



核酸分子计算及应用

沈心怡¹, 刘俊兰¹, 韩达^{1,2*}

1. 上海交通大学医学院分子医学研究院, 上海 200127

2. 中国科学院杭州医学研究所, 杭州 310022

*通讯作者, E-mail: dahan@sjtu.edu.cn

收稿日期: 2024-02-17; 接受日期: 2024-03-25; 网络版发表日期: 2024-03-28

国家杰出青年科学基金(22225402)资助项目

摘要 作为一种新兴计算技术, 核酸分子计算因其信息存储量大、并行性高、微型化、低能耗的优点而备受关注. 核酸分子计算系统主要包括信号输入、信号处理、信号输出三个模块, 其目前已经能够实现数学逻辑运算、信息处理加密、模拟神经网络的功能, 或与纳米材料结合构成分子机器, 有望广泛应用于数学、医学、计算机科学等各种领域. 本文概述了核酸分子计算的原理、功能及其生物医学应用, 重点介绍了其在构建分子机器方面的研究进展, 并讨论了当下面临的挑战.

关键词 核酸分子, 计算系统, 分子机器, 纳米技术

1 引言

计算机具有强大的数据存储和信息处理能力, 在现代科学技术的发展过程中必不可少. 但随着当今计算数据量的增大和各复杂运算系统的产生, 传统的硅基计算机已经不能满足当下日益增长的运算需求, 核酸分子计算应运而生^[1]. 其通过将需要处理的信息转化为输入链上特定的碱基序列, 基于碱基互补配对原理, 通过改变反应条件实现不同的数学运算, 并测量输出链的最终平衡浓度, 从而得到运算结果. 核酸分子计算的信息存储量大、并行性高, 且具有微型化、低能耗等优点, 近年来引起了广泛的研究兴趣.

自Adleman开创性地将DNA作为基本单元来解决哈密顿路径问题以来^[2], 各种计算方式被进一步成功地设计出来. 同时, 随着Ned Seeman开创的核酸纳米

技术的蓬勃发展^[3], 核酸分子不再被认为是一种纯粹的遗传密码载体, 而是成为一种高效的工程材料, 适用于分子机器的构建. 分子机器是指由分子尺度的物质构成且能实现机械运动、行使某种加工功能的机器^[4], 通过核酸分子实现运算功能就属于分子机器的一种.

本综述总结了核酸分子计算系统的基本类型及其实现的基本功能, 着重介绍核酸分子机器的构建与最新进展, 并详细列举了其在各领域的生物应用. 本文还讨论了当前其面对的障碍和挑战, 借此对目前的研究前沿提供概览.

2 核酸分子计算的原理

利用核酸分子进行计算, 本质上是输入信号后通

引用格式: Shen X, Liu J, Han D. Nucleic acid-based computation and its applications. *Sci Sin Chim*, 2024, 54: 777–787, doi: 10.1360/SSC-2024-0027

过核酸分子的反应输出相应的计算结果. 整个计算系统包括信号输入、信号处理、信号输出三个部分. 输入与输出的信号都是一定浓度、特定序列的核酸分子链. 输出信号通常通过荧光信号、可见光信号或磁性信号等可分辨的方法检测.

信号处理即计算模块, 主要通过逻辑门(图1)或数学运算完成. 本文重点介绍核酸分子逻辑门这类基本信号处理模块, 包括YES, AND、OR、NOT、NAND、NOR、XOR. 这些逻辑门可以进一步排列组合成为更加复杂的计算系统, 主要可以分为无酶参与和有酶参与的反应系统, 丰富了算法的应用.

2.1 无酶参与的信号处理系统

在无酶参与的信号处理系统中, 核酸分子相互反应完全依靠碱基自由互补配对, 对序列设计、反应物浓度、反应条件具有极高的要求.

2.1.1 基于链杂交反应的系统

最初的DNA计算系统始于1994年^[2]. 由于之前的最坏情况复杂度算法效率不够高(即多项式时间), 无法确定具有指定顶点的任意有向图是否具有哈密顿路径, Adleman构建了基于DNA链杂交的并行计算算法来解决上述挑战(图2). 他设计了特定的DNA序列来编码原始图中的各节点和边缘线段, 在混合了这些核酸分子的溶液体系中, DNA根据碱基互补配对原则进行链杂交并形成代表有效路径的核酸序列. 该研究引起了广泛的兴趣和注意, 预示了核酸分子计算在解决计算机科学中的困难组合问题的强大潜力^[5].

1995年, Lipton仿效Adleman的做法, 同样用链杂交反应解决了可满足性问题, 即通过一组值为“真”或

常用逻辑门及其运算								
名称	与门 或门 非门 与非门 或非门 异或门 同或门							
	AND	OR	NOT	NAND	NOR	XOR	XNOR	
符号								
输入信号		输出信号						
0	0	0	0	1	1	1	0	1
0	1	0	1	1	1	0	1	0
1	0	0	1	0	1	0	1	0
1	1	1	1	0	0	0	0	1

图1 常用逻辑门及其运算
Figure 1 Common logic gates and their operations.

“假”的布尔变量对公示进行适当赋值, 从而使整个布尔公式的输出为真^[6]. 此外, 1999年Roweis等人^[7]为了降低Adleman-Lipton模型杂交的错误率, 提出了Sticker模型, 其中一条核酸分子链可存储多个比特, 每步运算的输出链可进一步用于下一步计算, 且需要的试管数是Adleman-Lipton的一半. 2007年Lin等人^[8]结合微阵列技术进一步改善Sticker模型, 仅需要线性增加表示变量的核酸序列种类并增加数组的大小, 而无需延长序列, 更适用于处理大规模问题.

2.1.2 基于链取代反应的系统

随着研究的进展, 2000年Yurke等人^[9]将“toehold”的概念引入到核酸分子计算中(图3a), 开创了链取代的方法. 图中A链与B链反应生成C链, 反应继续进行

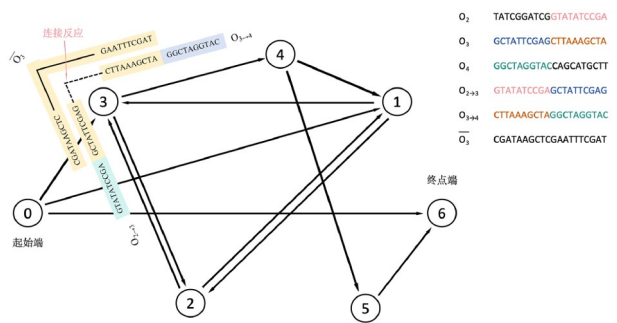


图2 求解哈密顿路径问题的算法示意图^[2] (网络版彩图)
Figure 2 Schematic illustration of the algorithm solving the Hamiltonian path problem ^[2] (color online).

