

DNA 双螺旋发现 50 周年

戴 闻

(中国科学院理化技术研究所 北京 100080)

2003 年 5 月 21 日中国登山队的 10 名成员先后成功地登顶珠穆朗玛峰。在人类征服珠峰 50 周年之际,这一壮举极大地鼓舞了中国人民战胜非典的勇气和信心。1953 年 5 月 29 日新西兰探险家希拉里和尼泊尔夏尔巴人诺尔盖首次将人类的足迹印在珠峰之巅,它提示人们:人类与自然相处既要以和谐为宗旨,不断丰富对自己家园的认识,又要有激情有挑战精神。

1953 年 4 月 25 日《Nature》周刊同时发表了 3 篇关于 DNA 螺旋结构的论文,作者分别是沃森和克里克、威尔金斯等、富兰克林(女)等。这 3 篇论文的发表,标志着人类在理解生命和保护生命的历程中迈出了关键的一步。

两个 50 周年在 2003 年初夏先后到来,引发了一系列的纪念活动。从 2003 年 4 月起,来自中、美、英、法、俄、德、日、韩国、印度、尼泊尔、伊朗、新西兰、瑞士、西班牙、匈牙利等国的 60 余支登山队便聚集在珠峰的南、北两侧,他们都在默默地准备着,等待权威组织机构为各支登山队安排的冲击顶峰的那一天。

4 月 14 日在北京中关村科技园区“新世纪科学论坛”的第一场报告会上,周光召院士以详实而丰富的资料向与会的数百名热情听众介绍了 DNA 双螺旋发现的前前后后。在报告的结尾部分周院士告诫青年科学工作者,科学上的每一次重大突破都是建立在许许多多历史成就基础之上的。实验是检验科学真理的惟一标准,能否造福于人类是衡量科学工作价值的重要标尺。要敢于思考,善于思考,抓住契机,一往无前,同时要坚持科学中的民主与法制,加强交叉学科之间的协作。

在 2003 年 1 月 23 日出版的《Nature》周刊上,特约编辑和主编专门组织了 18 篇文章,评述 DNA 双螺旋这一“永垂不朽”(eternal)的分子结构对历史、科学乃至文化的影响。这些文章的作者都是当今生命科学界的著名专家,本文作者阅读了其

中一部分文章,感到受益匪浅,有意向更多的读者推荐。

绵延的科学峰峦——DNA 研究大事记

1960 年 5 月中国登山队首次从北坡登上珠峰。在攀登第二台阶时,4 名队员互为人梯,最终使 3 人登顶成功。1975 年 5 月 27 日中国 9 名男女登山运动员成功登上珠峰。当年,为攀登第二台阶曾架设了 5m 长的金属梯,它至今仍旧保存完好,成了从北坡冲顶珠峰的“金梯”。

1859 年达尔文的经典著作《物种起源》正式出版。作者在书中指出,地球上的生物都是从一个共同祖先演化而来的。达尔文进化论与神创论截然对立,从而引发了生命科学的一场革命。“遗传变异[→]自然选择[→]适者生存”是达尔文思想的精髓,不过在当时的研究条件下,不可能对遗传变异的物质基础说出更多。

1865 年奥地利生物学家孟德尔公布了他关于豌豆杂交的研究结果,提出了基因遗传定律。孟德尔假定,纯种绿色豆荚豌豆具有成对遗传因子 YY,而黄色的则有因子 yy。两者杂交后的第一代必定是 Yy。令第一代在内部再杂交(即 Yy 与 Yy 自交),结果有 4 种可能,即 YY、Yy、yY、yy。因为 Y 相对于 y 是显性因子,所以 4 种可能性中有 3 种(YY、Yy、yY)将表现为绿色豆荚,仅有一种(yy)表现为黄色豆荚。孟德尔通过繁琐而极其严格的杂交程序,培育出了 500 多株第二代杂交豌豆。结果证实,绿色豆荚植株的数目近似等于黄色豆荚的 3 倍,不幸,孟德尔的发现并未引起包括达尔文在内的生物学界的重视,以至于他的论文在文献堆中静躺了 35 年。1900 年,孟德尔的基因定律被另外三位科学家重新发现,他们证明:遗传因子的分离与重新组合导致生命体性状的变异。

