

化学计量基因组学研究进展

蓝洋, 胡江涛, 张玉娟

重庆师范大学生命科学学院, 重庆 401331

摘要: 化学计量基因组学是一个新兴的研究领域, 研究基因组、转录组、蛋白质组及代谢组等组学数据中生物大分子的元素使用偏好。不同生物大分子的元素组成与数量不同, 当元素供应受到限制, 自然选择偏好性利用某些单体(氨基酸或核苷酸)来合成生物大分子(DNA、RNA 和蛋白质等), 从而减少元素成本。随着高通量测序技术和组装技术的大量应用, 越来越多的宏基因组、宏转录组数据被公开报道, 以及新的分析手段的应用, 使得该领域蓬勃发展。作为一门新兴的交叉学科, 化学计量基因组学整合了化学计量学、生态学、进化生物学、基因组学和生物信息学, 为研究分子进化和生态系统的互作以及后基因组时代数据的挖掘提供了一个全新的视角。本文从蛋白质和核酸的元素使用偏好性方面综述了化学计量基因组学的最新研究成果, 并探讨了新的研究方向, 以期为该领域的研究和应用提供有益参考。

关键词: 元素使用偏好性; 进化选择; 物种分类; 精准医学; 生命元素周期表

Research progress of stoichiogenomics

Yang Lan, Jiangtao Hu, Yujuan Zhang

College of Life Sciences, Chongqing Normal University, Chongqing 401331, China

Abstract: Stoichiogenomics is a newly arisen research field, which concerns the element usage biases of biological macromolecules at genome, transcriptome, proteome, metatranscriptome levels. Different biological macromolecules have different element compositions and contents. When the supply of some elements was constrained, natural selection might bias the usage of the monomers (amino acid or nucleotide) to reduce constrained element costs in the synthesis of biological macromolecules. This field is flourishing with the intensive applications of high throughput sequencing and assembly technologies, more and more available metagenomic and metatranscriptomic data, and the applications of new analysis strategies. As a newly emerged cross discipline field, stoichiogenomics integrates stoichiometry, ecology, evolutionary biology, genomics and bioinformatics to provide a whole new perspective for investigating the interactions of the macromolecular evolution and ecosystem, and data mining in the post genomic era. In this review, we summarize the latest research progress of stoichiogenomics from the aspect of the element usage biases in proteins and nucleic acids. Furthermore, new research directions are discussed to provide some valuable references for the

收稿日期: 2016-10-16; 修回日期: 2016-12-30

基金项目: 重庆市教育委员会科学技术研究项目(编号: KJ1600304)资助[Supported by the Key Project of Natural Science Foundation of Education

Department of Chongqing (No.KJ1600304)]

作者简介: 蓝洋, 硕士, 研究方向: 生物信息学。E-mail: 908805060@qq.com

通讯作者: 张玉娟, 博士, 副教授, 研究方向: 生物信息、媒介昆虫、比较基因组学。E-mail: yujuan.zhang418@gmail.com

DOI: 10.16288/j.yczs.16-343

网络出版时间: 2017/2/20 9:16:34

URI: <http://www.cnki.net/kcms/detail/11.1913.R.20170220.0916.009.html>

research and application of stoichiogenomics.

Keywords: element usage biases; evolutionary selection; taxonomy; precision medicine; life element periodic table

化学计量基因组学(stoichiogenomics)是一门新兴的研究领域,最早被 Elser 等^[1]在 2011 年明确提出来,旨在研究基因组、转录组、蛋白质组、代谢组等组学数据中生物大分子的元素组成和含量,即元素使用偏好性。作为遗传信息的携带者和生命活动最重要执行者的生物大分子(核酸、蛋白质),其分别由核苷酸或氨基酸等单体组成。不同单体(核苷酸、氨基酸)所包含的元素(如氮、硫、碳等)组成和含量不同,导致核酸和蛋白质的元素使用偏好性不同。核苷酸是核酸的基本组成单位,每个核苷酸由不同数目的碳、氢、氧、氮、磷原子组成,其中氮原子数量为 2~5 个,磷原子数量为 1 个。同理,氨基酸是组成蛋白质的基本单位。在生物体中组成蛋白质的氨基酸约有 20 种。各种氨基酸之间的区别在于侧链基团(R 基)的不同。R 基携带不同数目的氧、硫、氮、碳、氢原子。不同的原子组成形成不同种类的氨基酸,不同种类氨基酸的排列顺序千变万化,肽链的盘曲折叠及其形成的空间结构千差万别,从而形成细胞中种类繁多的蛋白质。

