编者按 超级计算是国家最核心的基础能力之一。与电子计算机相比，生物计算机在存储、计算、效能等方面超乎想象的巨大潜力，对我国未来的国家安全和基于“大数据”的经济社会发展具有十分重大的战略意义。本刊特推出专栏，从多侧面评述该领域的进展，希望其能引起读者和相关决策者的关注。

生物计算机时代即将来临*

文 / 许 进
北京大学信息科学技术学院 北京 100871

【摘 要】生物计算机是以核酸分子作为“数据”，以生物酶及生物操作作为信息处理工具的一种新颖的计算机模型。生物计算的早期构想始于1959年，诺贝尔奖获得者Feynman提出利用分子尺度研制计算机；1994年，图灵奖获得者Adleman提出基于生化反应机理的DNA计算模型；在生物计算机方面突破性工作是北京大学在2007年提出的并行型DNA计算模型，将具有61个顶点的一个3-色图的所有48个3-着色全部求解出来，其算法复杂度为3^{59}，而此搜索次数，即使是当今最快的超级电子计算机，也需要13 217年方能完成，该结果似乎预示着生物计算机时代即将来临。文章重点介绍了生物计算机的产生背景及意义；DNA计算机，特别是中州I-型DNA计算机的基本原理、计算方法与步骤；DNA计算机的研究进展，特别指出在密码分析与破译等领域的应用；分析了DNA计算机的能力，指出了研究中的难点、发展趋势，最后对我国生物计算机发展提出了一些建议。

【关键词】生物计算机，非枚举型DNA计算机，并行型DNA计算机，大规模型DNA计算机，密码分析与破译，研究进展，发展建议

DOI 10.3969/j.issn.1000-3045.2014.01.007

1 生物计算机产生背景与意义

计算工具是人类文明生活中不可缺少的工具之一。伴随着人类文明程度不断进步和发展，计算工具也随之进步与发展。人类文明时代可分为石器时代、铁器时代、蒸汽机时代、电气时代以及信息时代等阶段。在这几个阶段里，计算工具也

* 基金项目：973项目（2013CB329601,2013CB329602），国家自然科学基金重大仪器专项（61127005），国家自然科学基金项目（60974112, 30970960）
修改稿收到日期：2014年1月5日

历经了由简单到复杂、从低级到高级的不同演化过程，从“结绳记事”中的绳结、算筹、算盘、计算尺、机械计算机，直到当今的电子计算机，它们在不同的历史时期发挥了各自的历史作用。

电子计算机在其发展过程中，惊人地遵从摩尔定律[1]，为人类文明社会的发展做出了巨大贡献。但是，半个世纪以来，科学家们却一直在考虑新型计算机模型的研制，特别是2011年，在纪念图灵诞辰100周年的时候，就曾面向全世界征集超越图灵机的新型计算模型。追其原因，主要有两点：第一，电子计算机的工艺制造技术即将达到极限，如著名的理论物理学家Kaku在2012年预言，10年内电子计算机的工艺制造技术将达到极限；第二，由于图灵机模型所致，电子计算机一直不能处理规模较大的NP-完全问题。

在探索非传统的新型计算机模型研究中，相继提出了仿生计算（人工神经网络、进化计算、PSO计算等）、光计算、量子计算及生物计算等。而目前所有的仿生计算均依靠电子计算机来实现；光计算的计算模型就是图灵机模型，但实现的材料是光器件[2-4]，因此，很难超越当今的电子计算机；量子计算在处理NP-完全问题时的最好结果是：若在图灵机下算法复杂度是 n ，则量子计算可将复杂度降低为 $\sqrt{n}$ [5,6]。这就是说：量子计算模型实际上尚未超越图灵机模型。

生物计算是指以生物大分子作为“数据”的计算模型，主要分为3种类型：蛋白质计算、RNA计算和DNA计算。蛋白质计算模型的研究始于20世纪80年代中期，Conrad首先提出用蛋白质作为计算器件的生物计算模型[7]。1995年，Birge发现细菌视紫红质蛋白分子具有良好的“二态性”，拟设计、制造一种蛋白质计算机[8]。进而，Birge的同事，Syracuse大学的其他研究人员应用原型蛋白质制备出一种光电器件，它存贮信息的能力比目前电子计算机的存贮器高300倍，这种器件含细菌视紫红质蛋白，利用激光束进行信息写入和读取[9]。该蛋白质计算模型均是利用蛋白质的二态性来研制模拟图灵机意义下的计算模型，应属于纳米计算机“家族”的一员。

不同于蛋白质计算，RNA计算与DNA计算是利用生化反应，更确切地讲，是以核酸分子间的特异性杂交为机理的计算模型。由于RNA分子不仅在实验操作上没有DNA分子容易，而且在分子结构上也不如DNA分子处理信息方便，故目前对RNA计算的研究相对较少，有兴趣的读者可参见文献[10]和[11]。所以，近20年来，蛋白质计算与RNA计算少有进展，但DNA计算发展很快。故本文只介绍DNA计算与DNA计算机。

DNA计算是一种以DNA分子与相关的生物酶等作为基本材料，以生化反应作为信息处理基本过程的一种计算模式。DNA计算模型首先由Adleman博士于1994年提出[12]，它的最大优点是充分利用了DNA分子具有海量存储的能力，以及生化反应的海量并行性。因而，以DNA计算模型为基础而产生的DNA计算机，必有海量的存储能力及惊人的运行速度。DNA计算机模型克服了电子计算机存储量小与运算速度慢这两个严重的不足，具有如下4个优点：（1）DNA作为信息的载体，其贮存的容量巨大，1立方米的DNA溶液可存储1万亿亿的二进制数据，远远超过当前全球所有电子计算机的总储存量；（2）具有高度的并行性，运算速度快，一台DNA计算机在一周的运算量相当于所有电子计算机问世以来的总运算量；（3）DNA计算机所消耗的能量只占一台电子计算机完成同样计算所消耗的能量的十

