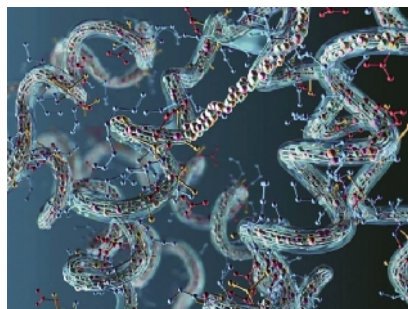


加入数据库 有了新名字 科学家加速“暗蛋白”研究

本报讯 人类基因组包含约两万个基因,它们携带着合成蛋白质的指令。不过,科学家发现,可能还存在数千种“暗蛋白”,它们在细胞中发挥着未知但很重要的作用。

这些蛋白质的编码来自基因组中被认为不会产生蛋白质的部分,因此未被纳入官方的基因组和蛋白质统计。但《自然》近日发表的一项研究为人类基因组编码的数千种此类分子赋予了一个正式名称——peptideins(肽蛋白),并将其纳入生命科学界广泛使用的基因和蛋白质数据库。研究人员表示,此次重新命名将为弄清不同肽蛋白在细胞中的作用带来助益。一些肽蛋白与包括儿童癌症在内的多种疾病以及基本的细胞功能有关。

“这是一个重大突破。”德国海德堡大学的生物信息学家 Christoph Dietrich 说,“这些微蛋白有可能开启一



根据最新研究,一些蛋白质需要在数据库中重新分类。

图片来源:Christoph Burgstedt

波新的研究浪潮。”

“暗蛋白”所含的氨基酸较少,因此长度极短,并且在其他物种中缺乏进化上的同源物。这也是它们被排除在数据库外的原因之一。不过,在大多数情况下,它们是由与已知蛋白质编码基因非常接近,甚至在某些情况下重叠的基因编码的。

致力于加速生物医学研究和创新的“TransCODE 联盟”研究团队对数千

种潜在“暗蛋白质”的实验数据进行了分析。他们从一份包含 7264 个疑似编码“暗蛋白质”的 DNA 序列入手,发现只有 15 个序列有足够的实验依据支持,可以被纳入蛋白质编码基因的官方目录。

在细胞中还能检测到数千种实验依据没那么充分、功能几乎未知的蛋白质片段。“目前我们还不确定它们是否在发挥作用,但它们确实存在。”欧洲生物信息研究所的生物信息学家 Jonathan Mudge 说。

Mudge 参与了 GENCODE 蛋白质编码基因数据库的工作。GENCODE 及其他数据库正在为肽蛋白创造一个新目录。Mudge 表示,这有助于吸引其他科学家关注它们,从而弄清其功能。

Mudge 说,此前被认为编码“暗蛋白质”的约 10 个序列已被移入由 GENCODE 维护的官方蛋白质编码基

因数据库。随着证据不断积累,预计会有更多的肽蛋白被添加其中。此外,一些典型的蛋白质可能被重新归类为肽蛋白。“这与发现数千颗围绕太阳运行的矮行星,迫使天文学家重新定义行星如出一辙。”Mudge 说。

在最新发表的论文中,研究人员发现有迹象表明,超过 50 种肽蛋白在某程度上对细胞至关重要。此前的研究表明,如今被称为肽蛋白的这些分子可能是某些癌症(如一种影响儿童的侵袭性脑癌)的重要驱动因素,而且对心脏功能也至关重要。

TransCODE 联盟的联合负责人、荷兰公主玛克西玛儿科肿瘤中心的系统生物学家 Sebastiaan van Heesch 预计,会涌现更多此类发现,“人们再也不能对它们视而不见了”。 (徐锐)

相关论文信息: <https://doi.org/10.1038/s41586-026-10459-x>

艾滋病曾推动人类快速进化

本报讯 在强效艾滋病药物出现之前,这种疾病给南非造成了严重的损失,以至于在短短十多年间,它就在人类基因组上留下了印记,改变了免疫系统基因的出现频率。20 年前,随着药物的普及,这种进化力量逐渐减弱,基因变化也随之放缓。近日,相关成果发表于美国《国家科学院院刊》。