1928 年美国生物学家摩尔根,在果蝇遗传实验的基础上,提出了基因连锁与交换的概念,从而提示:基因在其载体上是以线形方式排列的。

表 1 DNA 双螺旋研究大事年表

1869 年	医生米歇尔(Miescher F.)发现,在伤口脓液细胞的核中有一种酸性物质,它由蛋白质和另一种化合物构成。米歇尔为后者起名 为核酸,这便是今天我们所认识的脱氧核糖核酸——DNA
1919 年	雷文讷(Levene P. A.)提出,在 DNA 中有 4 种核苷酸,它们接续排列。
1928 年	格里菲斯(Griffith F)发现,在热灭活的肺炎双球菌中有一种物质,可使携带它的活体细菌发生遗传性变化。他称这种现象为 “转化”(transformation)
1938 年	西格纳(Signer R.)等发现, DNA 的分子量可高达 50 万~ 100 万。因此,雷文讷的 4 核苷酸结构必定是一条长链
1944 年	阿瑞里(Avery O.)等从化学上辨明,格里菲斯的“转化”是基于 DNA,并且明确提出:DNA 是基因材料
1949 年	查伽夫(Chargaff E.)报告,在不同的 DNA 片断中碱基的组成是不同的。全部嘌呤(即 A 和 G)的总合大约等于全部嘧啶(即 T 和 C)的总和。并且,A 与 T 数量相等,G 与 C 的数量也相等
1949 年	凡特列里(Vendrey R.)等发现,在人体性细胞核中 DNA 的数量只及体细胞的一半
1951 年	富兰克林通过 X 射线衍射区分了两种类型(A 型和 B 型)的 DNA
1952 年	何舍衣(Hershey A.)等发现,在病毒感染细菌的过程中,进入细菌细胞的只是病毒的核酸,而没有病毒蛋白,并且病毒可在细菌 内部繁衍后代
1952 年	富兰克林获得了极为清晰的 B 型 DNA-X 射线的衍射图。该图直接了当地显示出 DNA 的螺旋结构:螺旋的轴向周期是 3.4nm, 并且在 0.34nm 处另有很强的反射。这些定量结果与我们现在对 DNA 的认识完全一致:盘绕螺旋轴的每一圈有 10 个核苷酸, 螺距是 3.36nm
1953 年	4 月 25 日《Nature》刊出了沃森和克里克等具有里程碑意义的 3 篇论文
1954 年	物理学家伽莫夫(Gamow G.)提出, DNA 对蛋白质的合成起编码作用
1955 年	本策(Benzer S.)分析了一种病毒的 DNA 精细结构,区分了沿 DNA 链的单个碱基
1957 年	克里克提出“中心法则”:生命体的遗传信息在于 DNA 中核苷酸的次序,它决定了各种蛋白质中氨基酸的排列顺序
1958 年	梅赛尔森(Meselson M.)等用放射性同位素跟踪细胞分裂过程,展示了 DNA 链的半保留复制:双螺旋在复制之前必须先解开,旧 链作为模板为两条新链的合成提供信息,同时旧链本身又成为新双螺旋的一部分
1958 年	柯恩贝克(Kornberg A.)等分离出 DNA 聚合酶,从生物化学的角度证明:这种酶在 DNA 复制中起“复印机”的作用。新链的碱基 与模板链碱基互补配对,即 A 对 T, C 对 G;两条新链合成的运行方向是相反的
1961 年	尼伦贝格(Nirenberg M.)等阐明,核苷酸的次序可以编码氨基酸,从而为破译基因编码奠定了基础
1962 年	诺贝尔医学奖授予沃森、克里克和威尔金斯。然而,曾对 DNA 双螺旋结构的发现作出决定性贡献的女科学家富兰克林,却因英 年早逝(1958 年死于癌症,年仅 37 岁)未能登上诺贝尔领奖台

DNA 双螺旋研究大事年表向我们展示的仿佛是绵延无尽的科学峰峦。尽管所讨论的仅限于 DNA 结构,但我们却看到了医生、生物学家、生物化学家、物理化学家、物理学家,以及诸多其他领域科学家的足迹。对于 1953 年沃森和克里克论文发表,当时的社会反映并不强烈。在整个英国,只有一份报纸对双螺旋的发现做出了报道。但报道很有预见性:“为了弄清这些化学‘扑克牌’是如何被‘洗牌’和配对的,足够科学家们再忙上 50 年!”

