

·科学论坛·

生命科学的发展与当前的重要研究问题

江虎军* 冯 锋* 杨新泉* 王钦南* 夏家辉†

(* 国家自然科学基金委员会生命科学部,北京 100085;

† 中南大学湘雅医学院医学遗传国家重点实验室,长沙 410078)

[摘要] 生命科学的发展可大致划分为描述生命、认识生命、改造生命、创造生命、把握生命 5 个阶段。在生命科学研究的 4 个层次,即群体水平、个体水平、细胞水平、分子水平上,研究蛋白质间形成的多聚体及其蛋白质间的相互作用对阐明生命的奥秘具有重要意义。简要说明了生命活动过程中的信息流问题,以及基因组计划实施以来研究生命起源与进化中的问题。

[关键词] 信息流,生命起源,生命进化

地球约有 47 亿年的历史,据化石记载,距今约 35 亿年以前,地球上才有了生命,首先出现了原核蓝藻类;约 15 亿年以前,出现了真核藻类;约 5 亿年以前,地球上才产生了多细胞形式的生命形式;2.3 亿年以前,地球上产生了第一批古老的哺乳动物(如恐龙);5000 万年前,产生了现代哺乳动物;500 万年的时候,人和猿在进化上开始分离,各自走上了不同的进化道路。经过漫长的岁月,生命终于完成了从无到有、从单一到多样、从简单到复杂、从低级到高级的演化过程。外观看似多种多样千差万别的生命个体,然而组成生命最基本的分子却完全相同,即决定生命形式的遗传物质核苷酸(DNA 和 RNA)都是由腺嘌呤(Adenine, 缩写为 A, 以下相同)、鸟嘌呤(Guanine, G)、胞嘧啶(Cytosine, C)和胸腺嘧啶(Thymine, T)(在 RNA 中则为尿嘧啶,Uracil, U)四种分子组成,且都遵循完全相同的每 3 个核苷酸编码一个氨基酸法则,即中心法则。这四种核苷酸共能组成 64 种三联体密码,但只编码 20 种氨基酸,所有蛋白质均由这 20 种氨基酸组成,上述规律是生命界的公理。

现在分子生物学研究结果发现,不同物种间相同功能或相似功能的基因间具有同源性,进化距离越近同源性越高;功能越重要的基因保守性越强。如原癌基因 *ras* 的相关蛋白基因 *RAB1*,无论是高等的生物(如人)还是低等的生物(如细菌),其基因组

中都含有这个基因,它调节细胞器装配和运输期间膜小泡的铆定和融合,水解 GTP 变成 GDP,在细胞信号传递过程中发挥重要的作用。人(*H. sapiens*)的 *RAB1* 基因编码一个 204 个氨基酸的多肽,小鼠(*M. Musculus*)的 *RAB1* 基因编码的多肽与人的 *RAB1* 基因编码的多肽完全相同;而与少根根霉(*R. norvegicus*)、果蝇(*D. melanogaster*)、线虫(*C. elegans*)和酿酒酵母(*S. cerevisiae*)相比,同源性依然高达 99%、84%、79% 和 72%^[1],提示 *RAB1* 基因来源于同一个祖基因,故又称它们为同源基因,类似的例子在生物界比比皆是。

而且,即使在不同物种间,完成一个重要生命过程的信号通路也具有相似性。如在爪蟾(*X. laevis*)、线虫和人的 *wnt* 信号通路中,都包括了 *wnt*、*Fz*、*GSK*、*APC*、*b-catenin*、*TCF* 等主要成员(各成员在不同物种间均具有同源性)。而且信号通路的主要路径也相似,信号从 *Wnt* → *Fz* → *GSK* → *APC* → *b-catenin* → *TCF* 向下传递^[2]。爪蟾的 *wnt* 通路功能是参与背腹轴的形成;线虫的 *wnt* 通路功能是参与内胚层的形成;在高等的动物中,*wnt* 通路参与了胚胎发育中的多个问题,如脑、中胚层、肾、四肢等的发育。无论在低等生物还是在高等生物中,*wnt* 通路的功能主要与胚胎的发育有关,这说明了生物完成重要生命活动的机制也具有惊人的相似。而且同一个信号通路路径在低等生物中要比在高等生物中简

