

DNA-based programmable gate arrays for general-purpose DNA computing

用于通用DNA计算的基于DNA的可编程门阵列

Abstract

在过去的几十年里，电子和光子集成电路经历了从专用到可编程的演变^{1,2}。尽管液相DNA电路在算法的编码和执行中具有大规模并行的潜力^{3,4}，但通用DNA集成电路（DIC）的开发尚未得到探索。在这里，我们通过集成基于多层DNA的可编程门阵列（DPGA）来演示DIC系统。我们发现，使用通用单链寡核苷酸作为均匀的传输信号，可以可靠地集成泄漏最小、保真度高的大规模DIC，用于通用计算。具有 24 个可寻址双轨门的单个 DPGA 的重新配置可以通过布线指令进行编程，以实现超过 1000 亿个不同的电路。此外，为了控制分子的固有随机碰撞，我们设计了DNA折纸寄存器，为级联DPGA的异步执行提供方向性。我们通过由三层级联DPGA组装的二次方程求解DIC来证明这一点，该DPGA由30个逻辑门和大约500个DNA链组成。我们进一步表明，DPGA与模数转换器的集成可以对疾病相关的microRNA进行分类。在没有明显信号衰减的情况下集成大规模DPGA网络的能力标志着迈向通用DNA计算的关键一步。

Introduction

利用生物分子相互作用的液相生物计算由于其与生物系统的大规模并行性和内在兼容性而受到积极探索。例如，包括自动机⁵、逻辑电路⁶⁻⁸、决策机^{9,10}和神经网络¹¹在内的计算DNA反应网络已经实现，这些网络在分子信息处理^{4,12}，合成智能设备¹³和生物医学应用¹⁴⁻¹⁶方面显示出潜力。尽管取得了这些进展，但这些计算系统中的大多数都是在硬件上量身定制的，以实现特定的算法或有限数量的计算任务。

通用电子集成电路允许软件编程而不是特定于应用程序的定制硬件制造来执行某种功能，为计算机原型设计提供了一个更高级别的平台，而无需事先了解底层物理。值得注意的是，经典的硅基和新兴的基于碳纳米管的计算机和量子计算机经历了类似的演变，从特定应用（例如，专用集成电路）到通用（例如，现场可编程门阵列，FPGA）^{2,17-20}。可编程性和可扩展性是实现通用计算的两个关键因素。可编程性使设备规范能够执行各种算

法，而可扩展性允许通过向系统添加资源来处理越来越多的工作。与电子集成电路不同，电子集成电路中的栅极是物理定位的，通用电信号以定向方式传输，DNA集成电路

(DIC) 中的生物分子成分在溶液中扩散和混合²¹，这阻碍了可扩展和可编程生物计算设备的发展。在典型的DNA计算系统中，DNA组分正交性的限制以及难以控制分子的固有随机碰撞使得实现通用DNA计算实际上具有挑战性。

人们一直在努力探索DNA电路的可编程性^{22,23}。然而，由于这些液相系统的集成普遍缺乏方向性，可编程DNA系统的可扩展性还有待探索。与电子²⁴或量子²⁵电路中的子组件组装类似，在蜂窝²⁶和合成²⁷⁻²⁹分子反应系统中引入了空间区室以增加方向性，但在系统级别上显示出有限的可扩展性²⁹。受硅基FPGA的启发，其中电子形成用于设备内编程和设备间通信的通用信号，我们在这里开发了高度可扩展，基于DNA的可编程门阵列 (DPGA; 扩展数据图1) 采用通用单链DNA寡核苷酸作为均匀传输信号 (DNA-UTS)。

可扩展的 DPGA 架构

在由DNAUTS支持的可扩展DPGA架构的典型设计中，无法用一个DPGA实现的任务电路被划分为多个子电路，每个子电路映射到一个DPGA并由一个DPGA执行 (图1)。作为电子集成电路指令的模拟，我们建立了一个分子指令集，其中包含大约1,000条指令 (超过2,000个寡核苷酸)，定义了DPGA上的所有合法线路。每个子电路的分子指令与所涉及的计算单元混合，形成实现子电路功能的独特路由模式。设计了四种类型 (AND、OR、NOT 和 XOR) 的 24 个可寻址逻辑门作为计算单元，它们的组合构成了一套完整的布尔函数 (补充图 2)，在单个 DPGA 上提供了高编程空间 (补充文本 6.2)。DPGA的操作基于沿着程序配置的路径在门和DPGA之间接收和发送DNA-UTS。为了避免DPGA之间的串扰，我们进一步设计了一个DNA折纸寄存器来指导级联DPGA的异步计算处理。与其电子对应物类似，从上游DPGA计算的中间值通过DNA链置换写入DNA折纸寄存器，然后将其传输到下游DPGA。

