定量工程生物学的化学蛋白质组学支撑性技术

王蕾^{1,2†}, 黄建东^{1†}, 杨舒心¹, 黄术强^{1,2*}, 李楠^{1,2*}

1. 中国科学院深圳先进技术研究院, 深圳合成生物创新研究院, 中国科学院定量工程生物学重点实验室, 广东省合成基因组学重点实验室, 深圳 518055;

2. 中国科学院大学, 北京 100049

† 同等贡献

* 联系人, E-mail: shuqiang.huang@siat.ac.cn; nan.li@siat.ac.cn

2020-06-17 收稿, 2020-07-26 修回, 2020-07-27 接受, 2020-07-28 网络版发表

国家重点基础研究发展计划(2018YFA0902703)、深圳市科技创新委员会项目(JCYJ20170818164014753, KQTD2016112915000294, KQTD2015033117210153)、深圳合成生物学创新研究院项目(ZTXM20190015, ZTXM20190012, DWKF20190002, JCHZ20190002)、广东省科学技术厅项目(2018A030310685, 2018A030310035)、广东省合成基因组学重点实验室项目(2019B030301006)和国家自然科学基金(31800694, 31971354)资助

摘要 定量工程生物学是一门前沿交叉学科, 通过设计-合成-测试-学习-再设计路线将不同的生物元器件组合, 形成可以执行特定功能的基因线路, 再经过不断优化获得稳定的、可控的基因线路, 最后将设计优化后的线路引入不同的生命体, 以达到预设的目的. 这种变革性的方法可以创建一些能够灵敏感知和响应各种环境的工程系统, 但在其中的功能检测环节, 化学蛋白质组学技术则成为了测试工程改造生物功能和探究其作用机制的重要工具. 随着以非天然氨基酸嵌入、生物正交化学、高分辨率质谱等技术为手段的化学蛋白质组学方法的发展, 在复杂环境中解析工程生物的蛋白质组时空动力学变化成为可能, 为探究工程改造菌或工程改造细胞的工作原理及其在生物体内的作用机制提供了必要的技术支撑, 也为定量工程生物学研究中所需的深度功能测试提供了有效方法. 本文主要是概述化学蛋白质组学技术在定量工程生物学研究中的潜在应用.

关键词 定量工程生物学, 基因线路, 化学蛋白质组学, 生物正交化学

自20世纪末以来, 生命科学发生着日新月异的变化. 生命科学也逐步由定性、局部性的研究向定量、系统化、工程化的方向发展, 定量工程生物学(quantitative engineering biology: 即通过理性设计和精准地导入人工系统来定量地认知生命机制, 践行“造物致知”^[1])应运而生. 定量工程生物学基于对现有生命系统或各种元器件功能的认识, 按照工程化的设计原理对生命系统进行简化处理, 以最优化的方式重新编程, 构建出多种优化的基因线路, 使其发挥预先设定的功能. 例如, 生产新型生物能源和生物材料、合成多种临床药物以及特异性杀伤肿瘤细胞等. 另一方面, 工程生物

学以超越自然进化法则的方式对天然生物系统进行人工干扰、重建甚至创造新的生命体, 并以此研究复杂生物系统的运行规律以及生物进化的历程.

定量工程生物学在临床疾病治疗、化工生产等领域的广泛应用彰显出其强大的适应性. 但有一个问题不容忽视, 即将设计、优化后的基因线路放在底盘细胞中, 对细胞自身的正常生长以及细胞所处微环境是否有影响, 有哪些影响? 工程改造的细胞能否像天然细胞一样, 在其所处的微环境中正常生长, 拥有正常的细胞通讯功能, 还有待进一步验证.

蛋白质是生物功能的执行者, 生物体内蛋白质处

引用格式: 王蕾, 黄建东, 杨舒心, 等. 定量工程生物学的化学蛋白质组学支撑性技术. 科学通报, 2021, 66: 356-366

Wang L, Huang J D, Yang S X, et al. Chemical proteomics techniques and their perspective enabling roles in quantitative engineering biology (in Chinese). Chin Sci Bull, 2021, 66: 356-366, doi: [10.1360/TB-2020-0457](https://doi.org/10.1360/TB-2020-0457)

于动态变化的状态,生物体通过调控蛋白质的合成和降解的速率以及翻译后修饰来适应其体内外环境的变化。因此想要对生命的复杂活动有全面而深入的认识,必然要在整体、动态网络的水平上对蛋白质进行研究。近些年来,随着化学蛋白质组学的发展,高分辨率质谱技术的进步以及计算机算法的革新,使得我们能够以高通量水平从复杂的生物样品中鉴定和定量数千种蛋白质,获取生物体蛋白质组的动力学信息,进而得以解析各种生命活动的生物学机制,能够更加深入地理解生命体,为定量工程生物学研究中所需的深度功能测试提供了强有力的技术支撑。

1 工程生物的蛋白质组时空动力学

随着工程生物学的快速发展,诸多人工改造的工程生物面世。这些工程生物有些可以为人类生产复杂的有机分子,有些则有助于人们对于生命奥秘的探索。以酵母细胞作为底盘细胞的研究为例。2003年,Keasling实验室^[2]将带有10个酶的合成通路装载到大肠杆菌中,实现了青蒿酸(青蒿素的前体)的合成。随后他们又将该合成途径优化后引入到发酵效率更高的酵母体系中,实现在酵母菌中生产抗疟疾药物青蒿素,并于2013年实现了量产^[3]。斯坦福大学的Smolke团队^[4]将合成阿片类药物的5个基因导入酵母基因组,实现在酵母细胞中生产阿片类药物。2018年,中国科学院上海植物生理生态研究所的覃重军课题组与纽约大学的Boeke团队^[5,6]发表了关于酵母染色体融合重塑基因组结构的研究成果,分别将酿酒酵母的16条染色体融合为1和2条染色体。研究表明,重塑基因组结构的酵母细胞与野生型酵母具有相似的转录组水平及表型。2019年,Keasling团队^[7]将大麻中的相关基因插入啤酒酵母,使得酵母能够利用半乳糖合成四氢大麻酚(THC)和大麻二酚(CBD),实现在酵母中生产“大麻啤酒”,该研究有望帮助生物制药公司以较低的成本、批量生产大麻中的各种化学成分。