本文于 2001 年 2 月 27 日收到。

单,所体现出的功能也如此,说明高等生物的信号通路是在低等生物信号通路的基础上进化而来。

如何解释万千生命世界的这种表型多样性及其内在的统一性,一直是生物学家们的梦想。

1 生命科学发展的5个过程

纵观生命科学的发展,已经经历了由描述生命、认识生命到改造生命3个过程。科学家们一直力求在认识生命的基础上去改造生命,以求为人类带来更多的福利。过去,由于人类认识自然手段的落后,人们只能从外观简单地描述生命,因而只能很粗糙地认识生命。显微镜的发现使得人类认识到组成生命的基本单位是细胞;X晶体衍射技术的发展使得科学家们认识到组成生命遗传物质的DNA呈双螺旋结构;PCR技术的出现,使得科学家们可以从分子水平去操作生命分子DNA,因而大大加快了人类揭示生命奥秘的进程。现在生命科学的发展已能帮助科学家根据自己的意愿发展各种改造生命的手段和方法,而且效率更高针对性更强,如转基因技术、克隆技术、基因工程技术以及各种其他生物工程技术等等。今天,高度自动化的DNA测序仪使得科学家们有能力对生物的全基因组进行测序;生物芯片技术可以帮助科学家从基因组的全景角度,去了解生命过程的分子机制。

目前所发现的生物中基因组最小的是生殖道枝原体(*Mycoplasma genitalium*),共有580 070个碱基对和517个基因^[3]。通过对流感嗜血杆菌(*Haemophilus influenzae*)与生殖道枝原体两种细菌的全基因组的比较分析和对枯草芽胞杆菌(*Bacillus subtilis*)、生殖道枝原体、肺炎枝原体(*Mycoplasma pneumoniae*)等3种微生物的基因敲除(gene knock-out)实验,科学家们估计250个左右的基因数是组成一个生命个体的最少基因数^[4]。同时,科学家在实验室已实现组装DNA长度达28.5 kb的具有侵染功能的病毒^[5],这些研究为组装其他生命提供了方法和技术。这说明科学家们在朝着创造生命的方向已前进了一大步,这也是生命科学发展的一个必然。随着研究的不断深入,科学家们必将有能力创造出低级的简单的生命形式,直至高级的复杂的生命形式。但是,任由科学家的个人好奇随心所欲地创造出各种形式的生命,其后果将是令人担忧的,科学家们也认为这是在“玩上帝”,须小心才是。尽管如此,科学的发展必将完成对生命的完全解码和编码。因此,需要基于人类的法律、伦理和美学标准,去指导

人类的科学探索朝着正确的方向前进。

今后的生命科学发展必将使得人类有能力去创造生命,并把握住生命,当然这个阶段离我们仍有一定的距离。今日生命科学的研究重点依然是描述生命、认识生命和改造生命。只是现在的生命科学研究对象的侧重点发生了改变,过去的研究对象主要以模式生物为主,其目的在于认识生命的基本规律,而现在的研究对象已越来越重视与人类生活密切相关的过去难以研究的那些生物上,当然包括人自己。实际上,这些生物已成为今日生命科学研究的新的模式生物。这说明现在的生命科学的研究的发展也如同其他科学一样,在朝着复杂系统和复杂科学的方向发展。

具体说,在植物方面的研究对象主要是水稻、小麦、玉米、油料等农作物;在动物方面则是水产类和畜牧类等动物。解决的主要问题是产量、品质和营养形成的分子机制。这些问题既是重要的基础性问题,同时研究成果对生产又具有指导作用,这在以前是很难从分子水平去开展研究的。与此同时,从分子水平去整理、挖掘优异资源并加以利用,也已成为该领域中的一个重要的研究内容。

在医学方面,特别是基础医学,越来越受到人们的重视,研究的主要问题依然是关于人自身各种生理现象的分子机理,对各种疾病的诊断、预防和治疗,和人类自己如何更好地成长等一系列问题。人是高等生物,人自己无疑是了解生命奥秘的最好研究对象,而在低等生物中,由于缺少高等生物中的一些重要生理过程,研究的复杂性受到影响。国际上许多研究机构也将其原来的生物学研究称之为生物医学研究(biomedical science)。一方面,研究遗传结构相对复杂的生物可以大大丰富基础研究的内容;另一方面,所取得的研究成果可以提高人类自身的生活质量,为人类自己提供直接的健康保证,为社会服务,这种社会力量反过来可以大大促进生命科学的发展,推动社会的进步。

在微生物方面,主要研究与上述所提到的生物间的相互关系,找到并利用其中的互利关系,避免或改善其中的互害关系等。

对传统模式生物的研究依然很重视,其目的是通过研究模式生物的生命形式去认识复杂的生命形式,借鉴从模式生物的研究成果,去服务于人类对其他生物开展的研究。如许多一时无法在人身上进行的研究,可以先从模式动物开始。而且,通过物种间基因水平的同源性、功能的相似性比较去认识生命