为什么双螺旋能够引起公众持续的热情

在这次攀登珠峰的纪念活动中,包括“搜狐”在内的许多企业为中国业余登山队慷慨解囊。有人贬斥说:这使神圣的行动沾上了铜臭。但是从另一个角度看,正是由于引入了商业赞助机制,奥林匹克运动才获得了前所未有的生命力。奥运与经济相互促进,受益的是广大老百姓。人类基因组计划(HGP)是由多国政府联合支持的,由于私营机构“塞莱拉公

司”(Celera)参与竞争,HGP 加快了前进的步伐,终于在 2003 年 4 月提前 2 年完成了精细图谱的绘制。据悉,北京奥组委计划让 2008 年的奥运火炬传递路线跨越珠峰。

人文精神的魅力

对于普通公众来说,他们或许对陈景润的“1+2”不甚了了,或许对霍金的宇宙学理论一知半解。但这并不妨碍他们从两位科学家的精神世界中汲取力量。

对双螺旋发现做出决定性贡献的富兰克林,也是一位学术精湛、品德高尚的科学家。当时,沃森与克里克服务于卡文迪许实验室,而威尔金斯和富兰克林则就职于伦敦皇家学院,两家机构协作致力于 DNA 结构的研究。富兰克林完全没有意识到同行之间还有竞争,但沃森与克里克的心中是清楚的。作为双螺旋的发现者,沃森与克里克凭借他们卓越的直觉和相互间的默契,对科学做出了重大贡献。但是,不可否认,他们的发现也基于富兰克林未发表

现代物理知识

的实验证据。事实上,1953年1月在沃森访问伦敦皇家学院期间,富兰克林的合作者兼上司威尔金斯,在不经意之间曾向沃森展示过那张直接证明DNA双螺旋结构的X射线衍射图。这就是说,在1953年4月25日富兰克林的论文中所发表的那张决定DNA单胞结构的衍射图,在发表之前3个月就已经被沃森获悉了。尽管沃森和克里克曾在原始论文中笼统地对威尔金斯和富兰克林表示了感谢,但沃森始终不愿意承认富兰克林的X射线衍射图对DNA双螺旋发现所起的关键性作用。

1940年,年仅19岁的富兰克林在一封给父亲的信中写道:对我来说,科学和每一天的生活不可能也不应该被分开。科学一如既往地给予我人生的解答,这是基于事实,经验和实验。今天富兰克林的天才、特质和风貌已经使她成为众多科学女性的偶像。

因简洁而美丽

沃森和克里克于1953年4月25日发表的论文,其篇幅仅仅一页,其中还包括了一张DNA结构的示意图。该图就像是在北京中关村科技园区黄庄路口矗立着的双螺旋塔那个样子:两股并行盘绕的螺旋带,代表DNA的脱氧核糖-磷酸骨架,形如盘旋向上的楼梯扶手;扶手之间的水平联线,代表配对结合的碱基(A对T,C对G),恰似一级一级的台阶。这张示意图出自擅长绘画的克里克的妻子奥黛尔(Odile Crick)之手。今天,这一图像已经成了现代科学的象征,有人称其为科学世纪的“蒙娜丽莎”。

示意图尽管十分简洁,其中包含了深层次的科学信息。对于二条并行盘绕的螺旋,作者分别用一个“上行箭头”和一个“下行箭头”标出了单股DNA的行进方向。单股DNA以核苷酸为线性排列的单元,每个核苷酸包括:一个磷酸头(专业上称之为5'端)、一个糖基尾(专业称3'端)和一个碱基(A、T、C、G之一)。从下往上看,双螺旋的一条骨架是从5'到3',另一条则从3'到5'。在沃森-克里克模型中,上述“箭头”所指示的正是从5'到3'的化学方向。

在DNA复制之前,要求两股绞合在一起的螺旋先解开,形成树杈样的“Y”形,即所谓DNA复制叉,其中上边分开的两叉代表已经解开的部分。显然,左、右支叉的顶尖是不同的,一个是5'端,另一个是3'端。DNA聚合酶,作为制造新链的“复印机”,自有一套“规矩”;即新链的合成总是以旧链为模板,新链的碱基与模板链互补,新链的生长总是先有5'端再有3'端,新链与模板链的化学走向相反。对于枝杈顶尖为3'的