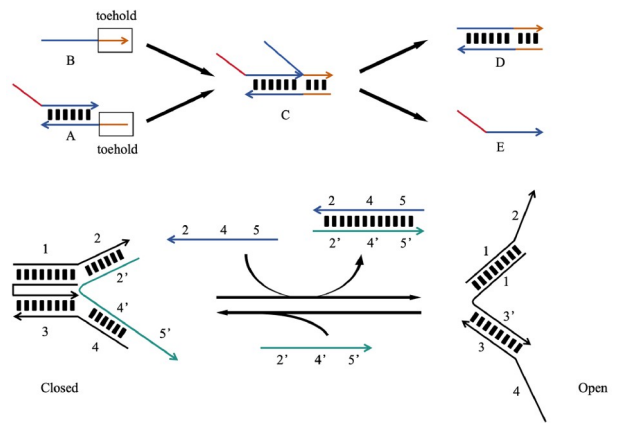


图3 (a) 链取代反应示意图; (b) 基于链取代反应的分子机器示意图^[9] (网络版彩图)
Figure 3 (a) Schematic illustration of strand displacement response; (b) schematic illustration of molecular machine based on strand displacement response ^[9] (color online).

并最终生成D链, 相当于B链将E链取代。

他们利用toehold介导的链取代反应开发了一种分子机器(图3b), 靶向的单链DNA促使开关在封闭状态与开放状态之间相互转换。该反应由取代后的新链吉布斯自由能的降低诱发, 反应速度与加入单链黏性末端长度、碱基对和链浓度相关^[10]。随后, 大量基于功能性核酸的逻辑门被开发出来。

进一步, Chen等^[11]构建了一种多米诺骨牌样的多输入与门计算系统(图4), 取代了先前复杂的多步操作, 降低了成本。当输入链a(tada)存在时, 与暴露的toehold ta*结合, 替换链da。由于较长的链部分处于未结合状态, 链tb将脱离, 从而暴露出tb*。此时输入链b(tbdb)与tb*结合, 置换过程与输入链a(tada)相同。最后输入链c(tc dc)取代链dc-report发出荧光信号。该反应的一个重要特性是可级联, 一个反应的输出可作为其下游反应的输入, 这一特性使反应能够被放大并执行更复杂的运算。

2.1.3 基于特殊核酸分子结构的系统

此外, 还有基于G-四联体、i基序、核酸适体结构的计算系统。G-四联体是四个鸟嘌呤通过Hoogsteen碱基配对后堆叠而成的特殊结构, 在相应螯合剂存在下可逆地形成或解离; i基序是由两个富含胞嘧啶(C)的平行链形成半质子化的C-C+碱基对稳定结合组成的二级结构, 是抗癌药物的潜在靶点; 核酸适体是通过SELEX程序筛选出的特异性与靶物质高度亲和的短DNA/RNA链, 有助于实现逻辑门引导的目标特异分析。这些特殊的序列结构都丰富了核酸分子计算的可能性^[12]。

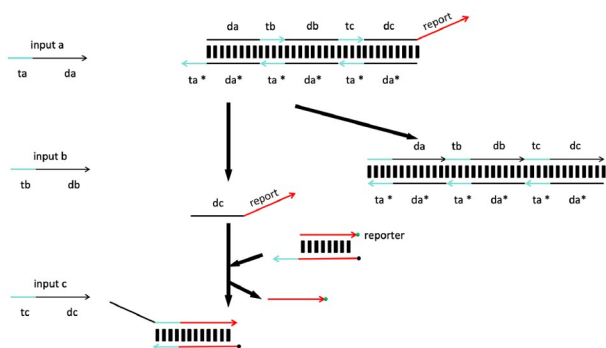


图4 多输入与门计算系统示意图^[11] (网络版彩图)

Figure 4 Schematic illustration of multi-input AND gate computing system ^[11] (color online).

2.2 有酶参与的信号处理系统

在核酸分子计算中, 不同的酶相当于不同的算子, 可用来模拟简单的计算: 内切酶相当于分离算子, 外切酶相当于删除算子, 连接酶相当于连接算子, 聚合酶相当于复制算子。有酶参与的系统使信号处理过程更加高效可控。

2.2.1 脱氧核酶反应系统

脱氧核酶(DNAzyme)具有RNA切割活性, 常用于切割底物RNA并进行检测。Stojanovic等将人工形成的核酸分子链进行链取代反应后产生脱氧核酶活性, 进一步实现下级酶促反应, 从而实现多种DNA逻辑门^[13](图5)。Willner及其同事构建了多个独立的计算模块并整合^[14,15], 在不同的输入情况下, 形成特定的DNAzyme, 并裂解相应的底物, 发出不同荧光信号, 实现半加法器、半减法等复杂运算。以类似的方式, Kolpashchikov等^[16]实现多输入与门, 可用于分析多种生物标志物。

2.2.2 内切酶反应系统

人们在对噬菌体的宿主特异性修饰现象研究中发现了内切酶^[17], 它可以在特定位点切割DNA并将其切成小的DNA片段。I型和III型酶均表现出限制性修饰作用, 但III型酶含有特定的裂解位点。II型酶, 称为限制性内切酶, 可以识别和切割DNA链上的特定核苷酸序列, 产生特定的DNA片段。

2017年, Pan等^[18]根据内切酶的特点, 构建了YES、NOT和AND逻辑门, 通过添加或移除toehold来特异性切割DNA链, 从而启动或抑制链取代反应, 可用于切换逻辑门和桥接蛋白质。该复杂电路的构建表

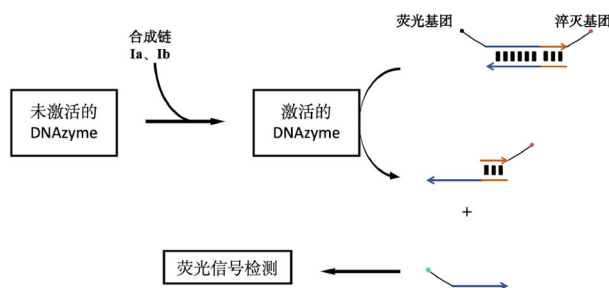


图5 脱氧核酶反应系统示意图^[13] (网络版彩图)

Figure 5 Schematic illustration of deoxyribozyme reaction system ^[13] (color online).