的本质。

2 生命科学研究的4个认识层次

科学家们在认识生命的过程中先后经历了个体形态水平、细胞水平直至现在的分子水平3个认识层次。今天的生命科学正朝着两个方向在发展,一是从更宏观的角度,即从群体或群落乃至全球生物圈去认识生命,即第4个认识层次;二是从更微观的角度,如从分子或原子(纳米)去认识生命。现在的生命科学研究更多地是从分子水平去诠释生命的奥秘,即从DNA水平和蛋白质水平去揭示生命现象背后的本质。

人类基因组计划实施的目的在于破译人基因组 3×10^9 bp个核苷酸对和约3万个基因的功能。人基因组的工作草图已经完成,到2001年将完成全部基因组的精确序列图。到目前为止,已经对149种生物进行全基因组测序,其中已完成了41种生物的全基因组测序^[6](截止于2001年1月6日),而且今后越来越多的其他生物的基因组也将完成测序。在蛋白质方面,已有14 062种蛋白质的空间结构得到解析^[7](截止于2001年1月3日)。所有的这些工作意味着人类在DNA和蛋白质两个层次上开始了对生命的数字化过程,这也将是今后生命科学研究中的一项基本的活动,说明生命科学已经到了大量产生数据的时代。尽管现在的生命信息量呈爆炸性增长,但今后依然要继续挖掘、收集、整理各种数据。同时,大量数据的积累也促使生命科学走入大规模比较、归纳、综合分析阶段,使得人类不仅可以从全基因组的高度去认识一种生命,而且还可以从整个生命界的高度去认识生命的起源与演化。呈指数增长的生命信息也给从事生命科学、数学、物理学、化学、信息学等领域的科学家提供了一个巨大的舞台。他们的工作就是力图从这些基本的数据出发,去分析生命的组成及其中的秘密,并还原成整体意义上的生命。

明白了每一个基因的结构和功能其实并不意味着清楚生命过程是如何完成的。生命究竟是如何像搭积木似的垒出一个完整的生命,从ATCG到形成一个有机生命的过程是如何完成的?组成生命需要多少种功能块的协调与配合?形成一种功能块哪些基因是必不可少的?基因间是否存在排列组合规律?这种规律是什么?等等。从已有的研究看,生命的形成远比搭积木式的方式要复杂得多。

已有研究表明,在细胞内的蛋白质有不少是以

多聚体的形式存在。构成多聚体的亚基(Subunit)可能完全相同,也可能完全不同,或只有部分相同。亚基越多的多聚体所表现出的功能也越复杂,如过氧化物歧化酶(SOD1),在细胞内以二聚体形式存在,组成的两个亚基完全相同,二聚体的主要功能是清除体内氧自由基^[8]。而COP9信号传导复合体则由9个不同的亚基组成,在信号传递过程中具有多种功能,组成它的亚基不仅参与调节多种细胞信号的传递和细胞周期的调控,而且还具有蛋白激酶活性^[9]。理论上,似乎亚基多的多聚体应由亚基少的多聚体进化而来,这主要是生命在进化过程中所体现出的功能综合的表现,以提高生命过程的效率。以单体发挥功能的蛋白质在细胞的往往能形成稳定的立体结构,发挥的功能也较单一。多聚体中核心的亚基在不同物种间的保守性很高,因为它的变化将直接影响多聚体的形成及多聚体的生物功能。因此,研究多聚体对于了解生命活动具有重要意义。

另一方面,蛋白质在执行功能的过程中,常与其他蛋白质,或与DNA,或与RNA相互结合(coupling),执行功能之后便解聚(uncoupling),即在细胞内形成暂时的多聚体。这种相互作用(interaction)是蛋白质发挥功能的主要方式。细胞内如此多的蛋白质间是如何有条不紊地相互作用,已成为细胞生物学家们研究的一个重点。

蛋白质间的相互作用传递着各种生命信息,使众多基因有秩序地开启和关闭。如果说研究蛋白质的三维结构是为了了解其功能,那么研究蛋白质间的相互作用则是为了了解生命活动过程,这些研究是了解生命过程本质的一个重要环节。

3 生命活动过程中的信息流

组成生命最小单位的细胞无论在其发育成生命个体的过程中,还是在其代谢过程中,均是通过各种生命信号的调控而完成的。其中每一个生理过程,都是多个基因共同相互作用相互配合的结果。基因翻译成蛋白质并通过它实现其功能,基因的开启和关闭指令着一组组蛋白质的先后登场。实际上,生命活动的本质就在于时间和空间上一个一个功能块间相互协调所完成的一个更大更复杂功能的过程,各个信号通路间相互联系相互配合形成一个庞大的网络,共同完成整个生命过程。