虽然近20年工程生物学的研究与应用已经取得了令人瞩目的成果,但是我们对于工程生物本身还知之甚少,这严重影响着对工程生物的广泛开发及其工业化应用。目前关于工程微生物的研究主要局限于通过观察表型或者转录组分析,而蛋白质是生物功能的执行者,因此基因组、转录组信息不能准确展现工程生命体的信息,并且生物体内蛋白质合成和降解处于动态变化的状态,当机体处于应激条件下或者其所处的内外环境发生变化时,蛋白质组会发生瞬时变化,既有

大量新合成的蛋白质,又有一些蛋白质会发生快速降解。因此我们可以利用化学蛋白质组学技术研究工程生物的蛋白质动力学,加深对工程生物的认识。下面将从3个层面阐述用于研究蛋白质组动力学的化学蛋白质组学技术。

1.1 定性鉴定新合成的蛋白质组

非天然氨基酸嵌入技术最早由Schultz团队^[8]提出,已有30多年的历史。随后,以Schultz实验室为代表的多个团队创建了原核及真核生物遗传密码子扩增的方法学^[9-12],将带有多种不同报告基团(如荧光素、光交联基团、细胞毒素等)的非天然氨基酸位点特异性地嵌入新合成的蛋白质^[13-15]。2006年,Dieterich等人^[16]提出生物正交非天然氨基酸标记(bio-orthogonal non canonical amino acid tagging, BONCAT)策略,用于亲和富集、鉴定细胞中新合成的蛋白质^[17,18]。利用细胞内源性翻译机制,将带有生物正交基团的非天然氨基酸嵌入到新合成的蛋白质中^[19],生物正交基团作为代谢标签对蛋白质进行化学标记,将新合成的蛋白与预先存在的蛋白区分开来,从而降低了样品的复杂度^[16,20](图1)。在蛋白质合成过程中,位点特异性地向蛋白质中加入带有化学探针的非天然氨基酸^[21],赋予被标记的蛋

图1 (网络版彩色)BONCAT标记策略流程图

Figure 1 (Color online) BONCAT labeling strategy flowchart

白独特化学功能。随后利用生物正交基团的特性,通过Cu催化[3+2]叠氮-炔基环加成反应^[22,23]对新合成的蛋白质进行选择分离纯化、富集,或者通过荧光成像等手段观察蛋白质的动态变化^[17,18]。BONCAT不仅可以用作狭义蛋白质组学(即基于质谱分析的蛋白质组研究)的研究工具,也可以用作广义蛋白质组学研究(例如基于荧光成像^[24]的新生蛋白质组学研究等)的技术手段。类似的概念近年来也被推广到核酸^[25,26]、糖类^[27]和脂类^[28]的代谢标记等领域中,用于对这些物质进行富集分析、动力学变化检测或高灵敏度示踪。

研究证明,BONCAT技术广泛适用于追踪各种模型系统中新合成的蛋白质,例如细菌^[29]、哺乳动物细胞^[30,31]和斑马鱼等动物模型^[32]。可用于检测蛋白质组的动态变化^[33]、核糖体的转换率^[34]以及在活体中原位、可视化监测大鼠海马区神经元中新合成蛋白的动力学变化等^[30,35],研究当多巴胺能神经元受到化学刺激时,轴突是怎样做出应激反应以维持自身稳定^[36]。Schuman团队^[32]在活体斑马鱼模型中已成功运用BONCAT技术,实现对复杂神经系统中的新合成的蛋白质进行可视化观察和亲和纯化分析,证明长期记忆的形成过程会伴随大量新的蛋白质合成。Tcherkezian等人^[30]观察到结直肠癌缺失基因受体(deleted in colorectal carcinoma, DCC)与蛋白质合成位点的共定位,为纺锤体蛋白作为刺激神经元中核外蛋白质产生的作用提供了支持。

1.2 定量分析新合成的蛋白质组

为了克服BONCAT时间分辨率不足,以及不能对蛋白质组进行定量分析的技术限制,Acuto团队将BONCAT技术^[16]与SILAC(stable isotope labeling with amino acids in cell culture,即在细胞培养的过程中,在培养基中加入稳定同位素标记的氨基酸,经过细胞代谢过程,使得蛋白带上同位素标签)技术^[37]联用,将它们的优势结合在一起形成定量非天然氨基酸标记策略(quantitative non canonical amino acid tagging, QuaNCAT)^[38]。Bagert等人^[39]在HeLa细胞中联合使用BONCAT和pSILAC技术(pulsed SILAC,是指在较短的时间段内(从5 min到若干小时)对蛋白质进行脉冲式的SILAC标记),在30 min的脉冲标记的时间区间内,成功鉴定到1484种新合成的蛋白。这是单独使用同位素标记技术无法达到的时间分辨率。Howden等人^[38]利用QuaNCAT技术检测当原代T细胞经过佛波酯和离子霉素

素共刺激后,T细胞蛋白质表达组的变化。

BONCAT技术与SILAC技术联用,通过将新生蛋白亲和纯化标签与蛋白质谱定量标签两种代谢插入相结合,有助于评估细胞蛋白表达量的短时应激变化,具有更高的定量准确度和检测灵敏度。克服了BONCAT技术定量鉴定的准确性较低,以及SILAC技术标记时间长且仅适用于研究稳态蛋白质组的变化缺陷。QuaNCAT技术不需要对细胞或生物体进行饥饿处理,减少了对样品造成的干扰,因而所测得的数据能更好地反映真实的生理条件下蛋白质的动态变化情况。