明, 内切酶辅助的toehold调节机制能够可靠、高效地实现多酶电路, 且大大降低了干扰可能性。

2.2.3 外切酶反应系统

DNA电路级联和分流过程中的信号衰减影响了计算结果的可靠性和DNA电路规模的进一步扩大。Zhang等^[19]利用提取自感染 λ 噬菌体的大肠杆菌内部的 λ 外切酶, 它可在5'磷酸基团修饰的dsDNA末端进行高效水解, 而无需额外的燃料链或能量。他们构建一系列可独立调节输出信号强度的并联DNA逻辑电路, 并提出了一种外切酶驱动的反应物再生策略。也有研究设计了一种双链分离(DSS)策略, 结合了toehold介导的链取代和外切酶III (Exo III)。前者产生的副产物链由于具有无碱基位点(AP)可被Exo III进一步水解, 从而产生额外的输出信号, 实现信号放大和多逻辑计算^[20]。

2.2.4 聚合酶反应系统

聚合酶具有的等温延伸特性可以填补缺失的单链片段, 一般通过与其他酶共同作用产生级联反应, 引发下游剪切实现目的结果。Orbach等人设计了一个聚合酶参与的自主激活系统, 聚合反应产生DNAzyme结构实现单链的剪切并输出荧光信号, 实现了与门、或门和可控与门的组装^[21]。Zhang等^[22]提出了一种生物传感平台, 巧妙地将链取代扩增(SDA)和聚合酶/切口酶信号扩增(NESA)集成为一步系统, 通过聚合、扩增、剪切的循环实现目标miRNA的信号识别与放大。Yin等^[23]将DNA聚合酶催化的等温聚合反应和启动信号转导的链取代聚合反应串联形成分子开关, 以实现特定miRNA的检测信号放大。

3 核酸分子计算的功能

以上基于不同信号处理系统的核酸分子计算均可实际应用于完成数学运算、解决逻辑难题、信息加密解码、神经网络模拟等具体功能, 本章将进行总结阐述。

3.1 数字运算与数论问题解答

核酸分子计算发展伊始便应用于数学问题的解决。1996年, Guarnieri等^[24]首次提出了一种利用DNA

链将两个正整数相加的算法。虽然该技术允许计算和, 但输出链结构与输入链结构不同。因此, 加法运算的输出链不能直接用作下一个求和的输入链, 且对位数有所限制。Fujiwara等人^[25]则提出了一种用于表示和操作二进制数的DNA模型, 可对DNA链中存储的变量进行基本运算(图6)。此外, Fukagawa和Fujiwara^[26]提出了运算乘法和除法的程序, 基本完成了分子水平上的四则运算。

另一方面, Ouyang等^[21]研究了一种利用分子生物学技术解决最大派系问题(MCP)的方法, 将集合映射到二进制数, 然后对其进行剔除和分类。在此基础上, 部分基础的集合问题得到解决^[27-30]。2006年, Yeh等人^[31]在Adleman和Lipton的基础上提出了一种新的DNA操作来解决二进制整数规划(BIP)问题, 证明了DNA计算在BIP问题上的前景。2017年, Yin等^[32]提出了一种DNA算法, 通过操纵质粒来解决特殊的整数规划问题。

3.2 图论与数学模型问题解答

自完成对哈密顿路径问题(HPP)^[2]和SAT问题^[6]的解答后, 核酸分子计算就被用于多种相似逻辑模型的研究。最经典的旅行商问题(HSP)即是在Adleman的研究基础上得到解决的^[33]。给定一系列城市 and 每对城市之间的距离, 求解访问每一座城市一次并回到起始城

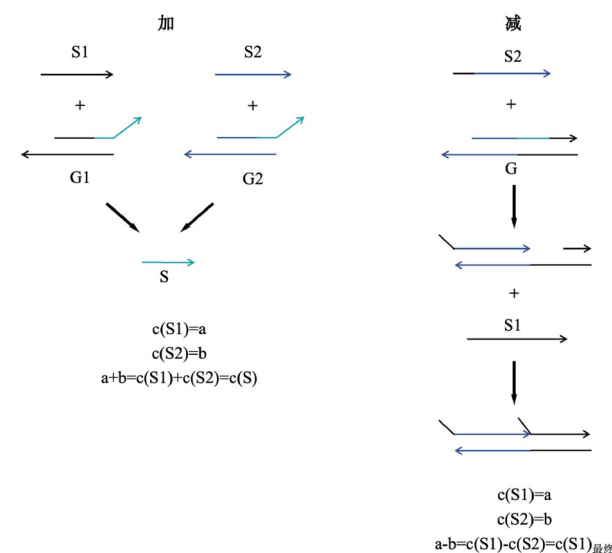


图6 DNA加减法运算示意图^[25] (网络版彩图)

Figure 6 Schematic illustration of addition and subtraction of DNA^[25] (color online).

市的最短回路. 一些变体例如广义旅行商问题(GTSP)^[34]、收奖旅行商问题(PCTSP)^[35]、配额旅行商问题(QTSP)^[36]、家庭旅行商问题(FTSP)^[37]也已经通过DNA计算找到最优解. 值得一提的是, Lee等人提出了一种基于DNA热力学性质的算法^[38], 并利用其DNA解链温度的变化来编码实际值. 他们将DNA计算的能力扩展到解决数值优化问题, 是一个重大的进步. 也有研究利用DNA折纸结构实现三色图形问题的逻辑计算, 通过三维折纸的自组装及原子力显微镜证实空间连接方法, 利于结果的可视化^[39].