那一股旧链,DNA聚合酶从模板链的顶尖开始构建新链,这条新链被称为前导链,它的化学结构发展方向正好是从5'到3'。不幸,DNA聚合酶的“规矩”不允许支叉顶尖为5'的那一股模板链被连续复制。

1970年代初,生物学家认识到,后一股的复制是从DNA复制叉的分叉点向顶尖逆向进行的,因为只有这样才能不违反DNA聚合酶的“规矩”。不过,此时新链(称为滞后链)只好一段一段地产生,它们被称为冈崎片段。最后,冈崎片段重新链接,构成完整的滞后链。尽管前导链与滞后链的合成程序不同,但合成的速率同样快——每秒500~1000碱基。

在沃森-克里克模型的示意图中,两个简单的“箭头”竟然引发出如此丰富的科学内容,这使人们不禁叹服。

与人类健康休戚相关

据信,一项有关肥胖症的基因专利价值上亿美元。这意味着,基因研究所提供的是一种高附加值产品。DNA的发现在达尔文的生物繁衍概念与现代的基因组功能概念之间架起了一座桥梁。它将基因组与生物化学、细胞生物学以及生理学结合在一起,从而一通百通,解释了长期困扰生物学家的一系列问题:突变、演化、变异、生物多样性、进化以及遗传等等。同时,DNA的发现也为研究环境与人类健康的关系打开了新的视窗。

与遗传基因有关的疾病,其来源有两种可能性,即单基因突变和多基因无序。前者(如血友病)服从孟德尔遗传定律,疾病的症状容易辨认,它的遗传规律可以从家族史予以跟踪。然而,对于多基因无序所导致的慢性病,有关研究一直进展较慢。这是因为,对于涉及多种基因的慢性疾病,没有一个特定的单基因在其中起主导作用,在相关的基因中每一突变都不是疾病发生的充分一必要条件。

进而,突变之间可能相互作用,DNA与环境以及生活方式之间也可能相互作用,结果问题变得更加复杂。肿瘤的发生在于特定的一组基因的错误表达,即基因的活性被改变了。它往往来源于环境与突变的联合作用,以至于中止了细胞新陈代谢的精确执行。每一个人都具有特定的基因组,并且每一个人都有他特定的生活经历。先天因素和后天因素在人类疾病中的作用应该被研究和界定。

基因影响人类性状的定量评估已经在全同双生子研究的基础上做出。结果表明,所有普通疾病都有它的基因基础,但基因的影响力有轻有重。对于

癌症和硬化症(宇宙学家霍金所患的即肌肉萎缩性侧面硬化症),基因的影响很弱,很少有一对双生子同患一种癌症的情况。相反,对于牛皮癣和精神分裂,环境和生活方式的影响很小,双生子同患此类疾病的可能性最大。至于其他性状或疾病,则与基因中等相关,按基因影响从强到弱排序,依次是智商、神经过敏、糖尿病和哮喘病等等。

在慢性病治疗方面,人们希望通过分析特定DNA的改变,认识个体对代谢酶药物的响应。然而,这在实践上十分困难,因为每一个人都生活在复杂的世界中,生活环境的影响有时比药物影响还要大。专家认为,近期的目标主要是针对群体,就环境和药物的作用做出预言和判断。显然,这种预言对于个体不具有同样的功效。

基因组测序的意义

你为什么要冒着生命危险去攀登世界上一座座高峰呢?一位登山探险家回答:“因为山在那里。”

你报名赴治疗“非典”一线参战,心里是如何想的?一位白衣战士回答:“因为非典正在肆虐中国。”

现在的问题是:美、英、法、德、日、中等国家为什么要下决心实施“测序人类基因组”这样一项耗资数十亿美元的庞大计划呢?简明的回答或许是:因为在每一个人的身上都有着相似程度达99.9%的DNA遗传密码。

中心法则

人体的细胞多达数万亿,每一细胞的核中都有23对染色体,遗传物质DNA长链以高度盘绕-折叠的形式存在于染色体中,单套DNA的总长度约1m,而细胞核的直径仅仅数 μm 。染色体对中的一条来自父亲,另一条来自母亲。在1~22号染色体中,成对的二条其尺寸是相同的,但性染色体(即第23号)比较特殊:女子是XX,男子是XY,X的尺寸比Y大得多。因此,人类基因组计划(HGP)实际上需测定总共24条染色体(1~22号单条,加上X和Y)中DNA链的碱基顺序。这24条染色体总共包含32亿碱基,它们在总长约1m的DNA链上各就各位。