亿分之一；(4)合成的DNA分子具有一定的生物活性，特别是分子氢键之间的引力仍存在。这就确保DNA分子之间的特异性杂交功能。

由此可见，DNA计算的每项突破性进展，必将给人类社会的发展带来不可估量的贡献。

第一，DNA计算机的研究在国防领域具有极为重要的意义。由于DNA计算的巨大并行性所导致的惊人速度，使得目前的密码系统对于DNA计算机而言已经失去意义。这就意味着，哪个国家在DNA计算机的研制中首先取得成功，这个国家在军事信息领域必将占据领先地位；

第二，DNA计算机的研制对理论科学的研究具有无法估量的意义，特别是针对数学、运筹学与计算机科学。这是因为，在理论研究中，许许多多的困难问题在DNA计算机的面前可能显得非常简单，如著名数学家Erdös认为人类要解决Ramsey数R(5,5)、R(6,6)是非常困难的。然而，若用DNA计算机，该问题将会很容易得到解决；

第三，DNA计算机必将极大地促使非线性科学、信息科学、生命科学等的飞速发展，进而推动诸如图像处理、雷达信号处理等巨大的发展；蛋白质优化结构的更深层认识乃至第二遗传密码的解决、天气预报更准确乃至整个气象科学的巨大发展等；也必将促使诸如量子科学、纳米科学等的巨大发展。

正是由于DNA计算机的上述重要意义，使得目前国际上关于DNA计算机的研究形成了一个新的科学前沿热点，正在极大地吸引着不同学科、不同领域的众多科学家，特别是生物工程、计算机科学、数学、物理、化学、激光技术以及信息等领域的科学家。2009年美国基金会启动超越摩尔定律的资助项目，主要用于资助DNA计算与量子计算。

2 DNA计算与DNA计算机的基本原理

DNA计算是以DNA分子作为信息处理的“数据”，相应的生物酶或生化操作作为信息处理“工具”的一种新型计算模型。基于DNA计算模型研制的DNA计算机，与电子计算机在硬件、原理等方面均不相同。DNA计算模型的一般原理图，可简要地通过图1所示的框图来描述：输入的是DNA片断和一些生物酶以及所需要的试剂等，然后通过可控的生化反应，输出的是DNA片断，这些DNA片断就是所需问题的解。

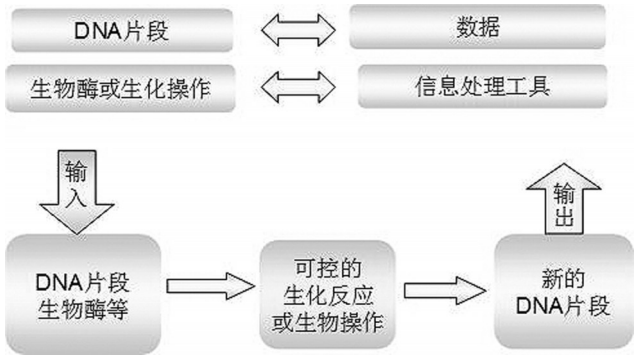

图1 DNA计算模型一般原理图

关于DNA计算机原理，我们以中州型DNA计算机模型为例给予说明。北京大学生物计算机研究组通过十几年的研究，建立了名为《中州I-型DNA计算机》硬件体系结构[13]，主要由4部分构成：存储系统、检测系统、运算系统及控制系统等，其中存储系统中含有“数据子库”及相应的探针子库；这4个系统之间的逻辑关系如图2所示。

中州I-型DNA计算机具有一定的通用性，可用于密码分析及大规模NP-完全问题等的求解。用该DNA计算机求解，与电子计算机唯一类似的是，将所给问题映射到DNA计算机模式上去。

DNA计算与DNA计算机的研究已有20年的

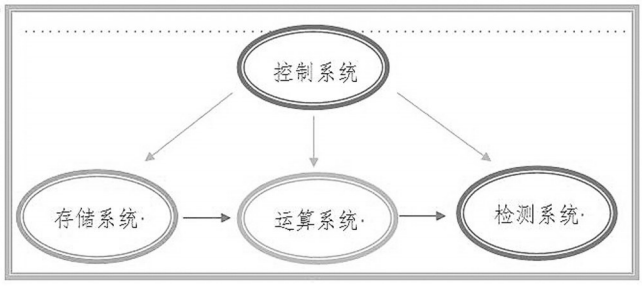

图2 中州I-型DNA计算机结构示意图

历史，目前虽仍处于实验室阶段，但已能解决一些复杂的困难问题[14]。下面，我们首先给出一般在实验室利用DNA计算求解问题的具体方法步骤，然后给出中州I-型DNA计算机求解问题的步骤。

一般用DNA计算模型求解步骤如下：

第一步 模型选择：针对问题，选择或建立DNA计算模型；

第二步 编程：在已有模型的基础上，进行编程；

第三步 编码：在DNA链的条数确定后，依据具体问题，建立相应约束条件（如解链温度值的约束，特异性杂交的约束，特别是要求链尽可能短等约束），进而进行编码；

第四步 合成DNA分子：对通过编码确定的DNA链，进行合成，并购置所需的生物酶以及相关试剂等；

第五步 建立计算平台：建立适应于生化反应、特异性杂交、无污染的良好生化操作环境；

第六步 实施计算：将所需DNA链、探针以及相关试剂等按照生化处理程序进行；

第七步 解的检测：通过PCR、测序、电镜，甚至光电等综合技术检测出所需要的解。

中州I-型DNA计算机的计算模型已固定，且需要的DNA链已合成，计算平台等均已建立，故只需要如下3步：

第一步 编程：针对问题直接进行编程；

第二步 实施计算：将通过编程后确定的DNA链、探针以及所需试剂等注入运算系统，进行运算；

第三步 解的检测：利用计算机中的检测系统，将所需问题的解全部检测出来，并输出。

限于篇幅，在此不赘述中州I-型DNA计算机的体系结构以及DNA计算的基本原理，相关内容可参见文献[15-22]。

3 DNA计算机研究进展

DNA计算机的研究可分为两大方面：（1）用于纳米机器人的研制。这方面的主要工作是充分利用DNA分子之间的特异性杂交开展的自组装技术。其研究成果重点应用于诸如疾病诊断治疗的自动化问题、癌细胞的消除等。如在2004年，以色列科学家在理论与实验上均证明了：DNA计算机是进行疾病诊断治疗的新有力手段[23-24]；（2）用于信息处理的计算机研制。主要研究快速实用化的、至少在某些方面超越电子计算机的新型计算机。