3.3 信息读取与图像加密

通过有酶/无酶的反应, 核酸分子计算可以实现信息(核酸序列)的储存、传递、读取^[40], 进一步实现信息加密, 用于商务等领域. 有研究设计了混沌系统进行图像加密^[41]. 2018年, Wang等^[42]提出了一种校正图像加密错误的新方法. 2021年, Namasudra等^[43]提出了一种新的基于DNA计算的加密方案, 用于保护云计算环境下的多媒体文件.

3.4 基于集成逻辑门的神经网络模拟

当核酸分子计算系统包含大量的逻辑门时, 如何进行大规模并行计算是必需考虑的问题. Qian等^[44]利用基于toehold交换原理的可逆链取代反应创建了一个“跷跷板”电路架构. 其反应可逆, 且一对“跷跷板”可完成催化循环、信号放大并进行与门和或门运算. 同时通过使用即插即用的分子组件, 门之间可任意组合, 大规模并行计算成为可能. He等^[45]提出了一种基于三组dsDNA结构的更稳定的逻辑门设计, 在室温下在微流控芯片器件中得到了验证, 适用于大型DNA纳米电路. Ma等^[46]也基于链取代反应设计了三输入逻辑门. 近日, 有研究设计出集成的无明显信号衰减的大规模DNA可编程门阵列(DPGA)网络, 标志着迈向通用DNA计算的关键一步^[47](图7).

最早尝试在化学介质中进行类似神经网络的计算的是Hjelmfelt等人^[48]. 作者设计了一个可逆的酶反应系统来模拟神经元放电和静止的工作模式, 神经元状态取决于输出信号与阈值的关系. 此后Qian和Winfree等^[49]实现了记忆四种分子模式的DNA链网络. Cherry和Qian^[50]利用“赢者通吃”的神经网络策略, 成功识别了九种分子模式. 此外, DNA分子计算还能用于卷积

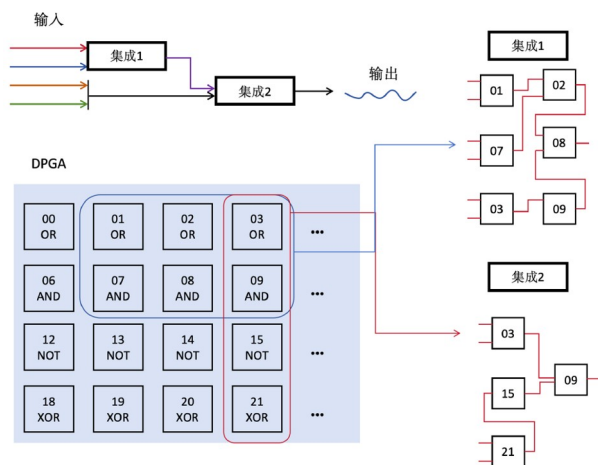


图7 DNA可编程门阵列示意图^[47] (网络版彩图)

Figure 7 Schematic illustration of DNA-based programmable gate arrays ^[47] (color online).

神经网络(CNN)进行模式识别和并行任务处理^[51,52].

4 核酸分子计算系统的应用

目前核酸分子计算系统已经逐渐投入实际应用, 在分子、细胞、临床等多个层面迅速发展.

4.1 构建核酸分子机器

核酸分子计算过程的动态性、可调控性与生物相容性使核酸分子与纳米材料结合构成分子机器成为可能. 其能够在特定刺激下执行类机械化运动, 包含DNA步行器、DNA镊子和DNA机器人等, 可用于多种生物分子的检测.

4.1.1 智能DNA逻辑纳米器件(CDLN)

DNA逻辑纳米器件是分子计算任务和智能生物分析应用的强大工具. Han等人^[53]开发了一种并行反逻辑的CDLN, 实现了聚腺嘌呤链及其聚合酶的智能体外分析. Mattath等^[54]将嘧啶突出末端加入DNA纳米团簇(NCs), 具有独特的自组装特性和可编程性, 可进行复杂逻辑运算(XOR and XNOR)和错误检测. 韩国的一项研究在铜纳米团簇合成过程中加入果糖得到了具有时间荧光稳定性的产物^[55]. Stojanovic等^[56]设计的分子自动机可以与玩家对战井字棋游戏. 其可相应对手的移动方式, 并通过荧光信号指示完成自己的移动,

且无法被打败。以上器件均具有动态读取、智能分析、输出信息的能力。

4.1.2 可调控DNA步行器

DNA步行器一般以DNA的水解为能量来源, 通过与轨道之间的结合-解离实现定向运动。Wang等人^[57]改善了金纳米颗粒(AuNP)行走器和轨道之间的无序组装问题, 将纳米机器构建在DNA树状大分子上, 底物链作为轨道、DNAzyme作为行走器, 可以精确调节底物链和DNAzyme在DNA树状大分子上的分布, 检测miRNA动态表达水平。DNA步行器同样可用于检测其他生物大分子^[58], 且可避免热循环扩增。Liu等人^[59]开发了调控工具clip实现对反应速率的细微调节、变构链取代、反应区域选择性激活、反应重置, 并应用于DNA步行器, 实现了可控行走、串联和可编程路径的行为。可运动的DNA纳米结构可进一步通过结构设计实现DNA货物运输等功能(图8)。Bath等^[60]通过限制性内切酶切断轨道使货物运动具有单向性。

4.1.3 基于多种识别元件的生物传感器

生物传感器中的分子识别元件可以是各种纳米级小分子, 需要具有高特异性、灵敏度^[61]。DNA四螺旋、三螺旋发光纳米传感器, 可以监测溶酶体腔内的

K⁺和pH, 解决了过往人们对溶酶体酸化的疑惑^[62]。三维DNA晶体的自组装已被证实具有实现逻辑门、进行大规模并行信息处理的潜力, 同时具有生物相容性、易于观察, 将来很可能成为纳米传感器的主要材料^[63]。

目前识别元件的固定、逻辑器件的设计、反应条件均会影响到分子机器的使用效果。Hu等^[64]开发了一种可以在保持传感能力的同时响应外部刺激的纳米器件, 并成功地应用于ATP传感器。且只有在外界刺激和目标分析物的共同作用下, 荧光信号才会发生变化, 避免了部分环境因素的影响。