这项研究生动展现了人类的快速进化。“这太令人惊叹了。”美国亚利桑那大学的 Michael Worobey 表示,“这清楚地展示了自然选择的作用,以及药物干预如何阻止自然选择。”

全球目前约有 4080 万艾滋病病毒(HIV)携带者,其中南非占 20%。基于 1998 年至 2025 年采集的近 1600 名感染或未感染 HIV 的母亲及 400 多名婴儿的血样,研究人员评估了南非受艾滋病影响最严重的夸祖鲁-纳塔尔省人口的基因变化。

英国牛津大学的 Philip Goulder 表示,在抗逆转录病毒药物 2005 年在夸祖鲁-纳塔尔省广泛使用之前,曾出现过一场“完美风暴”,多种因素相互交织,推

动了一场毁灭性的流行病。

20 世纪 90 年代初,HIV 进入南非,随后疫情在人群中暴发,导致夸祖鲁-纳塔尔省约 40% 的孕妇被感染——这一惊人的高感染率至今未变。由于遗传、医疗资源有限和可能存在的流行病毒亚型等多重原因,感染者在免疫系统受损的情况下迅速发展为艾滋病——这一过程大约需要 4.5 年,而在北美地区则是 10 年。

其他研究曾揭示疟疾和结核病等传染病如何改变人类基因组,但那些变化历经了数千年。

在夸祖鲁-纳塔尔省血液样本的 DNA 序列数据中,Goulder 和同事追踪了免疫系统中一种最精妙且有效的抗感染机制的变化。当细胞被 HIV 等感染时,人类白细胞抗原(HLA)分子会将病毒片段运送到细胞表面。在那里,它们就像有手一样将病毒片段拿起来展示。而被称为细胞毒性 T 淋巴细胞(CTL)的免疫系统“神枪手”会将它们识别为外来细胞,并杀死受感染的细胞。该研究重点关注了编码 HLA 分子的基因变异或

等位基因。

在 HIV 感染者体内,病毒会迅速演变出大量不同的变异株。一些 HLA 分子比其他分子更有效,能够捕获所有变异株共有的病毒片段;而另一些分子则更倾向于捕获病毒中频繁变异的片段,这在无形中削弱了 CTL 的防御能力。研究人员追踪了 3 种被称为“保护性”的 HLA 等位基因和 3 种被称为“易感性”的 HLA 等位基因的频率。

基于用数据构建的模型,研究人员发现,携带任何一种易感等位基因的人口比例从 1990 年的 28% 下降至 2004 年的 25%。在同一时期,携带任何一种保护性等位基因的人群比例从 23% 上升至 27%,这意味着存活人群中能够产生强效 CTL 反应的人口比例增加了。论文作者、美国弗雷德里克国家癌症研究中心的 Mary Carrington 表示,这些变化“幅度不大,但十分明确”。

目前,约 80% 的南非 HIV 感染者都能获得抗逆转录病毒药物,这虽然未能阻止这种进化趋势,但显著减缓了它的进程。研究人员预测,到 2035 年,22% 的

南非人口将携带易感等位基因,而 32% 的人口将携带保护性等位基因。研究指出,若夸祖鲁-纳塔尔省居民从未接触过抗逆转录病毒药物,到 2035 年,易感等位基因的流行率将从 1990 年的 29% 下降至 18%,而保护性等位基因的流行率则将跃升至 42%。

美国乔治城大学的 Tomoko Steen 表示,从更长的时间尺度看,从 HLA 频率中观察到的“定向选择”往往会被其他因素抵消。由于大多数等位基因在全球范围内均有分布,因此“似乎是迁移历史而非自然选择塑造了这些频率”。此外,由于保留不同的等位基因具有优势,即所谓的平衡选择,因此维持基因多样性的机制也会抵消大部分定向选择的影响。

Goulder 认同这一观点:“任何基因的存在,都是因为它经历了某种选择。实际上,那些对 HIV 不利的基因肯定对其他方面有益,只是我们目前还不知道。”

(文乐乐)

相关论文信息: <https://doi.org/10.1073/pnas.2502683123>