COFRADIC(combined fractional diagonal chromatography)是一种基于色谱分离多肽的技术,由Gevaert等人^[40]首次应用于分离及富集带有甲硫氨酸的肽段。Kramer等人^[41]对该技术进行了改造,使得该技术可以特异性分离叠氮基团标记的肽段。其工作原理为:叠氮基团(azide)标记的-肽段经三(2-羧基乙基)膦(tris(2-carboxyethyl)phosphine, TCEP)处理后将诱导两条竞争的反应途径。在此过程中肽段中的叠氮化物部分转化为胺,形成二氨基丁酸酯,使得肽段的亲水性增加。在第二条反应途径中,肽段在azide残基处裂解,产生C端具有高丝氨酸内酯残基的N端裂解产物和无修饰的C端裂解产物。Koster团队^[41]将COFRADIC技术和同位素标记相对与绝对定量技术(isobaric tags for relative and absolute quantitation, iTRAQ)联用,对Met缺陷型大肠杆菌进行15 min热激处理,检测其蛋白质组的变化。应用生物正交的方法对短时间内合成的蛋白质进行相对定量分析。该项工作鉴定到527种蛋白质的表达发生改变,在定量蛋白质领域实现了一项新的突破。Gevaert等人^[40]应用该技术检测到大肠杆菌K12细胞的872种蛋白质,鉴定到的蛋白质中不仅包含高丰度蛋白,还包括一些低拷贝的蛋白,例如乳糖操纵子的阻遏物,表明该技术适用范围较广。

COFRADIC技术的优点在于能够通过色谱技术和化学反应充分移除未标记的蛋白质;利用色谱技术分离和富集被标记的蛋白,有利于质谱分析和检测低丰度的标记蛋白;联合使用COFRADIC和iTRAQ可以鉴定和定量蛋白质表达的瞬时变化;在较短的标记时间内即可检测出数百种蛋白,解决了同位素标记技术存在的标记时间长、分辨率低等问题。

1.3 时空特异性蛋白质组动力学分析

美国加州理工大学的Tirrel团队利用FUNCAT

(fluorescence non canonical amino acid tagging)技术对大肠杆菌^[17]以及多种哺乳动物细胞系^[18]进行标记,借助荧光共聚焦显微镜,可视化新合成蛋白质的空间分布。2015年, Schuman团队^[42]联合使用非天然氨基酸标记策略和蛋白质邻近连接标记技术(proximity ligation assay, PLA),实现了原位可视化观察新合成蛋白质的起源、转运、分布及其降解过程。

Salic团队^[43]在2012年提出OP-puro标记策略(OP-propargyl-puromycin, 邻-炔丙基-嘌呤霉素、嘌呤霉素的炔基类似物)通过生物正交的方法对生物体新合成的蛋白质进行标记,进而利用生物正交基团的化学特性对标记蛋白进行分离纯化或者荧光成像分析(图2)。利用该技术可视化地观测3T3细胞系在标记不同时间后,新合成的蛋白质的空间分布;将OP-puro以腹腔注射的方式加入活体小鼠体内,标记1 h后,观察小鼠肠上皮细胞新生蛋白的空间排布。

Weintz等人^[44]和Sharma等人^[45]的研究证明,巨噬细胞经脂多糖(lipopolysaccharides, LPS)诱导活化后,其分泌蛋白质组信息显示,巨噬细胞的磷酸化模式相关蛋白发生了高度动态变化。Eichelbaum等人^[46,47]将BONCAT技术与pSILAC技术联用,在脂多糖诱导巨噬细胞活化的模型中,经过最短20 min的标记后,在巨噬细胞的分泌组中鉴定到4852种蛋白质,包含多种与抗炎症等免疫反应相关的蛋白,绘制出巨噬细胞活化后的分泌图谱,解析了巨噬细胞在胞内病原菌感染过程中的蛋白质分泌动力学,对体内环境被外源细胞(或细菌)扰动时变化的评估方法做出了重要的探索。2016年, Sitek课题组成功将SPECS(secretome protein enrichment

with sugars)技术^[48]应用于研究T细胞的分泌组^[49],使用两种不同的点击化学试剂(生物素、去硫生物素),获得了T细胞活化过程中其效应蛋白的分泌图谱。

2 复杂环境中解析工程生物的蛋白质组

近年来,工程生物学在临床疾病治疗领域应用越来越广泛,在肿瘤、代谢病、重大传染病等方面已经取得了引人注目的研究成果,并有望解决目前临床面临的诸多难题。

通过工程生物学改造的细菌可以提高其靶向肿瘤、有效荷载外源基因的能力,使肿瘤的细菌疗法焕发生机。例如,在细菌体内安装群体感应(quorum sensing, QS)开关调控的基因线路,只有当细菌种群达到阈值密度时才能激活效应基因的表达,该线路能够有效提高细菌靶向肿瘤细胞的特异性^[50-52]。对于工程改造的沙门氏菌,利用其T3SS分泌系统向肿瘤细胞递送抗血管生成蛋白,可有效抑制体内肿瘤的生长^[53,54],或者将工程菌上的肿瘤相关抗原传递给抗原呈递细胞,激活免疫细胞从而引发抗肿瘤免疫^[55]。2019年,哥伦比亚大学Tal Danino课题组^[56]对大肠杆菌进行改造,作为递送免疫治疗药物的载体,协助免疫系统从内部攻破肿瘤。相较于肿瘤的细菌治疗,溶瘤病毒(oncolytic virus, OV)在临床应用上已经先行一步,它们具有裂解肿瘤特异性细胞和激活免疫反应的能力,可作为潜在的原位肿瘤疫苗。2015年,美国食品药品监督管理局(FDA)批准溶瘤单纯疱疹病毒应用于晚期黑色素瘤治疗^[57,58]。2019年,谢震课题组^[59]优化了肝细胞癌特异启动子和microRNA 感应器,控制腺病毒选择性地肝癌细胞中

图2 (网络版彩色)OP-puro标记策略。OP-puro是一种带有生物正交功能基团的蛋白质合成抑制剂,能够进入核糖体A位,与新合成的多肽链末端形成共价连接。OP-puro掺入后肽链延伸终止,释放出C端带有OP-puro的不成熟多肽