精卵结合是生命的开始,第一个使人想到的问题是生命的初始信号究竟是什么?是如何激活使受精卵发育成一个成熟的生命个体的?今天的体细胞

克隆研究似乎在揭示卵细胞中的细胞质在启动已分化了的细胞核朝生命的再发育过程中具有十分重要的作用。从外观上看,受精卵的发育是细胞增殖、分化与器官形成的一个过程,实际上是基因间的相互配合完成整个生命组装的过程。在器官或组织的形成过程中基因的高度秩序性表达,其关键在于基因表达的调控信号,即调控因子的精巧调控,使得众多基因能够在适当的时候适当的地方得到适当的表达,但是科学家们依然不清楚这种生命信息在发育过程中是如何快速完成并相互配合的。

生命个体发育成熟以后,不再需要形成新的组织或器官,维持生命的正常运转是其主要问题。一方面要为已形成的组织或器官更新“零件”,即保持细胞的正常代谢(细胞分裂、分化、衰老、凋亡);另一方面,还要协调整个生命体的功能,应对本身及环境所给予的信号作出反应。实际上,细胞处在各种信号的包围之中,这些信号可以是蛋白质类信号,也可以是核酸类信号,或其他化学类和物理类信号。信号从细胞外跨膜传递到细胞内,再跨膜传递到细胞核内,激活相关基因的表达,启动细胞的分裂、分化和死亡。尽管科学家已对细胞代谢过程中的信号传递作了大量的研究,但对许多生命过程的信号传递依然有许多不清楚的地方,比如异常信号导致细胞异常代谢的信号传递。

细胞所接受的信号可以影响到细胞分裂过程中遗传物质的正确复制,导致一个或多个基因的突变、或缺失、或错误剪接,使得细胞的表型和功能发生改变。如果这种变化发生在性细胞形成期,则会因此使新生的生命个体某些性状的功能发生改变,并将此性状遗传下去;反之,如果这种变化发生在其它时期,则只会使局部细胞的表型发生改变,如肿瘤的产生,且该性状不能遗传下去。

在生命活动过程中的各种信号传递通路错综复杂。信号如何参与器官和组织的形成?如何跨膜传递?如何在细胞质中级联放大?如何进入细胞核内?如何激活相关基因表达?如何在时间和空间上实现各个信号通路间的协调?所有这些问题已是细胞生物学和发育生物学的研究重点。

4 生命的起源与进化

生命究竟是如何起源的?一直是困扰生物学家们的大问题。众所周知,理论物理学家们普遍认为,宇宙始于一个几何点,大爆炸之后演化成了今天的宇宙。但从目前的证据看,尚未找到证据或推

测出生命的起源也如此。我们可以想象,如果地球上的生命不是外星人或陨石带来的话,那么应在地球的多个地方先后产生形成了许多种有机分子。然而经过漫长的大自然的反复选择后,唯有 ATCG 四种有机分子结合在一起才能形成生命这一事实,说明生命的构成只可能由这四种核苷酸组成,且由它们构成的生命分子才能进行自我复制,并将进化过程中获得的信息累积起来遗传下去。

自我复制是生命分子开始进化的一个基本条件,而且要求尽可能使复制的分子保持一致。为了确保其遗传物质能得到准确的复制,生命个体具有 1 套校正遗传物质复制中可能出现的各种错误的机制。但是不同的物种间这种机制的功能有差异,高等生物具有比低等生物较多的纠错机制,所以高等生物遗传物质复制时的错配率较低,低等生物的错配率则较高,比如病毒的变异就是一个例子。从另一个角度看,允许遗传物质复制过程中出现错配是生命进化所必需。因此研究生命个体中的这种纠错机制对于了解生命进化具有重要意义。

从动物到植物,从低等的生物(如细菌)到高等的生物(如人),我们认为细菌和人似乎不应拥有同样一个祖先,而是有可能由不同的祖先进化而来。可以想象,在原始的生命海洋中,形成了许多包含有遗传物质的生命胞,而且通过相互间融合可以形成含有多个遗传物质团(有可能是染色体的前身)的大生命胞。只有具有了形成生命最基本组份的生命胞才能表现出生命,并开始进化。同一个生命胞在进化的过程中,可以演化成不同而又相近的后代,并形成一系统,而且某些重要功能的获得有可能使得进化轨迹表现为跳跃式进化。