从1995年起，由美国发起的生物计算机国际会议每年一届，一直延续至今。2006年，由中国、美国、日本以及一些欧洲国家发起了一个规模更大的国际生物计算机会议。该会议每年一届，已经召开了8届。其中第四届大会由北京大学承办，许进教授5次任生物计算机国际大会主席。此外，在生物计算机方面发表的学术论文数逐年呈指数上升，而且国际上已出版了多部生物计算机方面的学术专著。

经过20年的研究，DNA计算机无论在理论方面，还是在硬件研制方面，都取得了极大进展。特别是在2007年，我国成功地建立了搜索次数可达3^{59}的并行型DNA计算机模型[14]。而且，近两年利用自组装技术，由DNA分子构成的几种重要结构实现的成功，也极大地缩短了DNA计算机走向实用的周期。这些都表明一个新型的信息处理工具——生物计算机的时代即将来临！

目前关于DNA计算与DNA计算机方面的研究内容很多，其研究方向主要涉及诸如模型构建、编码、检测、控制技术等方面。另外，在诸如密码分析与破译，困难NP-完全问题求解上均有突破性的工作。下面，对

国内外研究现状给予简要介绍。

3.1 DNA计算中模型构建研究进展

1959年，诺贝尔奖获得者Feynman提出了分子计算的构想[25]；1973年，Bennett也曾设想制造一种用酶来催化的图灵机，限于当时技术发展水平，该思想并未引起人们的重视[26]。20世纪中期以来，由于生命科学中的几个重大突破工作，为DNA计算的诞生奠定了基础。于是，在1994年，Adleman建立了DNA计算模型[12]。从1994年至今，已建立了不少的DNA计算模型。

（1）基于DNA分子机理的计算模型主要有粘贴DNA计算模型、发夹DNA计算模型、质粒DNA计算模型和 k -臂DNA计算模型。其中粘贴模型由Roweis等人提出[27]，该模型采用单、双链混合型DNA分子进行编码，此类混合型的DNA序列唯一对应于一个0-1序列，具有一定的通用性。发夹DNA计算模型是Sakatomo等人巧妙地利用发夹DNA分子结构建立的一种用于求解可满足性问题的DNA计算模型[28]。质粒DNA计算模型是利用质粒分子上的DNA序列，通过核酸内切酶与连接酶的作用，可产生两种状态的DNA计算模型。每个质粒体与传统计算机中 k – 位数据寄存器作用相同。 k – 臂DNA计算模型是由Jonoska等人建立的利用3-臂和4-臂DNA分子稳定性进行优化计算的DNA计算模型[29]。

（2）基于生化操作机理的DNA计算模型主要有剪接系统模型和自组装模型，其中剪接系统模型是当前DNA计算研究模型中的另一个主要模型。该模型是在基因工程中酶切和酶连这两个基本操作的基础上，用形式语言抽象出来的一种模型。Benenson等人在2001年所给出的一台可编程的有穷自动机及在2004年提出“用于基因表达逻辑控制的分子自动机”模型[23]，均含剪接DNA计算模型中的思想。自组装计算模型是由Winfree等首先提出来的一种DNA计算模型[30]，通过DNA分子键的相互作用形成特定的构型来完成计算过程。在该模型中，人工合成单链DNA作为DNA杂交分子（“瓦片”）。这些DNA瓦片有粘性末端，它可优先与其他DNA瓦片的粘性末端匹配，促使进一步组装成瓦格。这是一个很有前途的模型，目前在生物计算中得到了良好的应用。

（3）从实验手段的角度，DNA计算模型可分为试管型、表面型和芯片型3种。试管阶段的主要任务是研究DNA计算基本原理的可行性问题；表面阶段的主要任务是走向实用化的过渡阶段；芯片阶段是DNA计算机研制的最终阶段，标志着DNA计算机研制趋于成熟。表面与芯片DNA计算模型是近10年备受关注的模型。在表面模型中，DNA链被固定在硅片或玻璃等经过严格化学处理后的表面上，而不是将DNA分子漂浮在溶液中。表面计算简化了对DNA链的操作，并减少了DNA链的丢失，使DNA计算朝芯片型和实用型方向发展。

（4）从应用角度，DNA计算模型可分为快速信息处理的计算机模型及纳米机器人的自组装DNA计算模型（或称DNA图灵机模型）。其中，前者计算能力有望超越电子计算机，用于求解电子计算机解决不了的困难组合优化问题；后者实际上是构建纳米级的机器人，用于疾病诊疗等。DNA图灵机型由Rothemund引入[31]，他定义了一个分子计算系统，该计算系统总处于一定的状态，具有一定的存储能力和执行无限数量的状态转换功能。2001年，Shapiro等人深化和实现了Rothemund提出的图灵机型DNA计算机模型，提出了一种可编程的有限自动机[23-24]。布尔DNA计算机模型是利用DNA分子的特性来模拟布尔运算的一种DNA计算机模型。目前这方面的研究甚多。Suyama等人提出基因表达分析的DNA计算机模型[32]。

（5）从实用化通用性的角度，DNA计算模型主要有并行型DNA计算模型和大规模型DNA计算机模型。北京大学生物计算机研究组在2007年建立了并行型DNA计算模型，用于求解图的顶点着色问题。2009年至今，在并行型模型的基础