在提出化学计量基因组学这个概念之前,2000 年 Elser 等^[2]提出了研究化学元素循环与生态环境平衡的生态化学计量学,并于 2002 年出版专著《Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere》。经过 10 多年的发展,生态化学计量学已成为了一门比较系统、成熟的学科^[3]。生态化学计量学是化学计量学(stoichiometry 或 chemometrics)在生物领域的一个分支,后者已经有几十年的发展历程^[4]。我们通过在 PubMed 数据库使用关键词搜索,发现化学计量学领域的文献数目为 17 880。生态化学计量学领域的文献数目为 187,化学计量基因组学领域文章数目为 25(统计数据截至 2016 年 8 月 31 日)。这表明近年来随着越来越多的基因组、转录组等组学数据的报道,化学计量基因组学越来越受到国内外研究者的重视。

在不同资源限制环境中,不同的生物和不同的生物大分子受到差异的选择力度^[5~7]。作为遗传信息的携带者和生命活动最重要执行者的生物大分子

(蛋白质、核酸),在不同资源限制条件下,其元素使用偏好性在不同的生物和生物大分子中发生了怎样的变化?对这些问题的探讨无疑具有十分重要的生物学意义。通过计算生物大分子中元素的使用偏好性,不仅可以掌握元素使用偏好性的进化规律,还可以此作为评估资源限制对生物大分子的直接影响。如果环境中某元素(如氮、硫、碳等)不够,自然选择可能会促使生物减少使用这些限制性元素,进而影响生物大分子的合成,并限制生物的生长和繁殖^[1]。已有的研究表明真核生物蛋白质与原核生物相比,其氧和硫元素具有使用偏好性^[8~10]。酵母细胞涉及硫代谢和碳代谢的酶,其相应元素含量也有偏好性现象^[11]。在微生物和植物中,还发现元素使用偏好性与蛋白质的表达具有相关性^[12, 13]。并且不同代谢途径中生物大分子的元素使用偏好性不同^[14, 15]。另外,还发现元素使用偏好性与基因组的不同转录区和基因组大小及核苷酸的长度有关^[16~18]。

通过对地球极端生活环境中生物使用元素偏好性研究,不仅可以解释环境对生物大分子元素使用偏好的影响,对于太空生物学(astrobiology)的研究也有很好的借鉴意义,有助于在外太空发现新的生命迹象及可能的元素使用规则。美国国家航空航天局(NASA)资助了 700 万美元的项目“Follow the elements”^[19]正是基于此项研究。但国内目前仍缺少该领域的相关介绍及研究。本文基于蛋白质和核酸水平,针对不同进化地位或生态环境中生物的化学计量基因组学研究进行综述,并对该领域的发展研究做出展望。

1 蛋白质的元素使用偏好性

相比核酸而言,蛋白质才是生命活动的主要承担者,其直接控制生物的性状,因此更直接的受到自然选择^[20]。合成一种蛋白质的成本与收益涉及生理和生态领域的相互作用。首先,不同氨基酸的合成成本不同。其次,不同生物所处的生态环境决定着资源(能量或元素)的利用效益。因此蛋白质的元素使用偏好受到合成不同氨基酸的能量成本和不同氨

积的效应), 因而运作效率低。

另一项研究^[15]也表明氮限制影响代谢途径和细胞成分的进化。通过检测人、小鼠、果蝇、拟南芥等合成代谢与分解代谢过程中主要功能蛋白的元素组成, 发现动物代谢途径中蛋白质的含氮量普遍高于植物代谢途径中蛋白质的含氮量; 氮限制条件下的拟南芥分解代谢途径的蛋白质的含氮量比合成代谢的蛋白质及整个蛋白组的含氮量显著降低。这表明植物比动物受到更强的资源限制。

1.4 蛋白质元素使用偏好性的宏进化规律

原核细胞(生物)的起源至少 35 亿年前, 而真核细胞(生物)的起源至少 15 亿年前, 真核细胞比原核细胞出现得更晚^[24]。从原核细胞(生物)进化出真核细胞(生物)是生物进化史上最重要的跨越之一。在这一宏进化过程中, 蛋白质的元素组成和功能基团发生了怎样的变化? 本实验室通过对 1051 种原核生物和 66 种真核生物全基因组和直系同源蛋白分析, 发现真核细胞中的蛋白质较原核细胞中的蛋白质具有更高比例的氧、硫元素(图 1)。功能基团分析结果显示真核细胞中的蛋白质趋向于含有更高比例的巯基、羟基和氨酰基, 但甲硫基和羧基的含量更低。以上结果表明, 生物从原核到真核这一宏进化过程中, 氧元素和硫元素含量的明显增加及其所构成的相关功能基团的变化, 使得真核蛋白发展出更多可用于执行复杂多样的生命活动的功能基团。该研