想象毕竟是想象,生命究竟是如何进化的?面对这个不能在实验室重复的实验,我们所能做的的确非常有限。今天,通过大规模基因组测序,科学家们完成了一个个物种的全基因组序列图,这为研究生命的进化提供了绝好的素材。由 ATCG 四种碱基构成的基因组全序列犹如上帝写的一本“书”,经过了反复修改后便成了现在的样子。我们需要了解的是在这本“书”中,上帝在其中究竟表达了一个什么中心思想?书中各章节间的关系是什么?每一段落中的句子是如何组成的?地球上如此众多的生物,每一个物种的全基因组序列犹如上帝写的一本本“书”,用的文字都一样(ATCG)。我们需要了解的是上帝分别在何时写出了哪些“书”?“书”与“书”间的关系如何?所以破译基因组序列中的遗传信息-遗

传语言和基因组序列间的进化信息 - 构建物种的系统进化树,将有可能帮助科学家揭开生命进化的秘密。因此系统地全面地挖掘隐藏在由 ATCG 所构成的四字天书中的各种信息将是科学家们今后需要研究的一个重大任务。

感谢董尔丹博士为本文提出的宝贵意见。

参 考 文 献

- [1] <http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?>
 [2] <http://stke.sciencemag.org/cgi/cm/CMP-5533>.
 [3] Clyde A, Hutchinson III, Scott N et al. Global transposon mutagenesis

- and a minimal mycoplasma genome. *Science*, 1999, 286: 2 165—2 169.
 [4] Eugene V K. How many genes can make a cell: the minimal-gene-set concept. *Annual Reviews Genomics and Human Genetics*, 2000, 1: 99—116.
 [5] Yount B, Kristopher M C, Ralph S. Basic strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *Journal of Virology*, 2000, 74(22): 10 600—10 611.
 [6] <http://www.genome.ad.jp/dbget-bin/get-htext>.
 [7] <http://www.rcsb.org/pdb/holdings.html#holdings>
 [8] <http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=147450>
 [9] 叶鑫生,沈倍奋,汤锡芳等. 细胞调控的探索 - 细胞信号传导、细胞凋亡和基因调控. 北京:军事医学科学出版社,1999.

THE DEVELOPMENT OF LIFE SCIENCES AND ITS CURRENT KEY RESEARCH AREAS

Jiang Hujun* Feng Feng* Yang Xinquan* Wang Qinnan* Xia Jiahui†

(* Department of Life Science, Aational Natural Science Foundation of China Beijing 100085;

† National Key Laboratory of Medical Genetics, Zhongnan University, Xiangya Medical School, Changsha, 410078)

Abstract According to the development of Life Sciences, it could be divided into five stages, which are as follows: describing life, realizing life, modifying life, creating life and controlling life, and scientists undertake research works at four different levels, including population level, individual level, cellular level and molecular level. Something about signal flow, life origin and life evolution are also discussed briefly.

Key words Life science, signal flow, life origin, life evolution

·资料·信息·

我委代表团圆满完成对巴基斯坦的科学访问

应巴基斯坦科技部长的邀请,为了促进两国以及我国同巴基斯坦科学基金会的科技合作,由王乃彦副主任率领的中国国家自然科学基金委员会代表团于2001年2月26日—3月5日对友好邻国巴基斯坦进行了科学访问。

抵达伊斯兰堡的第二天,在我国驻巴基斯坦陆树林大使等的陪同下,代表团访问了巴基斯坦科技部,拜会了世界著名科学家、巴基斯坦科技部长 Atta-Ur-Rahman 博士。科技部长赞扬了中国政府对巴基斯坦做出的巨大贡献(高速公路、核电站等50多个大项目)给予了高度评价。同时就进一步开展中巴两国的科技合作表示极大的关切。王乃彦副主任代表国家自然科学基金委员会表示,将拟就双方科学家感兴趣的研究领域,与巴基斯坦科学基金会共同资助双方科学家等问题进行深入地讨论。

巴基斯坦政府非常重视信息科学、生物技术等

高科技的发展。在经济实力并不十分充裕的情况下,巴基斯坦科学基金会2001年度的经费预算仍然达到近4000万美元。与此同时,巴政府还采取措施积极吸引留学生回国、开展对外交流、争取国外援助。并以此推动科技、教育水平的提高。巴方希望利用中国科学家的经验在基因技术、基因资源、盐碱地改良与种植、塑料降解、渔业、农作物高产、农业机械、中草药资源、海洋资源利用、核物理、半导体等方面开展合作。

代表团还对巴基斯坦科学基金会(PSF)进行了访问,王乃彦院士热情邀请巴基金会主任 Khalid Mahmood Khan 博士访问中国,并表示希望尽快促进基础研究领域的科技合作。代表团此次对巴基斯坦的访问取得了圆满成功。

(国际合作局 王逸 供稿)