上，进一步研制了大规模DAN计算机模型，以及实用化的中州型DNA计算机模型。

3.2 DNA计算中关于编码问题的研究进展

在DNA计算中，信息（如一个顶点或边，或者布尔公式中的一个变元的赋值等）总是通过特定的DNA序列来表示。由于DNA计算中作为“数据”的DNA序列在“运算过程”中是通过特异性杂交来实现的，因此，DNA计算中的编码要受到诸如解链温度、DNA序列的相似度等众多条件的约束，以及其规模、链的长度、实验环境等诸多因素的影响。因此，编码问题是一个非常复杂、困难的问题。许多学者视DNA序列中的编码问题为NP-完全问题。

最早认识到编码重要性的是Baum，为降低DNA序列间的相似度，他提出所有编码序列及其补序列间的最大相同子序列的长度应该小于某一特定长度[27]。其后，数学、热力学、各种智能算法、图论方法等方面展开了对DNA序列编码问题的研究。此领域有一些杰出的工作，发表了许多高质量的学术论文，其中具有代表性的有：Garzon提出了H-准则，研究了DNA序列的移位Hamming距离问题[33]；Feldkamp等引入了序列间相似度的新方法；Suyama等在提出了一种经过标准正交化的DNA编码数的方法；Deaton、Arita等将遗传算法应用于DNA计算中的编码搜索和优化；Frutos、Arita等建立模板编码方法；许进研究组提出了模板框方法；Hartemink等建立了一个基于热力学DNA杂交反应的模拟软件；Rose等提出了基于统计热力学的方法来研究大量DNA分子进行杂交的统计特性，并且采用统计物理学中的分割方程来量化DNA分子间的各种可能的杂交方式；Deaton首次将信息论中熵的概念引入DNA计算编码，以对这种基于杂交反应的信息处理过程中的“信道容量”的影响因素进行定性的研究；Garzon提出了利用超立方体方法对编码进行研究；Fumiaki Tanaka等人给出了设计具有统一解链温度的多重核酸序列的一种新的算法，并且在基于最小自由能量的情况下，不与非特定的物质进行杂交。相关方面的详细论述见文献[34]。

3.3 DNA计算中解的检测与生物操作研究进展

DNA计算中解的检测问题是当前DNA计算研究中最困难的问题之一，也是关键问题之一。目前在DNA计算研究中，检测问题一直是困惑DNA计算发展的一个障碍，基本上是采用常规的生物检测方法。这些方法主要有：电泳技术、毛细管电泳技术、PCR扩增、层析技术、磁珠分离、AcryditeTM技术、DNA传感技术、荧光标记方法、分子信标方法、测序技术以及生物酶技术等。在已有的研究中，大多是采用这些检测技术把所需问题的解检测出来。

从DNA计算机的基本检测方法来看，它的发展主要是依赖于基因工程检测方法的发展而发展。这显然阻碍了DNA计算机的发展。北京大学生物计算机研究组在关于检测问题的研究中提出了：将DNA计算机解的检测问题、编码问题、“解空间大小”等问题结合起来综合处理，并将诸如生物酶、DNA分子的发夹构型以及荧光标记，特别是分子信标技术和DNA库构建技术等有机地结合起来的方法。

3.4 微流控制系统研究进展

北京大学生物计算机研究组意在将存储系统、运算系统、检测系统与控制系统集合为一个整体，逐步形成一个真正意义上的中州型DNA计算机。而近20年发展起来的微流控制系统为实现此目标起到关键作用。微流控制系统大体包括3个部分：芯

片；分析仪，包括驱动源和信号检测装置；包含有实现芯片功能化方法和试剂盒。微流控制系统具有高度集成、分析速度极快、高通量、能耗低、物耗少、污染小、廉价、安全等特点。目前已广泛应用于有关DNA分析（如涉及遗传学诊断、法医学基因分型和测序等方面）、蛋白质分析、临床血细胞分析、药物分析、小分子分析以及化学有机合成和分析化学等。

微流控制系统可将生物和化学等领域中所涉及的样品制备、生物与化学反应、分离检测等基本操作单位集合在一块几平方厘米的芯片上，用以完成不同的生物或化学反应过程，并对其结果进行检测。计算机芯片使计算微型化，而微流控制系统使生化实验室微型化，因此，在生物医学领域其可使珍贵的生物样品和试剂消耗降低到微升甚至纳升级，而且分析速度迅速提高，成本急剧下降。

微流控制系统源于1990年Manz与Widmer提出的微全分析系统（TAS）。他们最初的想法是发展一种可以作为一个化学分析所需的全部部件和操作集成在一起的微型器件，强调“微”与“全”。所以将TAS看作是化学分析仪器的微型化。1993年Harrison和Manz等人在平板微芯片上实现了毛细管电泳与流动注射分析，借电渗流实现了混合荧光染料样品注入和成功电泳分离。但直到1997年这段时间里该领域的发展前景并不十分明朗。1994年始，美国橡树岭国家实验室Ramsey在Manz的工作基础上发表了一系列论文，改进了芯片毛细管电泳的进样方法，提高了其性能与实用性，引起了更广泛的关注。在此形势之下，第一届Lab-on-a-chip or μTAS国际会议在荷兰Enchede举行，起到了推广微全分析系统的作用。1995年美国加州大学的Mathies等在微流控制系统芯片上实现了DNA等速测序。微流控制系统芯片的商业开发价值开始显现，而此时微阵列型的生物芯片已进入实质性的商品开发阶段。同年9月，首家微流控制系统芯片企业Caliper Technologies公司在美国成立。1996年Mathies又将基因分析中有重要意义的聚合酶链反应（PCR）扩增与毛细管电泳集成在一起，展示了微流控制系统在生物医学研究方面的巨大潜力。目前全世界已至少有30多个重要的实验室（包括MIT、Stanford大学、加州大学伯克利分校、美国橡树岭国家实验室等）在从事这一领域的开发和研究。