究为原核生物到真核生物的宏进化提供了新的视角^[9]。

2 核酸的元素使用偏好性

资源限制影响蛋白质分子组成的同时, 也影响基因组和转录组的元素组成, 基因组、转录组和蛋白质组将协调一致地节约限制性元素。因为核苷酸的生物合成始于氨基酸, 影响氨基酸合成过程中的资源限制同样也会影响核苷酸进而影响内含子和其它非编码 DNA^[25]。DNA 的元素使用偏好性可能最小, 以氮元素为例, 双螺旋碱基互补配对原则缓冲了核苷酸中氮含量的变化(A 与 T 含 7 个氮, 而 G 与 C 含 8 个氮)。单链 RNA 的氮含量可能变化更大, 因为嘌呤比嘧啶含更多氮原子(A=5, G=5, U=2, C=3)。此外, 通常 RNA 对细胞生物量的贡献比 DNA 多 5~10 倍^[26], 所以“氮限制”或“磷限制”将主要影响细胞的转录组。

2.1 植物核酸的元素使用偏好性

由于植物基因组有内含子, 氮限制条件将首先影响内含子的含氮量, 从而促进整体上的“保氮效应”^[16]。在驯化条件下, 由于农作物获得了氮肥, 农作物转录组显示出比野生植物更高的含氮量。这表明, 长期营养缺乏将影响植物转录组的元素组成, 而密集的氮肥施用及其他驯化影响, 又将缓解作物中氮限制的自然选择压力^[16]。

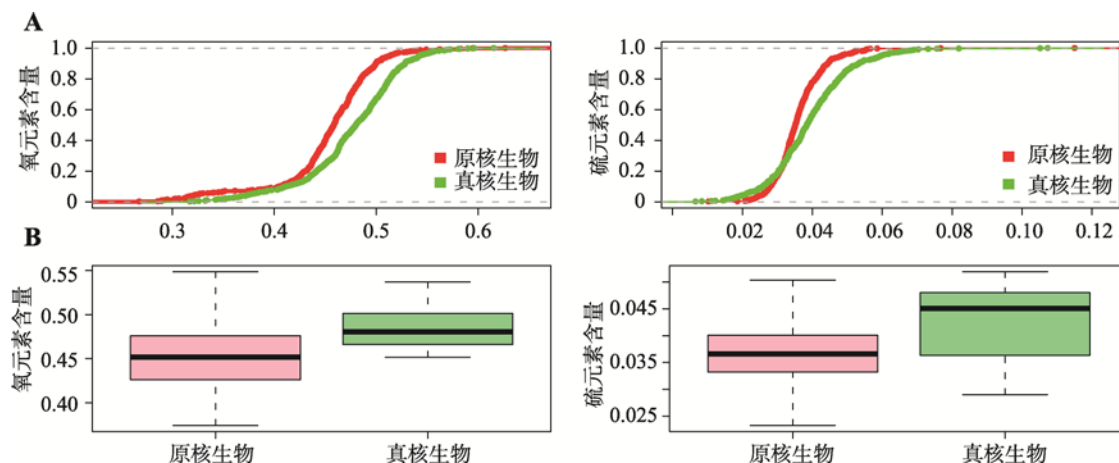


图 1 真核生物与原核生物蛋白质元素使用偏好性比较

Fig. 1 Comparisons of element usages in protein side-chains between prokaryotes and eukaryotes

A: 全基因组的蛋白质侧链氧元素和硫元素含量的累积分布曲线; B: 直系同源蛋白质侧链氧元素和硫元素含量的箱式分布图。根据参考文献^[9]的数据绘制。

2.2 资源限制影响基因组大小与核酸多聚体的长度

首先, 资源限制影响基因组大小。最新的研究通过对 99 种苦苣苔科(*Gesneriaceae*)植物基因组分析表明, 基因组大小与植株氮含量显著正相关, 但与植物磷含量的相关性不显著。这表明, DNA 合成和细胞功能对氮的营养竞争是苦苣苔科植物基因组大小变化的驱动者^[18]。其次, 每个核苷酸的合成都需要消耗氮和磷, 因而资源限制不仅影响核酸的元素组成, 也会影响核酸的长度。研究发现, 由于保氮、保磷效应, 拟南芥转录的非编码序列(内含子, 5'-UTR(untranslated regions), 3'-UTR)长度变短^[1]。虽然内含子长度与基因组大小相关, 但相同基因组大小的动、植物比较时, 植物平均内含子长度显著小于动物平均内含子长度。例如, 拟南芥和黑腹果蝇基因组大小相似, 但前者平均内含子长度却不到后者的 1/10^[1]; 对于 5'-UTR 和 3'-UTR 来说, 研究者发现植物基因中 5'-UTR 和 3'-UTR 序列更短且单个核苷酸氮含量更低。这些结果进一步说明了保氮、保磷机制塑造转录组的非编码区^[1]。