Figure 2 (Color online) OP-puro labeling strategy. OP-puro is a protein synthesis inhibitor with a bio-orthogonal functional group, which can enter the A position of the ribosome and form a covalent connection with the newly synthesized polypeptide chain. As long as OP-puro is incorporated, the peptide chain extension is terminated, and the immature polypeptide with OP-puro at the C-terminus is released

复制,裂解肿瘤细胞,实现了对肝癌细胞的特异性杀伤。

工程生物学研究在慢性代谢性疾病新疗法开发中也取得了令人惊喜的进展。2017年,叶海峰课题组^[60]结合光电信号,实现了远程调控胰岛素的释放,进而达到降低血糖的目的。2019年,该团队又设计合成了由绿茶次级代谢产物调控的胰岛素表达控制系统^[61],从而实现了通过引用绿茶来维持糖尿病小鼠血糖稳定。

在重大传染性疾病的新疗法方面,工程生物学同样取得了重大突破。2017年,北京大学邓宏魁实验室^[62]开发出一种利用CRISPR-Cas9对人成体造血干细胞进行基因编辑的技术体系。实验表明,经过基因编辑的造血干细胞能够在动物模型中长期稳定的重建造血系统,具有抵御HIV感染的能力。随后该研究组将供体来源的CD34⁺造血干细胞的CCR5敲除后移植到患者体内,经过改造的细胞在患者体内重建了被艾滋病破坏的机体免疫系统^[63]。该项研究成为世界上首例通过基因编辑干细胞治疗艾滋病和白血病患者案例。

诸多研究表明,工程生物学在疾病治疗方面具有极大的潜能,有望解决现阶段面临的医学难题。但是在将其应用于临床之前,需要解答相关的生物学问题:向宿主体内引入工程改造的细胞(或微生物等)后,对宿主造成了哪些影响?移植后的细胞能否与宿主体内建立正常的细胞通讯?了解其作用机制后,我们即可进一步优化基因线路,提高疗效,推进临床疾病治疗的进展。现已有大量有关细胞类型特异性化学蛋白质组学分析的研究报道,借助化学蛋白质组学技术研究工程生物体,绘制工程改造生物的完整蛋白质表达图谱,有助于增加对新生命体的生理活动的理解。

2.1 细胞特异性

在具有多种细胞类型的复杂环境中,想要研究某一种细胞的蛋白质组的动态变化,需要细胞特异性的分析技术。2016年,Barrett等人^[64]在OP-puro技术的基础上进一步优化,构建了PhAc-OP-puro化合物(OP-puro的类似物)。在表达青霉素G酰化酶(penicillin G acylase, PGA)的哺乳动物细胞中,PGA特异性水解PhAc的酰胺基团,随后苯乙酰基会自发断裂生成OP-puro,将OP-puro掺入新生多肽中。而对于野生型哺乳动物细胞来说PhAc-OP-puro是惰性的,因为OP-puro的 α -胺基被封闭了,不能够与肽链的C末端共价连接。通过在特定细胞类型中表达PGA(一种可以从PhAc-OP-puro中除去PhAc基团以生成OP-puro的酶),证明该策略

能够以细胞特异性方式标记和鉴定新生的蛋白质组。

随着BONCAT技术的发展,研究人员对其进行不同的改造使得该技术能够适用于各种复杂环境。2009年,Ngo等人^[65]对大肠杆菌的甲硫氨酸tRNA合成酶(MetRS)进行改造,通过高通量的筛选策略,筛选出一系列突变型MetRS,鉴别出NLL的特异性及活性最高,能够有效地将甲硫氨酸的类似物Anl嵌入到新合成的蛋白质中(NLL-MetRS^[65,66]; L13N、Y260L、H301L)。在改造型*E. coli*侵染巨噬细胞体系中,通过EcMetRS将Anl选择性嵌入到*E. coli*中,实现细胞特异性标记。2013年,该团队将来源于大肠杆菌的NLL-MetRS导入哺乳动物细胞中(NLL-EcMetRS^[67]),实现对哺乳动物细胞进行位点-特异性标记。因为来源于大肠杆菌的NLL-MetRS只能识别七元反密码子环,如细菌中tRNA^{Met}以及哺乳动物细胞中的起始tRNA^{Met},而哺乳动物细胞中的延伸tRNA^{Met}是九元反密码子,不能被活化。所以NLL-EcMetRS只能将Met的类似物Anl整合在肽链的N末端,但是因为大约80%蛋白质的N末端会被切除,故而该方法存在一定的局限性。

Tirrel课题组^[68]制备单基因突变型小鼠甲硫氨酸tRNA合成酶(MmMetRS_{L274G}),将其导入CHO、COS7、Hela细胞系。实验证实,非天然氨基酸ANL能够对表达突变型硫氨酸tRNA合成酶的细胞新合成的蛋白质进行标记。Dieterich团队^[69]将MmMetRS_{L274G}导入星形胶质细胞中,用于研究神经元-胶质细胞共培养的条件下,星形胶质细胞蛋白质组的动态变化;该研究团队对黑腹果蝇甲硫氨酸结合部位中的保守结合位点进行单核苷酸突变(DroMetRS_{L262G}^[70]),利用BONCAT/FUNCAT标记策略,用非天然氨基酸Anl标记黑腹果蝇的幼虫,可视化其神经发生及发育过程。Tirrel团队^[71]对线虫苯丙氨酸tRNA合成酶进行单核苷酸突变(Ce-PheRS_{T412G})时空特异性的、选择性标记启动子调控Ce-PheRS*的表达,联合使用BONCAT和SILAC技术,可以实现对活体秀丽隐杆线虫的特定类型细胞进行时空特异性标记,定量监测特定类型细胞在特定发育阶段的蛋白质表达组。此外,因为叠氮基团标记的蛋白质在紫外光的照射下能够与邻近蛋白形成共价连接,该标记策略也可以用于研究蛋白-蛋白间相互作用^[72]。2018年,Wyss-Coray团队^[73]改造了两种新型氨酰tRNA合成酶ScTyrRS_{Y43G}、MmPheRS_{T413G},首次将两种改造型合成酶同时放在一种细胞中。经实验证实,与表达单一突变型合成酶相比,同时表达两种(ScTyrRS_{Y43G}、MmPhe