近年来，国内有多家大学和研究所的实验室开始了这方面的研究。多数是从毛细管电泳或流动注射分析所得到的技术积累转移至芯片平台上进行研究。

微流控制系统具有如下几个功能：

（1）将泵、阀、管道、反应器等集于一体，呈高度集成化。该方面最具代表性的工作是美国Quake研究小组将3 000多个微阀、1 000个微反应器以及1 000多条微通道集成在尺寸仅有几十个平方毫米的硅质材料上，完成了液体在内部的定向流动与分配；

（2）由于不同样品分离检测的需要，分离通道表面呈现出多样性发展。用磺化、硝化、胺化及将带双官能团的化合物耦合到表面的胺基上的办法加以修饰可获得各种分子组分的表面；用EDA、PDA、CAB、SPH及有机硅烷和无机氧化物等加以修饰微通道表面，用改善吸附特性，改变疏水性和控制电动力学效应来提高分离效率；

（3）微流控制系统的检测技术朝着多元化发展。目前最常用的检测器是荧光和电化学检测器；

（4）微流控制系统已从主要应用的生命科学领域扩展到其他领域。如DNA、RNA、蛋白质等方面分析检测以及许多化学合成反应的研究。

3.5 实用化DNA计算模型的研究现状

从理论上讲，所有NP-完全问题是等价的。这里重点以图顶点着色这个NP-完全问题为例展开阐述。由于图顶点着色DNA计算模型也是2007年DNA计算研究中的突破点，故在此对图顶点着色DNA计算模型的研究进展给予专门介绍。

北京大学生物计算机研究组早期在图顶点着色DNA计算模型的研究上主要是理论模型。2005年，展开与生物实验紧密结合的研究，给出了“电泳型图顶点着色DNA计算机模型”；2006建立了“非枚举型图顶点着色DNA计算模型”。该模型是国际上首次建立的非枚举型DNA计算模型，并对12个顶点的图进行了实验。此模型所需初始解空间有283条DNA链，而按照枚举型DNA计算模型所需的DNA链为3^{12}=531 441条，因此，非枚举型DNA计算模型所需的DNA分子仅占原枚举型所需DNA分子的0.0532%。此模型为DNA计算机走向实用化克服了一个关键的难点[35]。在非枚举型DNA计算模型的基础上，项目组在2007年建立了另一种新的模型——“并行型DNA计算模型”[14]。在该模型的研制过程中，4位博士生在实验室花费了整整2个月时间，成功地解决了图3中所示61个顶点图的3-着色问题（共48个解）。由于该算法对任意61（甚至大于61）个顶点的图是同样的操作方法，这就意味着，该计算模型的搜索能力可达到3^{59}（因为其中顶点1和顶点16的颜色预先确定）。相比之下，对于61个顶点的图的3着色问题，用目前最快的计算机遍历搜索3^{59}次方也需耗费13 217年以上。上述所言的并行型图顶点着色DNA计算模型，在实验过程中，虽然采用了多种并行性，如实验员的并行性、生物操作自身的并行性，但所消耗的实验时间仍然很长，且试管量很大！因此，其实用性差。但从该模型业已看到DNA计算的能力在逐渐走向实用化。针对如何克服上述不足，建立可实用的DNA计算机模型，我们又提出了大规模型DNA计算机模型。近几年，在关键技术的研究上做了大量的实验性论证工作，其研究成果均发表在国际权威杂志上，如*Applied Physics Letter*，*Analytical Chemistry*等，影响因子超过3的6篇以上。鉴于篇幅，这里不详细讨论，详见文献[36-41]，其中图4展示了我们发表在*Analytical Chemistry*、*Applied Physics Letter*及*JCIS*上的成果。

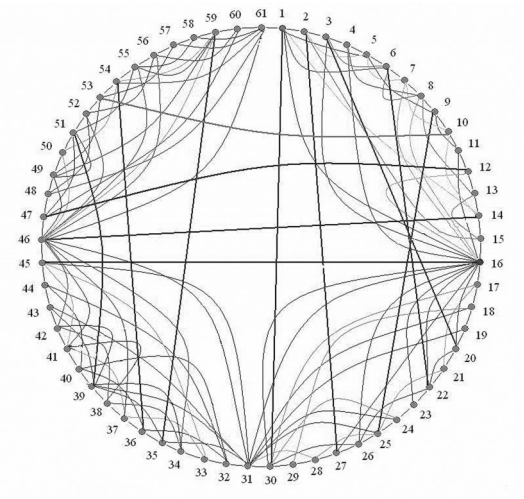

图3 并行型图顶点着色DNA计算模型实验图

3.6 基于DNA计算的密码学发展及现状

现代密码学建立在密钥搜索的时间复杂性上，即随着密钥长度的增长所需的搜索时间呈指数增长。而DNA计算的高度并行性、密集储存信息能力等特性，使其“天然”地适用于密码学领域。DNA计算求解图顶点着色等NP完全问题的意义在于对传统的基于计算安全的密码体系提出了挑战。因此，许多领域的科学家一直在关注着DNA计算机的发展，一旦DNA计算机的研究有突破性的工作，就意味着密码学领域将会受到“巨大的影响”。所以，DNA计算方法很快就融入到了密码系统的研究中。DNA计算在安全领域中的应用研究主要包括基于DNA计算的破译、DNA计算加密解密、基于DNA计算的信息隐藏及DNA认证4个方面。主要工作包括：(1)基于DNA计算的密码破译方法。Lipton等人最早提出了破译DES的DNA计算模型，所采用的是明文-密

