中国科学院苏州生物医学工程技术研究所(以下简称中科院苏州医工所)传来消息,该所肿瘤标志物中心高山实验室在国际权威期刊《癌症研究》发表研究论文,首次揭示了脂肪细胞相关膜蛋白(APMAP)可作为前列腺癌诊断和治疗的标志物,并阐述了其背后的分子机制。该研究为前列腺癌患者的治疗提供了新靶点,针对该靶点研发出的新药物,有望进一步降低前列腺癌的发展以及侵袭转移。

### 高胆固醇为何促发前列腺癌

通常,前列腺癌发展较为缓慢而易于治疗。

去势治疗是治疗前列腺癌的标准治疗方式之一,专家介绍,去势治疗目的是降低或阻断激素对肿瘤的作用,抑制肿瘤的生长。目前,国际上公认的研究表明,胆固醇对前列腺癌的发展起到很重要的作用。胆固醇促进了前列腺癌的高度进展,而抑制胆固醇合成的他汀类药物则降低了前列腺癌患者的死亡率。

然而,由于去势治疗后一直缺乏相应的治疗靶点,导致大部分的去势抵抗患者最终转变为转移性去势抵抗,而转移正是导致前列腺癌患者死亡的主要原因。因此,前列腺癌的侵袭转移成为诊断和治疗的重点研究对象,而独立于雄激素信号通路和去势抵抗的新靶点逐步成为前列腺癌领域的研究热点。

传统的肿瘤治疗方式主要通过手术加放化疗来实现。如今,随着科技进步,靶向治疗成为治疗肿瘤的新手段。即对肿瘤中能够促进该肿瘤生长和侵袭转移的靶点(标志物,通常为蛋白等),用药物和技术手段精准地抑制蛋白的表达,以达到阻止肿瘤进一步发展的目的。

“我们从高胆固醇促发前列腺癌这一现象入手,通过使用胆固醇刺激体外培养的前列腺癌细胞,将癌细胞打碎,提取细胞膜上的亚细胞组分进行检测,发现了两类靶向蛋白,分别是表皮生长因子受体(EGFR)以及脂肪细胞相关膜蛋白,高胆固醇能够诱导这两种蛋白的表达,从而促进前列腺癌的发展。”中科院苏州医工所副研究员蒋思远告诉科技日报记者。

### 或为已有抗药性的癌症患者找到新疗法

已有研究发现,前列腺癌与肺癌案例一样,癌细胞中的表皮生长因子受体的突变,大部分都是高表达,靶向治疗这个蛋白,就可以控制肿瘤发展进程。蒋思远与相关研究人员通过高通量转录组测序与生物信息学分析发现,高胆固醇诱导前列腺癌转化是通过上调脂肪细胞膜相关蛋白的细胞,进而激活表皮生长因子受体的通路实现的。

近年来,有论文报道,脂肪细胞相关膜蛋白和结肠癌转移有关,也有论文报道其在前列腺癌患者血液中表达量显著上升。中科院苏州医工所的研究人员通过进一步研究发现,在动物实验中,喂养高胆固醇的小鼠出现了明显的前列腺癌肝转移,而减少脂肪细胞相关膜蛋白数量,喂养同等含量高胆固醇的小鼠则表现出相对较低的前列腺癌肝转移。这些结果表明,脂肪细胞相关膜蛋白是一个抑制癌症细胞进一步发展的靶点,且独立于与胆固醇相关的雄性激素的信号通路,为单纯通过抑制胆固醇治疗前列腺癌而产生药物抗体的患者,找到了一条新的治疗方向。

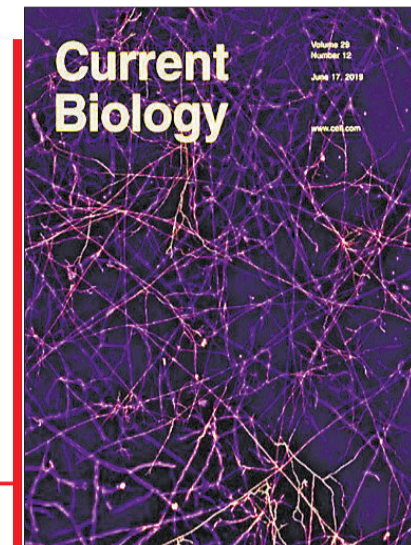
中科院苏州医工所研究员高山介绍,他们通过复杂的动物和一系列细胞实验,最终发现了前列腺癌发展背后的相关分子机制,脂肪细胞膜相关蛋白细胞是一个全新的靶点,并将在此基础上深入研究,以期揭开更多的癌症发展之谜,为抗癌新药研发探索新的路径。