RS_{T413G})具有更全面的蛋白质组覆盖度和较高的可信度。

2.2 组织特异性

2011年, Chin团队^[74]首次将遗传密码扩增技术成功应用于活体秀丽隐杆线虫, 将在大肠杆菌中进化的PylRS/tRNA_{CUA}对其衍生物掺入多种非天然氨基酸, 如将带有生物正交基团和光交联剂的氨基酸引入秀丽隐杆线虫的蛋白质中。2014年, 该团队使用SORT-M (stochastic orthogonal recoding of translation with chemoselective modification)技术将带有环丙烯基团的氨基酸类似物嵌入果蝇卵巢的生殖细胞中^[75], 利用环丙烯-四嗪反电子需求Diels-Alder环加成反应对标记的蛋白进行分离纯化、定量分析或荧光成像, 实现在活体果蝇中对特异类型的细胞在特定阶段合成的蛋白质进行标记(图3)。Diels-Alder环加成反应的速率比叠氮-炔基环加成反应高3~7个数量级^[76], 反应速率快意味着能够适用于研究快速变化的生物学过程, 对于活细胞或者组织分析的影响较小, 因而环丙烯-四嗪反电子需求Diels-Alder环加成反应逐渐得到了更广泛的应用。此外, 还可以联合使用多种生物正交反应^[77], 实现在同一个反应体系中同时对不同的蛋白质进行标记。2018年, 该团队将该技术应用于活体小鼠大脑^[78], 得以选择性

标记神经细胞或胶质细胞, 并且能够对大脑的不同区室进行选择标记, 可视化分析小鼠的神经系统。北京大学刘涛课题组^[79]系统地总结了遗传密码子扩展技术在免疫疗法、蛋白质疗法、抗体与药物相互作用等领域中的应用。

2.3 物种特异性

2014年, Tirrel团队^[80]将NLL-MetRS装载至耶尔森菌中, 以改造后的细菌侵染HeLa细胞, 鉴定耶尔森菌感染宿主细胞时其T3SS分泌系统分泌的效应蛋白的种类以及注射顺序, 利用非天然氨基酸标记策略, 实现细胞选择性富集、鉴定病原菌分泌的效应蛋白(图4)。Hang团队^[81]对大肠杆菌的苯丙氨酸tRNA合成酶的单一位点进行突变(A294G PheRS), 然后将其装配到沙门氏菌中, 在沙门氏菌感染小鼠巨噬细胞体系中选择性标记沙门氏菌的蛋白质组, 鉴别鼠伤寒沙门氏菌感染巨噬细胞过程中分泌的效应蛋白。近日, Selkirk等人^[82]利用化学蛋白质组学技术(BONCAT+pSILAC)研究鼠伤寒沙门氏菌感染小鼠巨噬细胞的生物学过程, 揭示了巨噬细胞在受到沙门氏菌感染后的不同时间段内, 其蛋白质组的时空动力学变化。

Schultz研究团队^[83]将源自产甲烷球菌的吡咯酰-tRNA合成酶/tRNA对在大肠杆菌中经过进化, 装载至

图3 (网络版彩色)SORT-M标记策略。吡咯酰tRNA合成酶(PylRS)/tRNA对能够识别多种非天然氨基酸, 因为PylRS不是通过识别tRNA携带的反密码子来氨酰化tRNA的, 所以PylRS能够活化多种PyltRNA_{xxx}, 因而PyltRNA_{xxx}能够将不同的非天然氨基酸嵌入到不同的密码子处

Figure 3 (Color online) SORT-M labeling strategy. The pyrrolysyl tRNA synthetase (PylRS)/tRNA pair can recognize a variety of unnatural amino acids. PylRS can activate a variety of PyltRNA_{xxx}, instead of recognizing the anticodon carried by tRNA to aminoacylate tRNA. Therefore, PyltRNA_{xxx} can incorporate unnatural amino acids to positions encoded by different codons

图 4 (网络版彩色)病原菌感染宿主细胞. 通过改造病原菌的tRNA合成酶, 将带有生物正交功能基团的非天然氨基酸嵌入到细菌的新生蛋白质中. 标记的蛋白通过细菌的分泌途径进入宿主细胞

Figure 4 (Color online) Pathogens infect host cells. By modifying the tRNA synthetase of pathogenic bacteria, unnatural amino acids are inserted into the newly synthesized protein of bacteria. The labeled proteins enter in the host cell through the bacterial secretory system

结核分枝杆菌中. 通过该体系将苯丙氨酸的类似物整合到结核分枝杆菌的琥珀无义密码子中, 对感染巨噬细胞的结核分枝杆菌的蛋白质表达组进行标记, 进而可以探究结核分枝杆菌通过向宿主细胞分泌了哪些效应蛋白, 以及各种效应蛋白分泌的时空顺序, 可以促进药物研发和疫苗的开发.

3 前景与展望

近些年来, 定量蛋白质组学研究技术的发展, 使得研究人员能够以前所未有的时间和空间分辨率监测蛋白质合成, 研究内在环境发生变化时, 生物体在极短时间内的应激反应. 利用现有的化学蛋白质组学技术获取生物体的蛋白质组动态变化的信息, 进而深入探

究各种生理病理现象的生物学机制, 推动定量工程生物学的发展. 促进药物研发、早期临床诊断, 成为临床疾病诊断及治疗的强有力辅助工具.

蛋白质组学的研究技术目前还存在诸多不完善之处, 新型定量蛋白质组学技术正处于研发阶段, 期待在不久的将来, 技术的发展可以实现在单细胞水平上监测生物体的生理和病理状态及其蛋白质组的动态变化; 以利于筛选疾病的生物标记物, 获取特定细胞的分泌组, 用于疾病的早期临床诊断; 为研究定量工程生物学中新改造的生物体提供方法支撑, 使得科研人员可以在系统层面上进行深度功能测试, 促进工程生物学的产品转化; 更好地推动定量工程生物学、生物医学及临床医学的发展.