适配体作为易于合成和长期储存的寡核苷酸, 其特异且多信号传导方式的优势使其一直以来受到研究者的青睐。然而目前适配体传感器的主要问题在于其对痕量检测物的灵敏度有限, 尤其是具有大 K_d 值的小分子。如何降低检出限(LOD)是目前亟待解决的问题^[12]。

4.2 细胞活动检测与调控

细胞层面上, 核酸分子计算已应用于细胞鉴定、胞内生物传感、逻辑电路调控细胞活动等。有研究利用模块化传感实现了细胞内源性miRNA的生物计算操作, 并结合多种分析物实现逻辑运算, 实现信号放大和针对细胞的准确诊断^[65]。Du等人^[66]利用溶酶体识别框架核酸(FNA)在活细胞内的溶酶体内建立纳米平台(图9), 通过装置上i基序和ATP结合适配体(ABA)对溶酶体pH和ATP做出反应, 从而触发FNA的结构变化, 并通过亚细胞成像确认。Siuti等人^[67]在活的大肠杆菌细胞中创建了所有16个双输入布尔逻辑函数, 并保持了长达90代的细胞记忆维持。

4.3 疾病诊断与治疗

血清miRNA、DNA、蛋白标志物、离子或局部组织细胞代谢物的含量在肿瘤、炎症情况下往往会发生改变。Wang K等设计的新型纳米机器, 通过串联链取代反应和DNA步行器相结合来检测DNA。用电化学适配体传感器作为识别分子, 可较准确地完成对不同患者血清中目标DNA的检测^[68]。Ma等^[69]开发了一个自动化的基于DNA计算的平台, 用于急性呼吸道疾病感染(ARI)病因分类。其将样品装载、标记物扩增、分类器实现和结果报告整合, 具有较快的检测速度和准确的检测结果。目前也有研究开发了完全自动化的用于

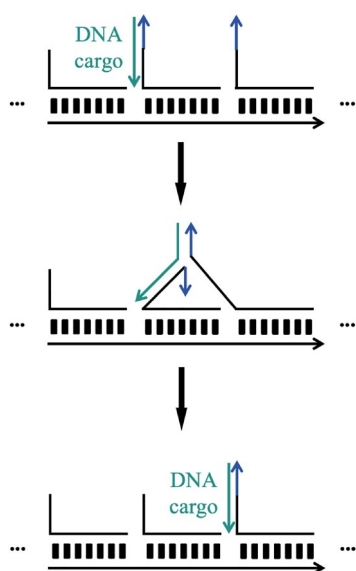


图8 DNA货物运输功能示意图^[59] (网络版彩图)

Figure 8 Schematic illustration of DNA cargo transports ^[59] (color online).

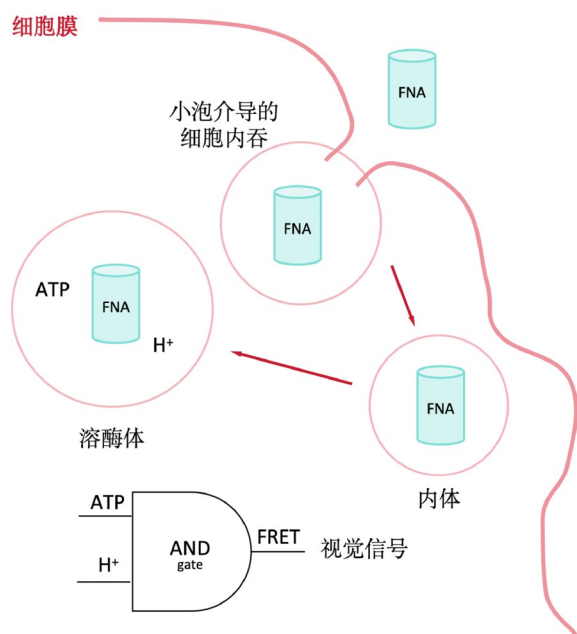


图 9 溶酶体识别框架核酸示意图^[66] (网络版彩图)

Figure 9 Schematic illustration of lysosome-recognizing framework nucleic acid [66] (color online).

癌症筛查的平台^[70](图10)。聚腺嘌呤mRNA在乳腺癌、胰腺癌、结肠癌和许多其他癌细胞中过表达。因此, DNA逻辑纳米器件对poly-A的识别分析及polyA聚合酶活性评估可提供强有力的生物学证据^[53]。目前的DNA纳米机器对目标蛋白质的识别种类和灵敏度都很有限, 为此Du等人^[71]构建了一个靶向结合诱导的DNA位移转换系统(BiDDCs), 用于将蛋白质识别结果输出为DNA链, 扩展DNA纳米机器的用途。

除以上检测, 目前利用核酸分子计算检测的疾病标志物多为血浆miRNA, 其对癌症早期诊断非常重要^[72,73]。Takeuchi等^[74]利用基于纳米孔的DNA计算技术检测miRNA表达模式, 检出限显著降低。Mameuda等^[75]通过将酶和DNA级联的方式模拟细胞质内的相关环境, 并成功检测出乳腺癌相关微量miRNA标志物。核酸分子计算系统可对微量疾病标志物进行高效检测, 根据疾病诊断标准设置合理的逻辑门策略, 从而达到对不同疾病准确、早期诊断的目的。疾病治疗方面, Li等^[76]设计的溶酶体识别框架核酸(FNA)可检测细胞内部pH和ATP水平, 同时FNA很容易地内化到活细胞中, 并在几个小时甚至几天内保持基本完整, 可能在可控药物释放和疾病治疗中得到广泛应用^[66]。Douglas

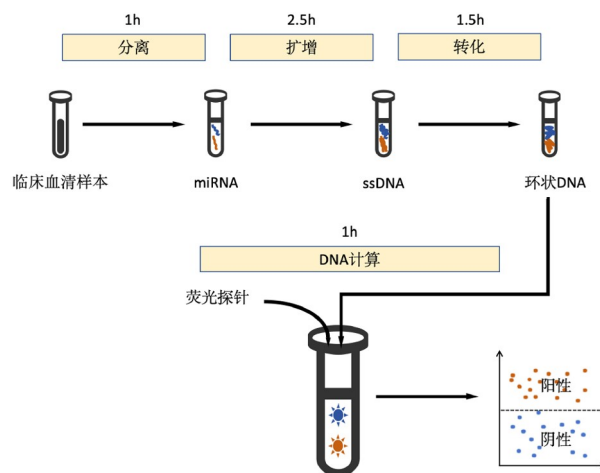


图 10 核酸分子计算用于癌症筛查示意图^[70] (网络版彩图)

Figure 10 Schematic illustration of cancer screening with nucleic acid molecular computation [70] (color online).