3 研究领域展望

元素使用偏好性的研究还处于初始阶段。一直以来由于具有完整注释的综合基因组数据的物种仍然有限, 迄今为止的研究工作仍集中在相对狭窄的范围内。然而由于高通量测序及组装技术的改进^[27, 28], 加快了不同物种及相同物种不同个体基因组测序的速度, 并且能够更加精确地量化基因和蛋白质的表达, 这些进步将有助于人们深化和拓展该领域研究。

3.1 宏基因组学、宏转录组学为元素使用偏好性研究拓展领域

由于高通量测序技术的革命^[29], 科学家获得越来越多的序列数据, 尤其是微生物^[30]的大量序列数据(DNA 与 RNA), 通常这些数据是结合监测关键环境变量所获得的^[31], 比如温度与化学条件(PH 盐度, 资源供给(氮、磷)等)。如在缺磷生长条件下, 海洋中的各种蓝藻(*Cyanobacteria*)用硫脂完全取代磷脂^[32]。在资源限制条件下, 可以预测生物大分子的某些特征, 如使用低生物合成成本、低含氮量的氨基酸偏好、较短的转录区及较短的聚合物(rRNA, 核糖体

蛋白等)。最近研究表明海洋表面与深海营养物可利用性的差异影响了适应不同深度的原绿球藻菌株蛋白质组的元素组成。深海菌株由于低可见光、高氮浓度, 其蛋白质组的元素组成支持含碳低、含氮高的氨基酸, 而海洋表面(光照丰富、低氮)的菌株恰恰相反^[33]。同样, 这些研究还可以拓展到土壤微生物和不同食性动物的肠道微生物, 如果它们的氮、磷总体上的相对利用率发生了变化。

3.2 揭示主要的进化选择事件

未来的研究将关注一些主要的进化事件。例如, 众所周知被人类作为粮食生产农作物中, 驯化引起了一系列相关基因的改变^[34]。因为部分驯化过程涉及土壤肥力的改变, 资源限制减缓了作物蛋白质组和转录组的选择压力。比较驯化物种与其祖先或近亲, 如栽培稻(*O. sativa*)与野生稻(*O. rufipogon*)或玉米与大刍草(*Purus frumentum*)的比较, 可使研究者评估资源限制减缓与选择高产品种对元素使用偏好性的影响^[16]。

3.3 推测物种的进化方向

关注物种进化的选择性事件, 其实质就是在考虑环境改变的压力条件下, 物种自身元素使用偏好的改变, 从而影响该物种适应性功能的形成^[35], 于是这就为推测物种的进化方向提供了一个新的研究思路。进化的本质动力就是环境选择, 现如今人们已经可以测算一定区域内环境条件因子的走向^[36, 37], 因此就可以模拟该地区的某物种在某种环境条件改变的情况下, 测序其基因组、转录组与蛋白质组, 以化学计量基因组学这一工具衡量, 物种自身情况对元素使用偏好的改变, 并确定该物种的哪些基因或蛋白表达增强或减少, 以推测该物种可能会形成哪种生物学功能, 从而大致演绎出其进化的方向。

3.4 辅助物种的分类研究

在整个生命进化的历程中仍然存在不少的盲点与困惑, 由于系统发育的亲缘关系无法确定, 往往导致了某些物种的分类地位在科学界各执一词^[38]。比如中生动物分类地位难以确定的问题, 在原生生物向后生生物进化过程中的存在众多节点问题, 究竟是哪一类的原生生物形成了动物(中生生物可能是进化中间的过渡态)? 类似于这样一系列的问题

困扰了进化生物学家与生物分类学家很多年^[39], 所以才有了 21 世纪启动的生命之树计划(Tree of Life web project, <http://www.tolweb.org/tree/>)。通过化学计量基因组学, 有希望对这些门类中各个物种元素使用偏向性的统计, 将之与具有类似的元素使用偏向性的物种对比总结在一起, 从而辅助物种的分类研究, 以确定该门类的系统发育地位。