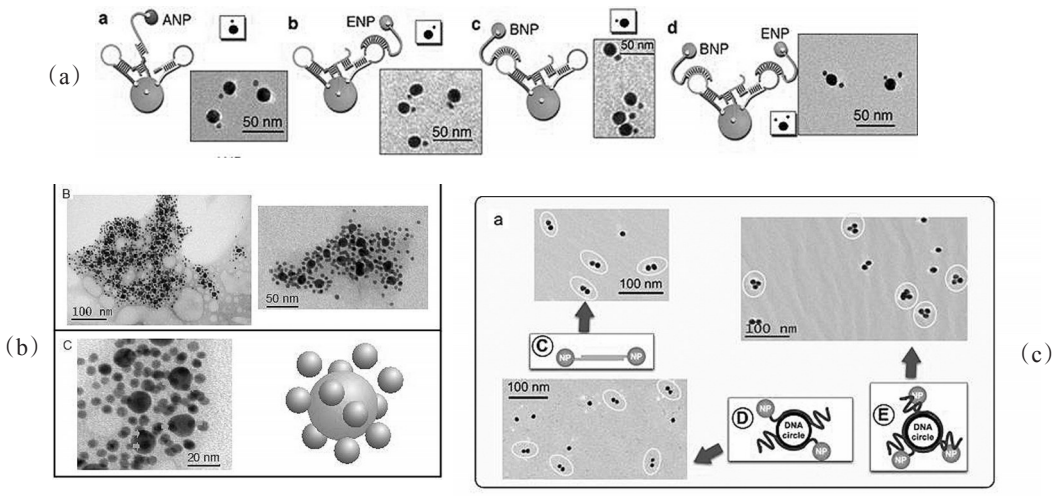

图4 2013年发表在*Analytical Chemistry*(a)、*Applied Physics Letter*(b)及*JCIS*(c)上的成果图示

文对破译法。其思想在理论上比较简单，所使用的生化操作主要有提取分离和粘贴等方法，具体生化操作方法在现实实验中不易实现；其后，Adleman等人又给出了使用粘贴模型破译DES的方法，这种模型主要使用两种基本的DNA分子记忆链和粘贴链来进行计算。他的方法仍然是明文-密文对破译法；Weng-Long Chang等人建立了一种破解RSA的整数分解DNA计算方法。但该方法目前仅仅是理论上的模型，对整数分解不能造成实质上的威胁；Burn提出使用自组装模型，通过组合加法和乘法，设计了大数分解系统来破译RSA。(2)基于DNA计算的加密解密体制，Ashish Gehani等人提出了一种基于一次性密码本的DNA加密和解密方法。他们设计了两种DNA序列的一次一密加密方法：一种是映射替代法，根据定义的映射表将固定长度的DNA明文序列单元替换成对应的DNA密文序列；另一种是异或法，采用生物分子技术进行DNA明文序列与密码本序列的异或操作。利用这两种方式实现的一次一密加密机制是具有绝对安全性的。Gehani将DNA计算引入到非对称加密机制中。他们提出利用DNA超强的并行计算能力以及其无与伦比的信息存储容量，采用比通常加密算法具有更高复杂度的算法来提高密码系统的强度；Andre Leier等人给出了使用DNA二元串进行加密和解密的方法；Jie Chen等提出基于DNA计算的分子密码设计。北京大学生物计算机研究组提出利用Y-junction carbon nanotubes（Y-CNT）探针进行DNA编码数据的读写，这样的纳米设备可以实现原子范围的访问和设置，描述了利用DNA引物扩增反应进行的二进制数的模2加法运算，并且利用DNA计算的并行性，实现了one-time-pads的加密解密。

我国在密码分析与破译方面的研究有很多，本专栏中有两篇评述文章，其中一篇是陈智华等人的工作。

2007年并行图顶点着色DNA计算模型的研究成功，表明DNA计算已经可以解决大规模模型一定范围内的NP-完全问题，这也预示着应用于破译密码DNA计算机模型研制成功的时机已经来临。这将威胁到现有的传统电子计算机加密安全。该领域的研究和开发对密码分析和密码安全具有巨大的意义。

4 展望与建议

在基于硅技术的高性能计算机发展方面，我国基础技术远远落后西方，而且历来受到发达国家的禁运和封锁。而DNA计算机研究领域，我国目前已经领先于世界。但是，由于发达国家在生

物技术方面的优势，其研究总体实力强于我国。若我国不注重生物计算机的研究，将来肯定又会落后于西方国家。为此，本小节中，作者提出成立生物计算机研究院的建议。另外，针对DNA计算机研制过程中遇到的两个难点，分离与检测方面的困难以及如何建立适应于DNA分子性能与当今分子生物技术的数学计算模型，提出如下3点建议：

（1）精准PCR仪的研制。PCR技术是1985年Kary Mullis在Cetus公司工作期间发明的一种用于放大扩增特定的DNA片段，可看作生物体外的特殊DNA复制。随后科学家很快发现，PCR技术不仅可扩增DNA片段，而且可提炼出目的DNA片段。但遗憾的是，PCR技术只能提炼出一个DNA计算平台中70%左右的目的DNA片段。如何100%将所需要的DNA片段从一个计算平台分离出来，是当前阻碍DNA计算机发展过程中最困难的一个“障碍”。如果此障碍被消除，DNA计算机必将迎来一个质的飞跃，即使是在实验室，其信息处理能力也将远远超越当今的电子计算机。因此，怎样研制出一个精准的PCR技术，进而形成精准PCR仪对DNA计算机商用发展至关重要；

（2）探针机。适应于DNA分子、已有生物酶、生化操作与仪器技术的数学计算模型的研究。二战期间的1940—1944年，美国急需用电子技术研制出一种计算机，用于密码分析与破译。当时已有几种类型的计算机，如1938年德国科学家朱斯成功制造了第一台二进制Z-1型计算机；1941年，朱斯研制出世界上第一台通用程序控制机电式计算机Z-3型计算机；1944年，美国科学家艾肯成功研制了一台机电式计算机MARK-I；1947年，艾肯研制出速度更快的机电式计算机MARK-II；1949年，艾肯研制出采用电子管的计算机MARK-III。但是都不够理想。1944年美国宾夕法尼亚大学邀请杰出的数学家冯诺依曼，在当时已有电子二极管、三极管的基础上，利用图灵机这个通用性的数学模型，建立了当今的电子计算机。自然，为研制以DNA分子与生物酶等为材料的DNA计算机，如何建立与之相适应的数学模型是当前DNA计算机研究中的另一个难点。我们已经找到了较为适应的数学模型，称为探针机；