参考文献

- 1 Zhang X E. Synthetic biology in China: Review and prospects (in Chinese). *Sci Sin-Vitae*, 2019, 49: 1543–1572 [张先恩. 中国合成生物学发展回顾与展望. 中国科学: 生命科学, 2019, 49: 1543–1572]
- 2 Martin V J J, Pitera D J, Withers S T, et al. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol*, 2003, 21: 796–802
- 3 Ro D K, Paradise E M, Ouellet M, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 2006, 440: 940–

943

- 4 Galanie S, Thodey K, Trenchard I J, et al. Complete biosynthesis of opioids in yeast. *Science*, 2015, 349: 1095–1100
- 5 Shao Y, Lu N, Wu Z, et al. Creating a functional single-chromosome yeast. *Nature*, 2018, 560: 331–335
- 6 Luo J, Sun X, Cormack B P, et al. Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. *Nature*, 2018, 560: 392–396
- 7 Luo X, Reiter M A, d’Espaux L, et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature*, 2019, 567: 123–126
- 8 Noren C J, Anthony-Cahill S J, Griffith M C, et al. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science*, 1989, 244: 182–188
- 9 Wang L, Brock A, Herberich B, et al. Expanding the genetic code of *Escherichia coli*. *Science*, 2001, 292: 498–500
- 10 Chin J W, Martin A B, King D S, et al. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc Natl Acad Sci USA*, 2002, 99: 11020–11024
- 11 Wang L, Zhang Z, Brock A, et al. Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc Natl Acad Sci USA*, 2003, 100: 56–61
- 12 Zhang Z, Smith B A C, Wang L, et al. A new strategy for the site-specific modification of proteins *in vivo*. *Biochemistry*, 2003, 42: 6735–6746
- 13 Chin J W, Ashton Cropp T, Anderson J C, et al. An expanded eukaryotic genetic code. *Science*, 2003, 301: 964–967
- 14 Deiters A, Cropp T A, Mukherji M, et al. Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J Am Chem Soc*, 2003, 125: 11782–11783
- 15 Liu C C, Schultz P G. Adding new chemistries to the genetic code. *Annu Rev Biochem*, 2010, 79: 413–444
- 16 Dieterich D C, Link A J, Graumann J, et al. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci USA*, 2006, 103: 9482–9487
- 17 Beatty K E, Xie F, Wang Q, et al. Selective dye-labeling of newly synthesized proteins in bacterial cells. *J Am Chem Soc*, 2005, 127: 14150–14151
- 18 Beatty K E, Liu J C, Xie F, et al. Fluorescence visualization of newly synthesized proteins in mammalian cells. *Angew Chem Int Ed*, 2006, 45: 7364–7367
- 19 Dumas A, Lercher L, Spicer C D, et al. Designing logical codon reassignment—Expanding the chemistry in biology. *Chem Sci*, 2015, 6: 50–69
- 20 Dieterich D C, Lee J J, Link A J, et al. Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat Protoc*, 2007, 2: 532–540
- 21 Prescher J A, Bertozzi C R. Chemistry in living systems. *Nat Chem Biol*, 2005, 1: 13–21
- 22 Meldal M, Tornøe C W. Cu-catalyzed azide-alkyne cycloaddition. *Chem Rev*, 2008, 108: 2952–3015
- 23 Rostovtsev V V, Green L G, Fokin V V, et al. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes. *Angew Chem Int Ed*, 2002, 41: 2596–2599
- 24 Couradeau E, Sasse J, Goudeau D, et al. Probing the active fraction of soil microbiomes using BONCAT-FACS. *Nat Commun*, 2019, 10: 2770
- 25 Wang P, Tang W, Li Z, et al. Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nat Chem Biol*, 2019, 15: 1110–1119
- 26 Bao X, Guo X, Yin M, et al. Capturing the interactome of newly transcribed RNA. *Nat Methods*, 2018, 15: 213–220
- 27 Gao L, Song Q, Liang H, et al. Legionella effector SetA as a general O-glucosyltransferase for eukaryotic proteins. *Nat Chem Biol*, 2019, 15: 213–216
- 28 Thiele C, Wunderling K, Leyendecker P. Multiplexed and single cell tracing of lipid metabolism. *Nat Methods*, 2019, 16: 1123–1130
- 29 Link A J, Tirrell D A. Cell surface labeling of *Escherichia coli* via copper(I)-catalyzed [3+2] cycloaddition. *J Am Chem Soc*, 2003, 125: 11164–11165
- 30 Tcherkezian J, Brittis P A, Thomas F, et al. Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation. *Cell*, 2010, 141: 632–644
- 31 Yoon B C, Jung H, Dwivedy A, et al. Local translation of extranuclear lamin B promotes axon maintenance. *Cell*, 2012, 148: 752–764
- 32 Hinz F I, Dieterich D C, Tirrell D A, et al. Noncanonical amino acid labeling *in vivo* to visualize and affinity purify newly synthesized proteins in larval zebrafish. *ACS Chem Neurosci*, 2012, 3: 40–49
- 33 Zhang M M, Tsou L K, Charron G, et al. Tandem fluorescence imaging of dynamic S-acylation and protein turnover. *Proc Natl Acad Sci USA*, 2010, 107: 8627–8632
- 34 Deal R B, Henikoff J G, Henikoff S. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science*, 2010, 328: 1161–1164
- 35 Dieterich D C, Hodas J J L, Gouzer G, et al. *In situ* visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci*, 2010, 13: 897–905
- 36 Hodas J J L, Nehring A, Höche N, et al. Dopaminergic modulation of the hippocampal neuropil proteome identified by bioorthogonal noncanonical