等^[77]开发了一种基于适体的DNA纳米机器, 它能够将小分子运送到细胞表面, 被特定的细胞表面信号激活后, 实现重构、进一步传递预先装载的分子, 进一步激活下游反应, 实现可控的细胞调控。Zhang等人^[78]设计了一种跨越细胞膜的基于DNA计算的纳米机器, 可实现对癌细胞表面受体的精准激活与调控, 有助于实体肿瘤的精准癌症治疗。

5 核酸分子计算面临的问题

核酸分子计算十分高效, 然而目前其障碍主要来自三个方面。

其一, 其对于编码的核酸序列要求较高。设计可靠的核酸序列要求研究人员从一个大的解集中找到满足多组冲突约束的DNA序列, 且没有某种固定的解法, 必须经过反复验证优化。目前核酸序列设计优化是核酸分子计算最大的挑战之一。近年来, 许多研究也提出了有助于序列优化的方法。Yang等^[79]提出多重GC约束的方法, 结合GC含量、Hamming距离、相似度、发夹结构等条件建立了新的条件函数, 提高DNA序列的稳定性和可靠性; Zhu等^[80]改进了IBPSO优化算法, 提高了效率, 并引入不同的约束避免了DNA二级结构的产生; Xie等人^[81]利用BHAOA优化算法, 先将解离散化以增强全局搜索能力, 再改变策略提升局部优化能力。以上研究均有各自的局限性, 仍有待进一步

优化。

其二, 反应的核酸分子并不能无限并行, 且数量和序列长度会影响到反应速度^[82]。为扩大反应容量, 可以通过模块化反应平台, 构建多组分、多通道的体系, 同时能更好地利用有限的荧光染料实现易于区分的多种输出信号。最近, 有研究通过冻融循环提高了局部分子有效浓度, 从而提升了反应速度^[83], 可用较少的核酸分子量实现尽可能快的计算。最新的纳米孔传感器阵列技术可以实现更加高度并行的检测^[84]。Chatterjee等^[85]开发了一种固定在DNA折纸表面的计算系统, 可以在一小时内计算出双输入双轨XNOR逻辑。他们将发夹结构定位在折纸支架链上, 利用级联式链交反应使发夹结构按照定位顺序打开并发出信号。折纸结构可能在高密度的细胞环境中会有更好的效果, 并有更广泛的应用前景; 但该技术更复杂、影响因素更多, 随着反应规模的扩大, 折纸之间可能存在相互作用影响反应速率。

其三, 大量核酸分子反应不可避免会存在串扰、突变、错配影响。Lv等^[86]构建的纳米器件利用杂交链反应作为信号放大器, 将弱输入信号转化为强输出信号, 大大提高了灵敏度, 降低了不同模块间的串扰。Famoori等^[87]提到一种基于剩余数表示系统(RNS)的具有纠错能力的DNA计算系统, 结合Sticker模型通过将大数的计算转化为余数运算, 降低了错误概率; Zheng的研究^[88]中也利用了这一点。进一步, 有研究通过使用磁镊子拉伸力来控制ssDNA链取代反应速度, 增强反应可控性^[89]。也可通过引入CRISPR系统增强反应

特异性。CRISPR基因编辑技术利用向导RNA(gRNA)引导Cas蛋白靶向识别和切割特定DNA序列, 其保证了CRISPR-RNA和目标DNA的一一对应。此外, CRISPR技术对靶标的序列和结构限制较小, 反应过程不可逆, 因此反应速度可保持较高水平。然而, 当后续反应需要集成大量的逻辑门时, 反应速度将受到链位移反应的动力学限制^[90]。目前大多数CRISPR应用集中在基因编辑和疾病诊断上, 在核酸分子计算方面的应用亟待进一步探索。

6 总结与展望

核酸分子计算改变了信息时代人们的计算方式。其通过有酶或无酶的反应, 可以实现包括数学运算、解决逻辑难题、信息加密解码、神经网络模拟在内的多种功能。将不同的运算模块组合成系统可实现更加复杂的运算网络, 进一步实现生物传感、可编程分子组装、模式识别等等, 应用于数学、医学、密码学等各个领域。尽管当下核酸分子计算仍具有一些明显的劣势, 在实际应用中无法完全取代电子计算机。但随着科学技术的飞速革新, 其依然具有很大的应用前景。

如今分子生物学技术和纳米技术同样在快速发展, 与核酸分子计算结合孕育的分子机器因其易于构建、可控制、生物相容性等优点而受到广泛关注。相信未来研究人员会逐步消除障碍, 实现更广泛的运算网络, 取得更多突破性的创新。

参考文献

- 1 Amos M. *Bull Eur Assoc Theor*, 1999, 67: 125–138
- 2 Adleman LM. *Science*, 1994, 266: 1021–1024
- 3 Chen YJ, Groves B, Muscat RA, Seelig G. *Nat Nanotech*, 2015, 10: 748–760
- 4 Balzani V, Credi A, Raymo FM, Stoddart J. *Angew Chem Int Ed*, 2000, 39: 3348–3391
- 5 Li F, Xiao M, Pei H. *ChemBioChem*, 2019, 20: 1105–1114
- 6 Lipton RJ. *Science*, 1995, 268: 542–545
- 7 Roweis S, Winfree E, Burgoyne R, Chelyapov NV, Goodman MF, Rothmund PWK, Adleman LM. *J Comput Biol*, 1998, 5: 615–629
- 8 Lin CH, Cheng HP, Yang CB, Yang CN. *Biosystems*, 2007, 90: 242–252
- 9 Yurke B, Turberfield AJ, Mills Jr AP, Simmel FC, Neumann JL. *Nature*, 2000, 406: 605–608
- 10 Yang S, Li J, Cui Y, Teng Y. *Chin J Biotechnol*, 2021, 37: 1120–1130 (in Chinese) [杨姗, 李金玉, 崔玉军, 滕越. 生物工程学报, 2021, 37: 1120–1130]
- 11 Chen X, Liu X, Wang F, Li S, Chen C, Qiang X, Shi X. *ACS Synth Biol*, 2022, 11: 2504–2512