（3）成立生物计算机研究院。建议在中科院成立生物计算机研究院，若该研究院成立，将是世界上第一个正规的生物计算机院。该院的主要职责是：研发实用化的生物计算机，这需要至少6个方面的专业科学家：

信号处理专家。主要任务是将生物信号转换成光信号，再将光信号转换成电信号，或者直接将生物信号转换成电信号，然后通过已有的电子计算机显示出最终结果；

控制工程专家。生物计算机中，其控制系统相当重要，从存储系统到运算系统、从编程到电子计算机的指令控制、从运算系统到检测系统，从检测系统到输出等均用到控制系统；

数学与运筹学家。主要参与适应于当今生物材料与技术的计算模型的研究，以及在生物计算机设计过程中的种种优化问题等；

电子计算机科学家。由于我们所设计的生物计算机输入/输出辅助系统是通过电子计算机来实现的，加之在设计过程中用到诸多的信息处理问题。因此，电子计算机是一个基本工具；

分子生物学家。生物计算机的核心硬件是分子生物材料与分子生物技术等；

密码学专家。生物计算机的研制成功，其第一个应用对象是密码分析与破译。

参考文献

1 Moore G. Progress in Digital Integrated Electronics. IEEE, IEDM Tech Digest, 1975, 11-13.

2 Feitelson Dror G. Optical Computing: A Survey for Computer Scientists. Cambridge, MA: MIT Press, 1988.

3 McAulay Alastair D. Optical Computer Architectures: The Application of Optical Concepts to Next Generation Computers. New York, NY: John Wiley & Sons, 1991.

4 Witlicki E H, Johnsen C, Hansen S W et al. Molecular Logic Gates Using Surface-Enhanced Raman-Scattered Light. J. Am. Chem. Soc., 2011, 133: 7288-7291.

5 David P D. Quantum Computation. Science，1995, 255-261.

6 Li H, Yao B, Tu T et al. Quantum computation on gate-defined semiconductor quantum dots. 2012, 57(16)：1919-1924.

7 Conrad M. On design principles for a molecular computer . Communication of the ACM , 1985 , 28(5)：464-480.

8 Birge R P. Protein-based computers. Scientific American , 1995, 272(3)：90-95.

9 Freemantle M. Protein devices may increase computer speed and memory. Chemical and Engineering News , May 22 , 1995，10-11.

10 Cukras A , Faulhammer D , Lipton R et al. Chess games:A model for RNA based on computation. BioSystems , 1998 , 52: 35-45.

11 Faulhammer Dirk , Cukras Ant hony R , Lipton Richard J et al. Molecular computation : RNA solutions to chess problems. Biochemistry , 2000 , 97(4)：1385-1389.

12 Adleman L M. Molecular computation of solution to combinational problem. Science, 1994, 266(5187)：1021-1024.

13 许进. 关于《中州I-型DNA计算机的研制方案》. 北京大学技术研究报告，2013.12.

14 Xu J, Qiang X L, Zhang K et al. A parallel type of DNA computing model for graph vertex coloring . IEEE Trans On NT.

15 许进，张雷. DNA计算机原理、进展及难点(I)：生物计算系统及其在图论中的应用. 计算机学报，2003, 26(1)：1-12.

16 许进，黄布毅. DNA计算机:原理、进展及难点(II)计算机“数据库”的形成———DNA分子的合成问题. 计算机学报，2005, 28(10)：1583-1591.

17 许进，张社民，范月科等. DNA计算机原理、进展及难点(III)：分子生物计算中的数据结构与特性. 计算机学报，2007, 30(6)：1-12.

18 许进，谭钢军，范月科等. DNA计算机:原理、进展及难点(IV)：论DNA计算模型. 计算机学报, 2007, 30(6)：881-893.

19 许进，李菲. DNA计算机:原理、进展及难点(V)：DNA分子的固定技术. 计算机学报，2009, 32(12)：2283-2299.

20 Xu J, Dong Y F. Sticker DNA computer model—Part I: Theory. Chinese Science Bulletin 2004, 49(8)：772-780.

21 Xu J, Li S P, Dong Y F et al. Sticker DNA Computer Model—Part II: Application. Chinese Science Bulletin, 2004, 49(9)：863-871.

22 Tanaka Fumiaki , Kameda Atsushi, Yamamoto M et al. Design of nucleic acid sequences for DNA computing based on at hermodynamic approach. Nucleic Acids Research , 2005 , 33(3)：903-911.

23 Yaakov Benenson, Binyamin Gil, Uri Ben-Dor et al. An autonomous molecular computer for logical control of gene expression. Nature, 2004, 429(6990)：423-429.

24 Binyamin Gil, Maya Kahan-Hanum, Natalia Skirtenko. Detection of multiple disease indicators by an autonomous biomolecular computer. Nano Lett., 2011，11(7)：2989-2996.

25 Feyman R P. There’s plenty of room at the bottom, in Minaturization. Gilbert D H, ed. New York: Reinhold, 1961, 282-296.

26 Bennett C H. On constructing a molecular computer. IBM Journal of Research and Development, 1973, 17:525-532.

27 Roweis S, Winfree E, Burgoyne R et al. A sticker based arcchtecture for DNA computation. in: Baum E B et al，DNA Based Computers，Proc. 2nd Annual Meeting，Princeton, 1999, 1-27.

28 Sakamoto K, Gouzu H , Komiya K et al . Molecular computation by DNA hairpin formation. Science, 2000, 288(49)：1223-1226.

29 Natasa Jonoska, Stephen A. Karl, Masahico Saito. Three dimensional DNA structures in computing. Biosystems, 1999, 52: 143-153.