3.5 助力医学的精准治疗

除此之外, 化学计量基因组学的研究手段也许可以应用在微观的生命世界、造福人类。在医学中不同的病变组织具有不同的组织微环境, 而这样的微环境正是影响病变组织发展的一个主要因素^[40]。比如就肿瘤细胞而言, 其通常生活在酸性的环境条件下^[41], 利用化学计量基因组学的手段了解不同化学物质导致的酸性微环境中肿瘤细胞化学元素使用的偏向性, 再将其富集到通路之中, 就可以明确其在发挥怎样的作用。这就可以为控制不同类微环境下, 肿瘤的靶向治疗提供切实可行的理论基础, 乃至解决方案。不仅仅在应对肿瘤时可以如此, 在不同病毒、寄生虫在人体特定的微环境中肆虐时, 化学计量基因组学的手段将为精准医学这一新世纪医学的弄潮儿添砖加瓦。

3.6 启迪太空生命探寻的新思路

窥探了化学计量基因组学在微观世界的应用, 再看看其在宏观世界的前景如何? 地球已经经历过 45 亿年的历程, 生命世界的形成也有了 24 亿年的历史^[42]。随着化学计量基因组学手段的广泛应用, 人类可以积累建立起地球上“生命元素的周期表”, 探究出从低等生物到高等生物不同元素的使用偏好性, 各大类生物生命元素的构成模式, 甚至是能够归纳推断出生命元素的演变脉络, 以及生命与环境元素的交互作用模式。这就可以给我们在外太空判断生命的存在, 探寻生命的诞生提供了一个参考标尺, 利用这把标尺比对外星球的气候、地质等等环境因子, 将会给我们打开探寻生命的新思路。

4 研究技术的革命与展望

4.1 变革中的信息储存方式

在生物信息学日新月异的今天, 数据量的日积

月累、更新变幻已经成为了生命科学领域高速发展的一大标志。目前我们所使用储存大数据的材料均是以硅为主体的半导体材料, 其以 0、1 编码的形式储存信息, 然而其存储的数据量密度有限, 无法运用极小的空间与资源储存大数据。就化学计量基因组学而言, 如果没有存储方式的变革很难保证处理基本数据的大量获取, 可以预见不久的将来这样的存储方式会极大地限制人类社会的发展, 所以有人提出科学设想可以用 2 斤 DNA 就能存储世界上所有的数据, 这样就可以极大地减少空间的占有量以及对环境资源的浪费^[43]。虽然这仅仅是一个设想, 但是已经有科学家在积极地探索与实验之中^[44], 相信在不远的未来, 信息储存方式的变革将会化学计量基因组学的研究, 提供强有力的基因组大数据或蛋白质组大数据的基本保障。

4.2 新涌现的生物信息手段

化学计量基因组学的研究需要核酸和蛋白质序列数据库的快速发展, 基因产物生物学功能的了解及基因表达数据的使用, 同时还需要生态学、进化和遗传学信息交叉的生物信息学知识库。比如在线资源 <http://www.graspdb.net>, 是最近建立的 12 种果蝇基因组的元素使用偏好性数据库^[45], 这些果蝇食性相异或具有在不同生境的进化历史^[46]。这些资源使生物学家能够在果蝇基因组中寻找元素使用偏好性的证据。另外, “Nutrilyzer”是一个用于展示差异表达的同源蛋白元素使用偏好性的工具^[47], 利用该软件可以展示资源限制和正常条件下差异表达的同源蛋白的元素使用偏好性。尽管已有上述工具, 未来还需要一个更灵活、可扩展的系统来涵盖完整的生态数据(描绘物种关键的生活史特征), 同时包含核酸和蛋白质的元素使用偏好分析。这些资源将会成为探索和分析生物大分子元素使用偏好性的工具, 并面向研究领域免费开发。

5 结 语

随着第二代高通量测序技术^[48]和组装技术^[49]的发展, 使得人们获得一个物种的基因组、转录组和表达数据更容易^[50-52]。如何从堆积如山的数据中挖掘出有用信息是人们面临的难题。化学计量基因

组学研究在研究大分子进化和种群遗传学方面的成功运用表明, 它为基因组、转录组和蛋白质组的进化和生态系统的互作以及后基因组时代数据的挖掘提供了一个全新的视角。本文综述了元素使用偏好性的相关研究, 已有的研究表明微生物、植物和动物直接生活在不同程度的资源限制环境中, 元素使用偏好性存在于它们的基因组、转录组和蛋白质组, 其中在转录组和蛋白质组更明显。然而目前化学计量基因组学的研究只看到一些组学层面上的数据趋势, 资源限制的作用机理仍然未知。未来怎样更好地结合测序数据与资源限制条件下物种特征, 揭示出关键的功能基因和通路将是人们面临的挑战。不仅对完善化学计量基因组学和响应资源限制的机理有重要理论价值, 也将有利于在外太空发现新的生命迹象及可能的元素使用规则。生命科学研究的本质就是让人们知道从何而来、向何处去, 这样才能够让人们更好地认识微观的自我与宏观的生命, 相信化学计量基因组学作为一项刚刚兴起不久的交叉学科工具, 必然能够为人类认识自身、探索真理、健康生活的各个研究领域起跑助力。