生命的遗传信息存在于DNA中,具体表现为4种碱基A、G、T、C的排列顺序。DNA的复制保证了遗传基因可以从父母传给子女,它是生命繁衍的基础。此外,个人的成长、发育还要求各种蛋白质的不断再生。在这一过程中,信息的流动遵从“DNA \rightarrow RNA \rightarrow 蛋白质”的方向。RNA也是由4种碱基构成的,只不过将DNA中的T换成了U。信使RNA

(mRNA)承担着转录DNA信息的任务。mRNA的合成同样是以DNA链为模板,只不过碱基互补配对的原则变成了A-U和C-G,并且RNA链的长度比DNA短得多。在mRNA合成的过程中,它暂时与DNA模板链结成双链,当中止合成的信号出现,新合成的mRNA即脱离模板链。与此同时,被解开的两股DNA再度绞合在一起。

从mRNA到蛋白质的过程叫做“翻译”,具体工作要靠核糖体RNA(rRNA)和转运RNA(tRNA)的帮助,并以氨基酸为原料,但mRNA中的碱基信息与蛋白质中氨基酸的排列顺序是严格对应的。1961年,尼伦贝格(Nirenberg A.)等使用单种碱基的mRNA(即UUUUU……)合成了一种蛋白质多肽链。研究者发现,后者只包含一种氨基酸,即苯丙氨酸。对mRNA中的碱基和多肽链中的氨基酸进行计数,结果表明:碱基U的数目恰好是苯丙氨酸数目的3倍。之后,三联密码子与氨基酸之间的关系得到了确认:例如:GAA和GAG对应谷氨酸,AAA和AAG对应赖氨酸,UUU和UUC对应苯丙氨酸等。

4种碱基可以结成64种不同的三联密码子($4^3=64$),但氨基酸只有约20种,于是一种氨基酸往往对应多种三联密码子。具体说,61种三联密码子的功能是编码氨基酸,另外3种三联密码子在蛋白质合成的过程中起“标点符号”的作用——对合成中的蛋白质链给出停止生长的信号。

人类基因组计划

对于一个生命体来说,它身上的每一个细胞都包藏着同样的全套遗传信息,这就是所谓细胞的全能性。原则上,供者的体细胞核可以通过克隆技术,植入到已经去除了细胞核的卵子中,进而发育成与供者完全相同的个体。克隆羊,克隆牛便是根据这个道理。

HGP的产品是一本由32亿字符构成的天书,纵然读者有一目十行的能力,也不可能在有生之年将它读完。况且,字母只有4种,单调乏味,读完它又有什么用呢?上述看法在DNA双螺旋发现后的前20年比较流行。尽管科学界已经认识到A-T,C-G配对的意义,许多生物学家的研究兴趣仍然局限于蛋白质本身,因为后者与人类生命的联系似乎更为直接。到了19世纪70年代,当DNA分子在细胞中的行为被弄清,它在蛋白质合成中所起的信息提供者的作用被确认,DNA的测序工作终于被提到日程上来了。

随着 DNA 复制过程在实验室中的实现, 1975~1977 年桑杰(Sanger F.) 和吉尔伯特(Gilbert W.) 等首先对几百碱基长的 DNA 进行了测序。1977~1978 年, 测序一个病毒的完整基因组(约 5000 碱基) 也获得成功。1985 年聚合酶链式反应(PCR) 的发明, 使得 DNA 的片段可以扩增到百万份以上。这些进展导致了 HGP 于 1990 年启动。中国作为合作计划中惟一的发展中国家, 承担了 32 亿碱基中 1% 的测序任务。1998 年, 集成工业水平的全自动测序机研制成功, 从而大大加快了 HGP 的步伐。2003 年 4 月, 多国政府首脑同时宣布: 人类基因组的完成图已经全部绘制成功。