- amino acid tagging (BONCAT). *Proteomics*, 2012, 12: 2464–2476
- 37 Ong S E, Blagoev B, Kratchmarova I, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, 2002, 1: 376–386
- 38 Howden A J M, Geoghegan V, Katsch K, et al. QuaNCAT: Quantitating proteome dynamics in primary cells. *Nat Methods*, 2013, 10: 343–346
- 39 Bagert J D, Xie Y J, Sweredoski M J, et al. Quantitative, time-resolved proteomic analysis by combining bioorthogonal noncanonical amino acid tagging and pulsed stable isotope labeling by amino acids in cell culture. *Mol Cell Proteomics*, 2014, 13: 1352–1358
- 40 Gevaert K, Van Damme J, Goethals M, et al. Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis. *Mol Cell Proteomics*, 2002, 1: 896–903
- 41 Kramer G, Sprenger R R, Back J W, et al. Identification and quantitation of newly synthesized proteins in *Escherichia coli* by enrichment of azidohomoalanine-labeled peptides with diagonal chromatography. *Mol Cell Proteomics*, 2009, 8: 1599–1611
- 42 tom Dieck S, Kochen L, Hanus C, et al. Direct visualization of newly synthesized target proteins *in situ*. *Nat Methods*, 2015, 12: 411–414
- 43 Liu J, Xu Y, Stoleru D, et al. Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proc Natl Acad Sci USA*, 2012, 109: 413–418
- 44 Weintz G, Olsen J V, Frühauf K, et al. The phosphoproteome of toll-like receptor-activated macrophages. *Mol Syst Biol*, 2010, 6: 371
- 45 Sharma K, Kumar C, Keri G, et al. Quantitative analysis of kinase-proximal signaling in lipopolysaccharide-induced innate immune response. *J Proteome Res*, 2010, 9: 2539–2549
- 46 Eichelbaum K, Winter M, Berriel Diaz M, et al. Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat Biotechnol*, 2012, 30: 984–990
- 47 Eichelbaum K, Krijgsveld J. Rapid temporal dynamics of transcription, protein synthesis, and secretion during macrophage activation. *Mol Cell Proteomics*, 2014, 13: 792–810
- 48 Kuhn P H, Koroniak K, Hög S, et al. Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons. *EMBO J*, 2012, 31: 3157–3168
- 49 Witzke K E, Rosowski K, Müller C, et al. Quantitative secretome analysis of activated jurkat cells using click chemistry-based enrichment of secreted glycoproteins. *J Proteome Res*, 2017, 16: 137–146
- 50 Anderson J C, Clarke E J, Arkin A P, et al. Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol*, 2006, 355: 619–627
- 51 Swofford C A, Van Dessel N, Forbes N S. Quorum-sensing *Salmonella* selectively trigger protein expression within tumors. *Proc Natl Acad Sci USA*, 2015, 112: 3457–3462
- 52 Din M O, Danino T, Prindle A, et al. Synchronized cycles of bacterial lysis for *in vivo* delivery. *Nature*, 2016, 536: 81–85
- 53 Huh J H, Kittleson J T, Arkin A P, et al. Modular design of a synthetic payload delivery device. *ACS Synth Biol*, 2013, 2: 418–424
- 54 Shi L, Yu B, Cai C H, et al. Angiogenic inhibitors delivered by the type III secretion system of tumor-targeting *Salmonella typhimurium* safely shrink tumors in mice. *AMB Expr*, 2016, 6: 56
- 55 Xu X, Hegazy W A H, Guo L, et al. Effective cancer vaccine platform based on attenuated salmonella and a type III secretion system. *Cancer Res*, 2014, 74: 6260–6270
- 56 Chowdhury S, Castro S, Coker C, et al. Programmable bacteria induce durable tumor regression and systemic antitumor immunity. *Nat Med*, 2019, 25: 1057–1063
- 57 Johnson D B, Puzanov I, Kelley M C. Talimogene laherparepvec (T-VEC) for the treatment of advanced melanoma. *Immunotherapy*, 2015, 7: 611–619
- 58 Kohlhaup F J, Zloza A, Kaufman H L. Talimogene laherparepvec (T-VEC) as cancer immunotherapy. *Drugs Today*, 2015, 51: 549–558
- 59 Huang H, Liu Y, Liao W, et al. Oncolytic adenovirus programmed by synthetic gene circuit for cancer immunotherapy. *Nat Commun*, 2019, 10: 4801
- 60 Shao J, Xue S, Yu G, et al. Smartphone-controlled optogenetically engineered cells enable semiautomatic glucose homeostasis in diabetic mice. *Sci Transl Med*, 2017, 9: eaal2298
- 61 Yin J, Yang L, Mou L, et al. A green tea-triggered genetic control system for treating diabetes in mice and monkeys. *Sci Transl Med*, 2019, 11: eaav8826
- 62 Xu L, Yang H, Gao Y, et al. CRISPR/Cas9-mediated CCR5 ablation in human hematopoietic stem/progenitor cells confers HIV-1 resistance *in vivo*. *Mol Ther*, 2017, 25: 1782–1789
- 63 Xu L, Wang J, Liu Y, et al. CRISPR-edited stem cells in a patient with HIV and acute lymphocytic leukemia. *N Engl J Med*, 2019, 381: 1240–1247
- 64 Barrett R M, Liu H W, Jin H, et al. Cell-specific profiling of nascent proteomes using orthogonal enzyme-mediated puromycin incorporation. *ACS Chem Biol*, 2016, 11: 1532–1536
- 65 Ngo J T, Champion J A, Mahdavi A, et al. Cell-selective metabolic labeling of proteins. *Nat Chem Biol*, 2009, 5: 715–717