参考文献(References):

- [1] Elser JJ, Acquisti C, Kumar S. Stoichiogenomics: the evolutionary ecology of macromolecular elemental composition. *Trends Ecol Evol*, 2011, 26(1): 38–44. [DOI]
- [2] Elser JJ, Sterner RW, Gorokhova E, Fagan WF, Markow TA, Cotner JB, Harrison JF, Hobbie SE, Odell GM, Weidner LJ. Biological stoichiometry from genes to ecosystems. *Ecol Lett*, 2000, 3(6): 540–550. [DOI]
- [3] Zeng DP, Jiang LL, Zeng CS, Wang WQ, Wang C. Reviews on the ecological stoichiometry characteristics and its applications. *Acta Ecol Sin*, 2013, 33(18): 5484–5492. 曾冬萍, 蒋利玲, 曾从盛, 王维奇, 王纯. 生态化学计量学特征及其应用研究进展. *生态学报*, 2013, 33(18): 484–5492. [DOI]
- [4] Li ML, Shi LM, Li ZL, Yu RQ. Chemometrics study and application. *Anal Lab*, 1991, 10(2): 55–58. 李梦龙, 石乐明, 李志良, 俞汝勤. 化学计量学研究及其应用. *分析实验室*, 1991, 10(2): 55–58. [DOI]
- [5] Gutteridge A, Pir P, Castrillo JI, Charles PD, Lilley KS, Oliver SG. Nutrient control of eukaryote cell growth: a systems biology study in yeast. *BMC Biol*, 2010, 8: 68. [DOI]
- [6] Francois CM, Duret L, Simon L, Mermillod-Blondin F, Malard F, Konecny-Dupré L, Planel R, Penel S, Douady CJ, Lefébure T. No evidence that nitrogen limitation influences the elemental composition of isopod transcriptomes and proteomes. *Mol Biol Evol*, 2016, 33(10): 2605–2620. [DOI]
- [7] Secco D, Whelan J. Toward deciphering the genome-wide transcriptional responses of rice to phosphate starvation and recovery. *Plant Signal Behav*, 2014, 9(4): e28319. [DOI]
- [8] Acquisti C, Kleffe J, Collins S. Oxygen content of transmembrane proteins over macroevolutionary time scales. *Nature*, 2007, 445(7123): 47–52. [DOI]
- [9] Zhang YJ, Yang CL, Hao YJ, Li Y, Chen B, Wen JF. Macroevolutionary trends of atomic composition and related functional group proportion in eukaryotic and prokaryotic proteins. *Gene*, 2014, 534(2): 163–168. [DOI]
- [10] Stamati K, Mudera V, Cheema U. Evolution of oxygen utilization in multicellular organisms and implications for cell signalling in tissue engineering. *J Tissue Eng*, 2011, 2(1): 2041731411432365. [DOI]
- [11] Baudouin-Cornu P, Surdin-Kerjan Y, Marlière P, Thomas D. Molecular evolution of protein atomic composition. *Science*, 2001, 293(5528): 297–300. [DOI]
- [12] Bragg JG, Wagner A. Protein material costs: single atoms can make an evolutionary difference. *Trends Genet*, 2009, 25(1): 5–8. [DOI]
- [13] Elser JJ, Fagan WF, Subramanian S, Kumar S. Signatures of ecological resource availability in the animal and plant proteomes. *Mol Biol Evol*, 2006, 23(10): 1946–1951. [DOI]
- [14] Carlson RP. Metabolic systems cost-benefit analysis for interpreting network structure and regulation. *Bioinformatics*, 2007, 23(10): 1258–1264. [DOI]
- [15] Acquisti C, Kumar S, Elser JJ. Signatures of nitrogen limitation in the elemental composition of the proteins involved in the metabolic apparatus. *Proc Biol Sci*, 2009, 276(1667): 2605–2610. [DOI]
- [16] Acquisti C, Elser JJ, Kumar S. Ecological nitrogen limitation shapes the DNA composition of plant genomes. *Mol Biol Evol*, 2009, 26(5): 953–956. [DOI]
- [17] Bohlin J, Snipen L, Hardy SP, Kristoffersen AB, Lagesen K, Dønsvik T, Skjerve E, Ussery DW. Analysis of intra-genomic GC content homogeneity within prokaryotes. *BMC Genomics*, 2010, 11: 464. [DOI]
- [18] Kang M, Wang J, Huang HW. Nitrogen limitation as a driver of genome size evolution in a group of karst plants. *Sci Rep*, 2015, 5: 11636. [DOI]