完成图告诉我们的, 不仅是其中清晰可见的文字, 更重要的是分子生物学家对于这些文字的解读。例如, 单条 21 号染色体中的 DNA 包含 3300 多万碱基对, HGP 的研究人员通过对测序数据的分析, 已经辨认了其中 225 个基因(包括 127 个已知基因和 98 个预测基因) 以及 59 个假基因。127 个已知基因的平均长度为 57000 碱基, 最长的基因(符号为 DSCAM) 含有 833627 碱基。基因组完成图以详细的列表给出: 基因符号、登记号、基因描述、类别、取向、起始和终止碱基对位置、基因长度和相应的基因克隆名称等。总之, 以数据库形式给出的人类基因组信息直接推动了生物学的发展, 使之成为一门“发现科学”(discovery science), 即成为一门由假说驱动、由实验和综合分析验证的科学。

另一方面, HGP 的实施也为生物技术领域带来了一场革命, 其中重要的技术突破有: 单基因的克隆, 基因的蛋白质表达, 基因在染色体中的定位, 以及用于研究 DNA—mRNA—蛋白质相互作用的生物芯片。这些进展既为研究物种进化提供了工具, 又为保护人类健康提供了危险性预测、临床诊断和新药设计的途径。此外, DNA 碱基次序中的突变因人而异, 因此对变异的探测形成了 DNA 指纹学的基础。今天, DNA 指纹学已经被广泛用于刑侦和亲子鉴定。

目前, 全自动测序机每 24 小时可以完成 150 万碱基的排序。如果只使用一台这样的机器, 完成一遍 32 亿碱基的测序约需 5 年半的时间(为了保证测序精度, HGP 实际上要对每一个 DNA 片断测量 6~10 遍)。今天, 对一个个体人完成全基因组测序耗资达 5000 万美元以上。随着测序技术的不断发展(指向微流体, 微电子以及纳米技术), 人们期待着将上述费用降低到 1 万美元以下。

对于水稻、果蝇、小鼠以及 100 多种细菌的基因组研究表明: 人类的蛋白质有 61% 与果蝇同源, 46% 与酵母菌同源; 人类第 17 号染色体上的全部基因几乎都可以在小鼠 11 号染色体上找到。各个物种间的 DNA 如此相似, 为什么会有细菌、植物和动物之分, 为什么会在形体上产生巨大差异呢? 最近的研究揭示: 就整个基因组而言, 尽管物种间的基因相似, 但它们的调控网络却是千差万别的。调控网络作为基因组重要的组成部分也是一种数字信息。调控网络包括: 启动、加强、沉寂等多种转录因子, 以及这些因子在 DNA 上的位点。这些位点通常是个体基因的控制区。正是转录因子和所处位点的联合作用决定了个别基因的开启、关闭以及表达的时间、速率、空间和幅度。

在人类与猴子之间基因组的比较表明: 三种非洲猴(猩猩、大猩猩和黑猩猩) 是人类的近亲, 其中黑猩猩的基因组与人类仅有 0.6% 的差异。这 3 种非洲猩猩先后于约 700 万年和约 500 万年前与人类分道扬镳; 而亚洲的长臂猿早在距今 1400 万年前就与人类分手了。

线粒体存在于细胞核之外, 在人类生殖细胞(即卵子和精子)形成的过程中, 线粒体 DNA 与 Y 染色体 DNA 都不参与减数分裂, 因而没有重组问题。对于这两种 DNA 进行测序和分析, 可以分别建立起展示女性和男性遗传路径的谱系图。例如, 对现代中国 1 万多名男性 Y 染色体的研究表明, 他们几乎无一例外地具有一个 DNA 突变位点——M168G; 而这一突变位点仅仅是今天一部分非洲人的遗传标记。事实上, 现代非洲人之间基因组的离散度大于亚洲人, 也大于欧洲人。有关研究的直接推论是: 在 5 万~20 万年前, 有大约 1 万个人类祖先走出非洲, 这些人分别在欧洲、亚洲、美洲和澳洲继续繁衍。无论是当年走出非洲的, 还是未走出非洲的, 他们的后代共同组成了今天 60 亿人口的世界大家庭。

DNA 中的数字信息分别在三个不同的时间尺度上发挥调控作用: 进化(数十年到数百万年)、发展(数小时到数十年)、生理学变化(数毫秒到数星期)。其中发展期所涉及的(对于人类而言)是从受精卵发展到成人的阶段; 生理学变化则涉及特定功能的触发, 如免疫反应。物种的进化主要是调控网络的进化, 而不仅是基因本身的进化。破译这类功能是分子生物学作为一门新的信息科学所面临的挑战。