- 12 Fan D, Wang J, Wang E, Dong S. *Adv Sci*, 2020, 7: 2001766
- 13 Stojanovic MN, Mitchell TE, Stefanovic D. *J Am Chem Soc*, 2002, 124: 3555–3561
- 14 Orbach R, Wang F, Lioubashevski O, Levine RD, Remacle F, Willner I. *Chem Sci*, 2014, 5: 3381–3387
- 15 Orbach R, Remacle F, Levine RD, Willner I. *Chem Sci*, 2014, 5: 1074–1081
- 16 Gerasimova YV, Kolpashchikov DM. *Chem Commun*, 2015, 51: 870–872
- 17 Meselson M, Yuan R. *Nature*, 1968, 217: 1110–1114
- 18 Pan L, Wang Z, Li Y, Xu F, Zhang Q, Zhang C. *Nanoscale*, 2017, 9: 18223–18228
- 19 Zhang X, Liu X, Yao Y, Liu Y, Zeng C, Zhang Q. *ACS Synth Biol*, 2023, 12: 2107–2117
- 20 Zeng C, Liu X, Wang B, Qin R, Zhang Q. *Analyst*, 2023, 148: 1954–1960
- 21 Orbach R, Mostinski L, Wang F, Willner I. *Chem Eur J*, 2012, 18: 14689–14694
- 22 Zhang Q, Chen F, Xu F, Zhao Y, Fan C. *Anal Chem*, 2014, 86: 8098–8105
- 23 Yin BC, Liu YQ, Ye BC. *Anal Chem*, 2013, 85: 11487–11493
- 24 Guarnieri F, Fliss M, Bancroft C. *Science*, 1996, 273: 220–223
- 25 Fujiwara A, Matsumoto KI, Chen W. *Int J Found Comput Sci*, 2004, 15: 461–474
- 26 Fukagawa H, Fujiwara A. Procedures for multiplication and division in DNA computing. In: *Proceedings of the 2006 International Conference on Foundations of Computer Science*. Las Vegas, USA, 2008
- 27 Chang WL, Guo M. *Biosystems*, 2003, 72: 263–275
- 28 Guo M, Ho MSH, Chang WL. *Parallel Computing*, 2004, 30: 1109–1125
- 29 Chang WL, Guo M, Wu J. *Parallel Process Lett*, 2005, 15: 469–479
- 30 Chang WL, Ren TT, Feng M. *IEEE Transon NanoBiosci*, 2015, 14: 121–128
- 31 Yeh CW, Chu CP, Wu KR. *Biosystems*, 2006, 83: 56–66
- 32 Yin Z, Cui J, Yang J. *Chin J electron*, 2017, 26: 1284–1288
- 33 Wang Z, Zhang Y, Zhou W, Liu H. *Appl Math Comput*, 2012, 219: 2267–2270
- 34 Ren X, Wang X, Wang Z, Wu T. *IJCIS*, 2021, 14: 228
- 35 Wang ZC, Liang K, Bao XG, Wu TH. *IEEE Transon NanoBiosci*, 2024, DOI: 10.1109/TNB.2023.3307458
- 36 Wang Z, Wu X, Wu T. *Comput Intel Neurosc*, 2022, 2022: 1450756
- 37 Wu X, Wang Z, Wu T, Bao X. *IEEE Transon NanoBiosci*, 2022, 21: 75–85
- 38 Lee JY, Shin SY, Park TH, Zhang BT. *Biosystems*, 2004, 78: 39–47
- 39 Xu J, Chen C, Shi X. *ACS Synth Biol*, 2022, 11: 2456–2463
- 40 Yin Q, Zheng Y, Wang B, Zhang Q. *IEEE ACM Trans Comput Biol Bioinf*, 2022, 19: 3384–3394
- 41 Signing VRF, Tegue GAG, Kountchou M, Njitacke ZT, Tsafack N, Nkapkop JDD, Etoundi CML, Kengne J. *Chaos Solitons Fractals*, 2022, 155: 111777
- 42 Wang B, Xie Y, Zhou S, Zheng X, Zhou C. *Molecules*, 2018, 23: 1878
- 43 Namasudra S, Chakraborty R, Majumder A, Moparthy NR. *ACM Trans Multimedia Comput Commun Appl*, 2020, 16: 1–19
- 44 Qian L, Winfree E. *Science*, 2011, 332: 1196–1201
- 45 He S, Cui R, Zhang Y, Yang Y, Xu Z, Wang S, Dang P, Dang K, Ye Q, Liu Y. *ACS Appl Mater Interfaces*, 2022, 14: 10721–10728
- 46 Ma J. *Sci Rep*, 2023, 13: 15210
- 47 Lv H, Xie N, Li M, Dong M, Sun C, Zhang Q, Zhao L, Li J, Zuo X, Chen H, Wang F, Fan C. *Nature*, 2023, 622: 292–300
- 48 Hjelmsfelt A, Weinberger ED, Ross J. *Proc Natl Acad Sci USA*, 1991, 88: 10983–10987
- 49 Qian L, Winfree E, Bruck J. *Nature*, 2011, 475: 368–372
- 50 Cherry KM, Qian L. *Nature*, 2018, 559: 370–376
- 51 Zhong G, Li T, Liu W, Chen Y. 2018, doi: 10.48550/arXiv.1810.10687
- 52 Xiong X, Zhu T, Zhu Y, Cao M, Xiao J, Li L, Wang F, Fan C, Pei H. *Nat Mach Intell*, 2022, 4: 625–635
- 53 Han J, Lv X, Zhang Y, Wang J, Fan D, Dong S. *Anal Chem*, 2023, 95: 16725–16732
- 54 Mattath MN, Zhang H, Ghosh D, Govindaraju T, Shi S. *Nanoscale*, 2023, 15: 17386–17397
- 55 Kim S, Lee ES, Cha BS, Park KS. *Nano Lett*, 2022, 22: 6121–6127