- [19] NASA Astrobiology Institute. EPO Activity: Follow the elements EPO Formal/ Informal Education and Public Outreach. <http://nai.nasa.gov/annual-reports/2012/asu/epo/follow-the-elements-epo-formal-informal-education-and-public-outreach/>, 2008-12-12. [DOI]
- [20] Diz AP, Martínez-Fernández M, Rolán-Alvarez E. Proteomics in evolutionary ecology: linking the genotype with the phenotype. *Mol Ecol*, 2012, 21(5): 1060–1080. [DOI]
- [21] Akashi H, Gojobori T. Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc Natl Acad Sci USA*, 2002, 99(6): 3695–3700. [DOI]
- [22] Scheibel T, Bell S, Walke S. *S. cerevisiae* and sulfur: a unique way to deal with the environment. *FASEB J*, 1997, 11(11): 917–921. [DOI]
- [23] Li N, Lv J, Niu DK. Low contents of carbon and nitrogen in highly abundant proteins: evidence of selection for the economy of atomic composition. *J Mol Evol*, 2009, 68(3): 248–255. [DOI]
- [24] Martin W, Russell MJ. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos Trans R Soc B Biol Sci*, 2003, 358(1429): 59–85. [DOI]
- [25] Hessen DO, Jeyasingh PD, Neiman M, Weider LJ. Genome streamlining and the elemental costs of growth. *Trends Ecol Evol*, 2010, 25(2): 75–80. [DOI]
- [26] Sterner RW, Elser JJ. Ecological stoichiometry: the biology of elements from molecules to the biosphere. Princeton, NJ: Princeton University Press, 2002. [DOI]
- [27] Martin JA, Wang Z. Next-generation transcriptome assembly. *Nat Rev Genet*, 2011, 12(10): 671–682. [DOI]
- [28] Pendleton M, Sebra R, Pang AWC, Ummat A, Franzen O, Rausch T, Stütz AM, Stedman W, Anantharaman T, Hastie A, Dai H, Fritz MHY, Cao H, Cohain A, Deikus G, Durrett RE, Blanchard SC, Altman R, Chin CS, Guo Y, Paxinos EE, Korbelt JO, Darnell RB, McCombie WR, Kwok PY, Mason CE, Schadt EE, Bashir A. Assembly and diploid architecture of an individual human genome via single-molecule technologies. *Nat Methods*, 2015, 12(8): 780–786. [DOI]
- [29] DeLong EF. The microbial ocean from genomes to biomes. *Nature*, 2009, 459(7244): 200–206. [DOI]
- [30] Riesenfeld CS, Schloss PD, Handelsman J. Metagenomics: genomic analysis of microbial communities. *Annu Rev Genet*, 2004, 38(1): 525–552. [DOI]
- [31] Jones N. Undersea project delivers data flood. *Nature*, 2010, 464(7292): 1115. [DOI]
- [32] Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblížek M, Lomas MW, Mincer TJ, Moore LR, Moutin T, Rappé MS, Webb EA. Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature*, 2009, 458(7234): 69–72. [DOI]
- [33] Lv J, Li N, Niu DK. Association between the availability of environmental resources and the atomic composition of organismal proteomes: evidence from *Prochlorococcus* strains living at different depths. *Biochem Biophys Res Commun*, 2008, 375(2): 241–246. [DOI]
- [34] Burke JM, Burger JC, Chapman MA. Crop evolution: from genetics to genomics. *Curr Opin Genet Dev*, 2007, 17(6): 525–532. [DOI]
- [35] Hecht M. Evolutionary biology. US: Springer Science & Business Media, 2012. [DOI]
- [36] Qin DH, Luo Y. Global climate change causes and future trends. *Impact Sci Soc*, 2008, (2): 16–21.
秦大河, 罗勇. 全球气候变化的原因和未来变化趋势. 科学对社会的影响, 2008, (2): 16–21. [DOI]
- [37] Duan QY, Xia J, Miao CY, Sun QH. The uncertainty in climate change projections by global climate models. *Chin J Nat*, 2016, 38(3): 182–188.
段青云, 夏军, 缪驰远, 孙巧红. 全球气候模式中气候变化预测预估的不确定性. 自然杂志, 2016, 38(3): 182–188. [DOI]
- [38] Oliver PM, Adams M, Lee MS, Hutchinson MN, Doughty P. Cryptic diversity in vertebrates: molecular data double estimates of species diversity in a radiation of Australian lizards (*Diplodactylus*, Gekkota). *Proc R Soc B Biol Sci*, 2009, 276(1664): 2001–2007. [DOI]
- [39] Song DX, Zhou KY. Assessing biodiversity: a heavy task. *J Nanjing Normal Univ (Nat Sci)*, 2002, 25(2): 1–6.
宋大祥, 周开亚. 生物多样性的评估仍是一项艰巨的工作. 南京师大学报(自然科学版), 2002, 25(2): 1–6. [DOI]
- [40] Montalvo-Javé EE, Olguín-Martínez M, Hernández-Espinosa DR, Sánchez-Sevilla L, Mendieta-Condado E, Contreras-Zentella ML, Oñate-Ocaña LF, Escalante-Tattersfield T, Echegaray-Donde A, Ruiz-Molina JM, Herrera MF, Morán J, Hernández-Muñoz R. Role of NADPH oxidases in inducing a selective increase of oxidant stress and cyclin D1 and checkpoint 1 over-expression during progression to human gastric adenocarcinoma. *Eur J Cancer*, 2016, 57: 50–57. [DOI]
- [41] Damaghi M, Gillies R. Phenotypic changes of acid adapted cancer cells push them toward aggressiveness in