## 封面故事

主持人:本报记者 陆成宽

### 真菌用营养物质「扶贫」应对资源不平等

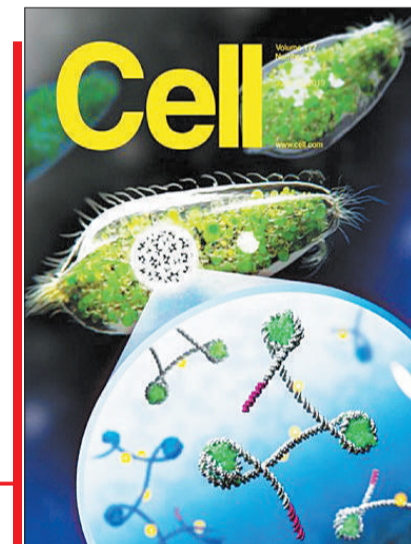
《当代生物学》 2019.6.17



世界生态系统的特点是资源分配不均,而不同物种之间的“贸易伙伴关系”可以帮助个体应对这种资源不平等。“贸易”允许个人以较低的成本交换商品,从而获得其他方式不可能或更难获得的资源。然而,随着资源在时间或空间上变得越来越不完整,生物体如何改变其交易策略仍是未知的。荷兰语布鲁塞尔自由大学马修·怀特塞德等研究人员,调查了共生真菌如何协调与宿主的“贸易”,以应对其网络中不同程度的资源不平等。研究团队利用量子点技术来标记碳,展示了真菌如何将营养物质从丰富的地方转移到贫瘠的地方,从而增加资源的价值。

### 甲基化影响淡水生物尖毛虫的转录活性

《细胞》 2019. 6.13



单细胞生物尖毛虫是一种生活在绿藻上的淡水原生生物,其转录活性核包含数百万个碱基大小的染色体,每条染色体都具有短的核小体阵列和紧密的端粒。美国哥伦比亚大学的莱斯利·本等研究人员,通过重建在原位或异位具有腺嘌呤N6甲基化(6mA)的完整的尖毛虫染色体,研究了6mA对体外核小体占据的影响,发现6mA不利于核小体在体外的占有率。

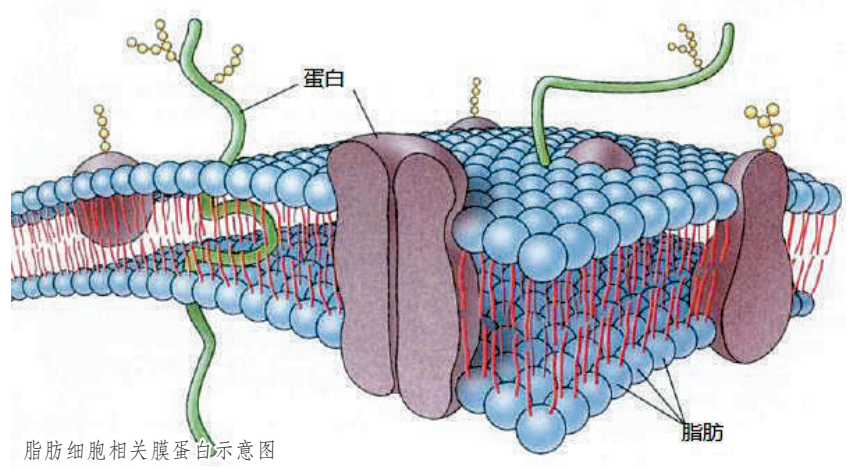
### 无脊椎动物是否感染细菌与年龄有关

《寄生虫学趋势》 2019.6



个体水平的年龄效应已广泛记录在包括昆虫和软体动物等无脊椎动物中。以色列特拉维夫大学佛罗里达·艾米讨论了无脊椎动物宿主年龄效应的流行病学、生态学和进化方面的问题,并将其与人类和其他脊椎动物已知的年龄相关现象进行了对比。研究表明,幼年水蚤比成年水蚤更容易受到巴斯德氏菌的感染,当暴露于两种菌株的寄生虫环境中时,幼年水蚤允许两者共存和传播,而在成年个体中可以观察到竞争性排斥。

(本版图片除标注外来源于网络)



脂肪细胞相关膜蛋白示意图