30 Winfree E. Design and self-assembly of two-dimen sional DNA crystals. Nature, 1998, 394: 1223-1226.

31 Rothemund P. A DNA and restriction enzyme implementation of

turing machine, In DNA based Computers: Proceedings of the
DIMACS Workshop, 1995, Princeton University.
32 Akira Suyama. DNA Chips- Intergrated Chemical Circuits for DNA Diagnosis and DNA computers. Proc. 3rd International Micromachine Symp., 1997, 7-12.
33 Garzon M, Neathery P, Deaton R J et al. A new metric for DNA Computing. Proceedings of the 2nd Annual Genetic Programming Conference, 1997,472-487.
34 许进. 核酸序列设计:理论、算法及应用. 北京大学技术报告, 2011.
35 Xu J, Qiang X, Yang Y et al. An Unenumerative DNA computing model for vertex coloring problem. IEEE Transactions on Nanobioscience, 2011,10(2):94-98.
36 Zhang C, Wu L Q, Yang J et al. A Molecular Logical Switch beacon controlled by thiolated DNA signals. Chemical Communications, 2013, 49: 11308-11310.
37 Zhang C, Ma J J, Yang J et al. Binding assistance triggering attachments of hairpin DNA onto gold nanoparticles. Analytical Chemistry, 2013, 85(24): 11973-11978.
38 Yang J, Shen J J, Ma J J et al. Fluorescent nanoparticle beacon for logic gate operation regulated by stranddisplacement. ACS Appl. Mater. Interfaces, 2013,5(12): 5392-5396.
39 Zhang C, Ma J J, Yang J et al. Nanoparticle Aggregation Logic Computing Controlled by DNA Branch Migration. Applied Physics Letter, 2013, 103: 093106.
40 Zhang C, Ma J J, Yang J et al. Selective control of gold nanoparticles based on circular DNA strand displacement. Journal of Colloid and Interface Science, 2014, 418: 31-36.
41 Shi X L, Lu W, Wang Z Y et al. Programmable DNA tile self-assemblyusing a hierarchical sub-tile strategy. Programmable DNA tile self-assembly using a hierar-chical sub-tile strategy Nanotechnology 2014,25(in press).

Forthcoming Era of Biological Computer

Xu Jin
(School of Electronics Engineering and Computer Science, Peking University, Beijing 100874, China)

Abstract Biological computer is a novel computer model, which uses nucleic acid molecular as "data" and uses enzyme and biological operations as informational processing tools. In 1959, Feynman conceived a kind of computer in molecular scale. In 1994, Adleman proposed a DNA computing model based on biochemical reaction. The breakthrough of the biological computer is Peking University's parallel DNA computation model, it was carried out in 2007, in which the 3-coloring problem of a 3-chromatic graph with 61 vertices was solved. The computation complexity is 3^{59}, which means that it would take 13217 years to complete the computation process even by the fastest supercomputer. This fact seems to herald the era of biological computer. In this paper, we mainly present the background and significance of biological computer, and give an instruction on the basic principles and calculation method of DNA computer, especially Zhongzhou I-type DNA computer. Then we briefly summarize its applications in the fields of cryptanalysis and decipher. Furthermore, we discuss the computational ability of DNA computers, and point out some difficulties in the study of biological computer. Finally, we predict the possible development trends of DNA computing in the future, and propose some advices to the further research on DNA computing in China.

Keywords biological computer, non-enumeration DNA computing, parallel DNA computer, large scale DNA computer, cryptanalysis and decipher, research progress, proposals

许 进 北京大学信息科学技术学院教授，博导。1959年出生于陕西乾县。1993年获西安交通大学管理工程专业工学博士；1995年获北京理工大学数学系理学博士；1995年在西安电子科技大学师电路与信号处理领域进行博士后研究。国际生物计算机学术会议组委会核心成员；中国电子学会电路与系统学会委员；中国电子学会图论与系统优化专业委员会理事长。1996年访问香港中文大学半年；1997－1998年应邀到新加坡国立大学作为计算机系教授开展研究工作。2006—2010年先后2次任国际生物计算机大会主席。主要研究方向：DNA纳电子技术、分子信号检测、DNA传感器、DNA计算和图论与组合优化等。在 *Discrete Mathematics*、*International J. of Graph Theory*、*Graph and Comb.*、*IEEE Trans. On AC*、*IEEE Electronical Letters*、*J. of Chemistry and Physics*、*J. of Chemical Information and Computer Sciences*、*BioSystems*、《中国科学（E辑）》、《中国科学（F辑）》以及《科学通报》等权威与重要刊物发表学术论文200余篇，其中*SCI*检索110余篇，*EI*检索50余篇，他引1 236次（截至2008年2月）；出版学术专著4部。1996年开始从事DNA纳米技术和分子计算机模型、理论以及结构体系学习与研究。2000年负责组建我国第一个生物计算机研究所。目前已培养该领域我国首批博士后9名、博士28名；正在培养的在站博士后5名、在读博士生10名。其生物计算机研究成果，获2013年国家自然科学奖二等奖。E-mail:jxu@pku.edu.cn

（接137页）

Demand Promotes Innovation, and S&T Drives Development —Perspectives on the S&T Support of CAS on Geological Disaster Prevention in Zhangmu, Tibet, China

Duan Xiaonan[1] Wei Fangqiang[2] Feng Renguo[3]

（1 Bureau of Frontier Science and Education, Chinese Academy of Sciences, Beijing 100864, China

2 Insitute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China

3 Bureau of Science & Technology for Development, Chinese Academy of Sciences, Beijing 100864, China）

Abstract In this paper, the S&T support project of CAS on geological disaster prevention in Zhangmu, Tibet, was introduced, some new thoughts on collaborative innovation were also presented. In the course of S&T-driven regional development, how to acquire scientific issues from actual demand, with the synergistic advantages in the system of policy-production-education-research-application, was also illustrated. Based on the chain design of innovation-promoted development, the paper aims to provide a reference and enlightenment on the support role of S&T innovation for the national and local development of economy and society.

Keywords Zhangmu, Tibet, geological disaster, demands of science and technology, collaborative innovation

段晓男 中科院前沿科学与教育局地球科学处副处长，生态学博士。1979年出生。主要从事地球科学相关领域的科研管理与服务工作。E-mail：xnduan@cashq.ac.cn

（相关图片请见封二）