- their evolution in the tumor microenvironment. *Cell Cycle*, 2016, doi: 10.1080/15384101.2016.1231284. [DOI]
- [42] Wu XY, Chen SL, Ge Md, Chen YZ. General biology. 4th ed. Beijing: China Higher Education Press, 2014.
吴相钰, 陈守良, 葛明德, 陈阅增. 普通生物学. 第4版. 北京: 高等教育出版社, 2014. [DOI]
- [43] Zhirnov V, Zadegan RM, Sandhu GS, Church GM, Hughes WL. Nucleic acid memory. *Nat Mater*, 2016, 15(4): 366–370. [DOI]
- [44] Church GM, Gao Y, Kosuri S. Next-generation digital information storage in DNA. *Science*, 2012, 337(6102): 1628. [DOI]
- [45] Gilbert JDJ, Acquisti C, Martinson HM, Elser JJ, Kumar S, Fagan WF. GRASP [Genomic Resource Access for Stoichioproteomics]: comparative explorations of the atomic content of 12 *Drosophila* proteomes. *BMC Genomics*, 2013, 14: 599. [DOI]
- [46] Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN, Ruby JG, Brennecke J, Harvard FlyBase Curators, Berkeley *Drosophila* Genome Project, Hodges E, Hinrichs AS, Caspi A, Paten B, Park SW, Han MV, Maeder ML, Polansky BJ, Robson BE, Aerts S, van Helden J, Hassan B, Gilbert DG, Eastman DA, Rice M, Weir M, Hahn MW, Park Y, Dewey CN, Pachter L, Kent WJ, Haussler D, Lai EC, Bartel DP, Hannon GJ, Kaufman TC, Eisen MB, Clark AG, Smith D, Celniker SE, Gelbart WM, Kellis M. Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature*, 2007, 450(7167): 219–232. [DOI]
- [47] Lotz K, Schreiber F, Wünschiers R. Nutrilizer: a tool for deciphering atomic stoichiometry of differentially expressed paralogous proteins. *J Integr Bioinform*, 2012, 9(2): 196. [DOI]
- [48] Buermans HPJ, den Dunnen JT. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta*, 2014, 1842(10): 1932–1941. [DOI]
- [49] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma FD, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*, 2011, 29(7): 644–652. [DOI]
- [50] Zhang YJ, Hao YJ, Si FL, Ren S, Hu GY, Shen L, Chen B. The *de novo* transcriptome and its analysis in the world-wide vegetable pest, *Delia antiqua* (Diptera: Anthomyiidae). *G3 (Bethesda)*, 2014, 4(5): 851–859. [DOI]
- [51] Chen B, Zhang YJ, He ZB, Li WS, Si FL, Tang Y, He QY, Qiao L, Yan ZT, Fu WB, Che FY. *De novo* transcriptome sequencing and sequence analysis of the malaria vector *Anopheles sinensis* (Diptera: Culicidae). *Parasit Vectors*, 2014, 7: 314. [DOI]
- [52] Gilissen C, Hehir-Kwa JY, Thung D, van de Vorst M, van Bon BWM, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BBA, Kleefstra T, Brunner HG, Vissers LELM, Veltman JA. Genome sequencing identifies major causes of severe intellectual disability. *Nature*, 2014, 511(7509): 344–347. [DOI]

(责任编辑: 于黎)