- 66 Tanrikulu I C, Schmitt E, Mechulam Y, et al. Discovery of *Escherichia coli* methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine *in vivo*. *Proc Natl Acad Sci USA*, 2009, 106: 15285–15290
- 67 Ngo J T, Schuman E M, Tirrell D A. Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells. *Proc Natl Acad Sci USA*, 2013, 110: 4992–4997
- 68 Mahdavi A, Hamblin G D, Jindal G A, et al. Engineered aminoacyl-tRNA synthetase for cell-selective analysis of mammalian protein synthesis. *J Am Chem Soc*, 2016, 138: 4278–4281
- 69 Müller A, Stellmacher A, Freitag C E, et al. Monitoring astrocytic proteome dynamics by cell type-specific protein labeling. *PLoS One*, 2015, 10: e0145451
- 70 Erdmann I, Marter K, Kobler O, et al. Cell-selective labelling of proteomes in *Drosophila melanogaster*. *Nat Commun*, 2015, 6: 7521
- 71 Yuet K P, Doma M K, Ngo J T, et al. Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA*, 2015, 112: 2705–2710
- 72 Fleet G, Porter R, Knowles J. Affinity labelling of antibodies with aryl nitrene as reactive group. *Nature*, 1969, 224: 511–512
- 73 Yang A C, du Bois H, Olsson N, et al. Multiple click-selective tRNA synthetases expand mammalian cell-specific proteomics. *J Am Chem Soc*, 2018, 140: 7046–7051
- 74 Greiss S, Chin J W. Expanding the genetic code of an animal. *J Am Chem Soc*, 2011, 133: 14196–14199
- 75 Elliott T S, Townsley F M, Bianco A, et al. Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat Biotechnol*, 2014, 32: 465–472
- 76 Lang K, Davis L, Wallace S, et al. Genetic encoding of bicyclononynes and *trans*-cyclooctenes for site-specific protein labeling *in vitro* and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. *J Am Chem Soc*, 2012, 134: 10317–10320
- 77 Willems L I, Li N, Florea B I, et al. Triple bioorthogonal ligation strategy for simultaneous labeling of multiple enzymatic activities. *Angew Chem Int Ed*, 2012, 51: 4431–4434
- 78 Krogager T P, Ernst R J, Elliott T S, et al. Labeling and identifying cell-specific proteomes in the mouse brain. *Nat Biotechnol*, 2018, 36: 156–159
- 79 Huang Y, Liu T. Therapeutic applications of genetic code expansion. *Synth Syst Biotechnol*, 2018, 3: 150–158
- 80 Mahdavi A, Szychowski J, Ngo J T, et al. Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci USA*, 2014, 111: 433–438
- 81 Grammel M, Dossa P D, Taylor-Salmon E, et al. Cell-selective labeling of bacterial proteomes with an orthogonal phenylalanine amino acid reporter. *Chem Commun*, 2012, 48: 1473–1474
- 82 Selkrig J, Li N, Hausmann A, et al. Spatiotemporal proteomics uncovers cathepsin-dependent macrophage cell death during Salmonella infection. *Nat Microbiol*, 2020, 5: 1119–1133
- 83 Wang F, Robbins S, Guo J, et al. Genetic incorporation of unnatural amino acids into proteins in *Mycobacterium tuberculosis*. *PLoS One*, 2010, 5: e9354

Summary for “定量工程生物学的化学蛋白质组学支撑性技术”

Chemical proteomics techniques and their perspective enabling roles in quantitative engineering biology

Lei Wang^{1,2†}, Jiandong Huang^{1†}, Shuxin Yang¹, Shuqiang Huang^{1,2*} & Nan Li^{1,2*}

¹ Guangdong Provincial Key Laboratory of Synthetic Genomics, Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China;

² University of Chinese Academy of Sciences, Beijing 100049, China

† Equally contributed to this work

* Corresponding authors, E-mail: shuqiang.huang@siat.ac.cn; nan.li@siat.ac.cn

Quantitative engineering biology, which uses tools to modify cells for various purposes, is a rapidly emerging interdisciplinary field at the frontier of current biological research. The technology lends itself to many diverse applications, ranging from biomedical uses such as vaccine development, engineered tissue and diagnostics, to industrial uses in numerous sectors including but not limited to agriculture, textiles, and bioenergy.

Engineering biologists use rational design to assemble different biological components such as genes, promoters and other regulatory elements to produce genetic circuits with a wide range of innovative functions. To develop a novel circuit, engineering biologists use an iterative strategy of Design-Build-Test-Learn-Redesign, typically cycling through several iterations and optimizations before the final goal is achieved. Once the genetic circuit has successfully been engineered, it is inserted into cells that have been synthetically simplified to only contain essential genes. These basic cells are called chassis cells. They act as the host cell for the genetic circuit, allowing for the creation of many different types of engineered systems that are able to precisely perceive and respond to a variety of environments. By employing this rational design approach, engineering biologists have successfully created vital life-saving technologies like anti-tumor bacteria and virus.

Proteomics is an essential tool in quantitative engineering biology. Proteins are the functional elements of almost all biological processes and engineered organisms are no exception. The ability to observe protein expression is necessary for understanding the working mechanism of the engineered organisms as well as the molecular interactions between the organisms and their surroundings. Quantitative proteomics provides a high-throughput way to study the spatio-temporal dynamics in the engineered organisms and their environments. However, classical proteomics lacks the ability to easily distinguish the newly synthesized proteome from the accumulated proteome, which is essential to proteome dynamics studies.

Fortunately, chemical proteomics provides new possibilities for labeling and quantifying proteome dynamics. It is quickly becoming an essential tool for the functional characterization and mechanism discovery of newly-engineered organisms. Innovative chemical biology tools, such as unnatural amino acids and bio-orthogonal chemistry, along with the emergence of high-resolution mass spectrometry, have contributed to the incredible rise in popularity of chemical proteomics. These new methods have the unique ability to reveal the spatio-temporal dynamics of the engineered organisms' proteome in complex environments. Chemical proteomics is an indispensable technique to explore how engineered cells function in their host organisms. Furthermore, chemical proteomics allows us to develop effective methods for the complex functional studies required for quantitative engineering biology research. Here, we review the potential applications of chemical proteomics in quantitative engineering biology.

quantitative engineering biology, genetic circuits, chemical proteomics, biological orthogonal chemistry

doi: [10.1360/TB-2020-0457](https://doi.org/10.1360/TB-2020-0457)