- 56 Stojanovic MN, Stefanovic D. *Nat Biotechnol*, 2003, 21: 1069–1074
- 57 Wang X, Sun H, Liu B, Jiang K, Li Z, Meng HM. *Anal Chem*, 2022, 94: 17232–17239
- 58 He MQ, Wang K, Wang WJ, Yu YL, Wang JH. *Anal Chem*, 2017, 89: 9292–9298
- 59 Liu L, Hu Q, Zhang W, Li W, Zhang W, Ming Z, Li L, Chen N, Wang H, Xiao X. *ACS Nano*, 2021, 15: 11573–11584
- 60 Bath J, Green SJ, Turberfield AJ. *Angew Chem Int Ed*, 2005, 44: 4358–4361
- 61 Khan S, Burciu B, Filipe CDM, Li Y, Dellinger K, Didar TF. *ACS Nano*, 2021, 15: 13943–13969
- 62 Chen F, Lu Q, Huang L, Liu B, Liu M, Zhang Y, Liu J. *Angew Chem Int Ed*, 2021, 60: 5453–5458
- 63 Zhang C, Paluzzi VE, Sha R, Jonoska N, Mao C. *Adv Mater*, 2023, 35: e2302345
- 64 Hu Y, Jia Y, Yang Y, Liu Y. *RSC Adv*, 2023, 13: 9003–9009
- 65 Gong X, Wei J, Liu J, Li R, Liu X, Wang F. *Chem Sci*, 2019, 10: 2989–2997
- 66 Du Y, Peng P, Li T. *ACS Nano*, 2019, 13: 5778–5784
- 67 Siuti P, Yazbek J, Lu TK. *Nat Biotechnol*, 2013, 31: 448–452
- 68 Wang K, He MQ, Zhai FH, Wang J, He RH, Yu YL. *Biosens Bioelectron*, 2018, 105: 159–165
- 69 Ma Q, Zhang M, Zhang C, Teng X, Yang L, Tian Y, Wang J, Han D, Tan W. *Sci Adv*, 2022, 8: eade0453
- 70 Moerman PG, Schulman R. *Nat Nanotechnol*, 2020, 15: 626–627
- 71 Du H, Yang P, Hou X, Zhou R, Hou X, Chen J. *Chem Commun*, 2019, 55: 3610–3613
- 72 Chen C, Chen X, Li X, Shi X. *Front Genet*, 2022, 13: 1064715
- 73 Lv WY, Li CH, Lin HR, Li LL, Zou HY, Li CM, Zhen SJ, Huang CZ. *Biosens Bioelectron*, 2022, 207: 114183
- 74 Takeuchi N, Hiratani M, Kawano R. *JACS Au*, 2022, 2: 1829–1838
- 75 Mameuda A, Takinoue M, Kamiya K. *Anal Chem*, 2023, 95: 9548–9554
- 76 Li J, Fan C, Pei H, Shi J, Huang Q. *Adv Mater*, 2013, 25: 4386–4396
- 77 Douglas SM, Bachelet I, Church GM. *Science*, 2012, 335: 831–834
- 78 Zhang Y, Chen W, Fang Y, Zhang X, Liu Y, Ju H. *J Am Chem Soc*, 2021, 143: 15233–15242
- 79 Yang X, Zhou C. *IEEE Transon NanoBiosci*, 2024, doi:10.1109/TNB.2023.3316431
- 80 Zhu D, Huang Z, Liao S, Zhou C, Yan S, Chen G. *IEEE T Nanobiosci*, 2023, 22: 603–613
- 81 Xie L, Wang S, Zhu D, Hu G, Zhou C. *Interdiscip Sci Comput Life Sci*, 2023, 15: 231–248
- 82 Zhou C, Song Y, Jin X, Li B, Pang C. *Nanoscale Horiz*, 2023, 8: 176–184
- 83 Zhu Y, Xiong X, Cao M, Li L, Fan C, Pei H. *Sci Adv*, 2023, 9: eaax7983
- 84 Zhang K, Chen YJ, Wilde D, Doroschak K, Strauss K, Ceze L, Seelig G, Nivala J. *Nat Commun*, 2022, 13: 4904
- 85 Chatterjee G, Dalchau N, Muscat RA, Phillips A, Seelig G. *Nat Nanotech*, 2017, 12: 920–927
- 86 Lv WY, Li CH, Yang FF, Li YF, Zhen SJ, Huang CZ. *Angew Chem Int Ed*, 2022, 61: e202115561
- 87 Famoori F, Molahosseini AS, Zarandi AAE. *IEEE T Nanobiosci*, 2023, 22: 329–336
- 88 Zheng X, Wang B, Zhou C, Wei X, Zhang Q. *IEEE T Nanobiosci*, 2016, 15: 499–507
- 89 Pei Y, Bian T, Liu Y, Liu Y, Xie Y, Song J. *Nano Lett*, 2022, 22: 3003–3010
- 90 Zhang J, Liu C. *ACS Synth Biol*, 2021, 10: 3148–3157

Nucleic acid-based computation and its applications

Xinyi Shen¹, Junlan Liu¹, Da Han^{1,2*}

¹*Institute of Molecular Medicine, Renji Hospital, School of Medicine, Shanghai JiaoTong University, Shanghai 200127, China*

²*Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou 310022, China*

*Corresponding author (email: dahan@sjtu.edu.cn)

Abstract: Nucleic acid-based computation has garnered significant attention as an innovative computing paradigm, distinguished by the vast data storage capacity, superior parallel processing capabilities, compact design, and energy efficiency of nucleic acids. The molecular computing system mirrors the structure of electronic computation systems, encompassing three essential components: signal input, signal processing, and signal output. It has been successfully employed to perform arithmetic and logical operations, data processing, encryption, and the development of analog neural networks. Furthermore, its integration with other nanomaterials has paved the way for the creation of molecular machines, which hold promise for extensive applications across mathematics, medicine, and computer science. This comprehensive review delves into the principles, functionalities, and biomedical applications of nucleic acid computation. It underscores the advancements in molecular machine construction and concludes with an exploration of the challenges that lie ahead.

Keywords: nucleic acid, computing system, molecular machine, nanotechnology

doi: [10.1360/SSC-2024-0027](https://doi.org/10.1360/SSC-2024-0027)