本设计采用双轨输入/输出端口的双轨逻辑门，允许代表高低信号的两条DNA链同时通过 (图2a和扩展数据图3)，以实现DPGA。所有门的输入/输出端口的分子设计的一致性允许DPGA的任意路由和积分，这是基于给定的逻辑函数，然后公式生成相应的DNA序列 (补充文本3.3和3.4)。例如，双轨AND门由两个杂交的DNA分子实现，具有三个特定结构域，它们充当开关：一个结构域 (蓝色) 分别响应任一输入的低信号，而两个串联结构域 (黄色) 分别响应每个输入端口的高信号 (图2b)。在DPGA的实施过程中，DNA链的输入触发链置换反应 (SDR) ³⁰ 以实现逻辑功能 (扩展数据图4)。DPGA 的配置是通过寻址和连接目标电路所需的门来实现的。在这些情况下，配置的DPGA可以抽象为更

高级别的计算单元，其中输入层逻辑门的输入端口和输出层逻辑门的输出端口分别用作DPGA级输入和输出端口（图2a）。

顺序SDR用于实现串联开关，其中第一个输入（in2H）与“S3”混合，后者取代“S5”以暴露in1H的立足点;in1H然后杂交并释放“S4”，其作为输出（图2c）。我们优化了分子设计的结构，发现由S3、S4和S5形成的间隙结构，S3中有一个不成对的碱基，可以抑制泄漏反应途径，并允许较高的计算速度（图2d，e，扩展数据图5和补充文本4.1）。

我们首先测试了单个双轨门，其功能为DPGA编程奠定了基础。归一化高电平（outH）和低电平（outL）输出信号都生成了正确的结果，与所有四种类型的栅极在20分钟内的真值表一致（图2f和补充图25），表明1 nt间隙设计提供的计算速度高于先前报道的大规模DIC中基于SDR的AND门6，31。重要的是，即使在链退火后没有任何纯化（补充图21-23），我们观察到所有栅极的泄漏率非常低（低于0.1），这反映了DPGA的容错能力。为了解释一个给定执行的结果，双轨结果由 outH 和 outL 之间的差异定义（双轨结果， $(outH - outL + 1) / 2$ ）。从outH和outL到双轨结果的转换图如图2g所示。特别是，在所有四种输入组合下，AND门的双轨结果都落在正确的ON和OFF状态区域内。所有类型的双轨栅极的测量显示，对于所有逻辑 FALSE 条件，双轨结果均低于 0.2，在所有 TRUE 条件下，双轨结果均高于 0.8（图 2h）。

在证明了单门内的可编程性之后，我们接下来探讨了DNA-UTS是否可以连接栅间和DPGA间传输以实现计算电路。对于单个DPGA内的栅间传输，通过输入端口进入配置电路的输入值被处理以生成输出值，传输到输出端口进行读出。因此，DPGA编程需要三种类型的导线——输入端口到栅极、栅极到栅极和栅极到输出端口（图3a）。我们建立了一个分子指令集，定义了DPGA上的所有合法导线，包含三种类型的接线指令：类型1指令（WIR1）将输入传导到门;类型 2 指令（WIR2）将89输出信号从上游栅极传导到下游栅极的输入端口;类型 3 指令（WIR3）将输出信号从门传导到 DPGA 的输出端口（图 3b 和补充文本 2.1）。所有合法导线都可以事先合成和组装，允许通过添加相应的导线对 DPGA 进行编程。WIR1 是通过进入入口门的 ssDNA 输入实现的。执行该门后，产生的DNA-UTS作为输出通过WIR2传输到下游门，该门由十条ssDNA链实现（图3b）。我们设计了噪声阈值和信号放大器，用于布线指令的分子实现（补充文本2.2和扩展数据图 6），以抑制沿反应途径传输过程中的信号衰减，这可能会限制DICs的深度27,32。

WIR2的信号转导是通过SDR实现的。陡峭斜坡的展示表明，我们可以有效地抑制泄漏的信号（低于0.4）以接近0，并将真实输出（高于0.6）放大到接近1（图.3d）。类似地，

WIR3用十条DNA链实现，以通过SDR转导信号（补充图4）。具有增加层数的级联闸门被要求测试类型 1、2 和 3 布线指令的协作性（图 3e 和补充图 28-31）。所有输入组合的低输出和高输出信号在1小时内进入正确的状态。双轨结果对于应该是1的人高于0.8，对于那些应该是0的人，低于0.2。因此，我们确定可编程和可靠的内部DPGA DNA-UTS传输可以通过布线指令的子集来实现。

对于 DPGA 间布线，我们采用了第四种类型的布线指令（WIR4）来实现 DPGA DNA-UTS 间传输，即来自上游 DPGA1 的输出信号传输到下游 DPGA2。为了抑制DPGA之间的串扰和信号衰减，每个WIR4都以类似信号中继的方式实现，先是DNA折纸寄存器，然后是WIR2（图3c）。DNA折纸的空间限制和可寻址性33,34控制了溶液中分子相互作用的内在随机性。在典型的WIR4中，DNA折纸寄存器暂时存储DPGA1的输出，然后下游WIR2将输出传输到DPGA2的入口门，为级联DPGA的异步执行提供方向性。输出信号通过SDR写入DNA折纸，SDR由暴露的脚趾（5 nt）介导，该脚趾由另一端具有较长脚趾（7 nt）的检索链读出，并在视觉上（图3f和补充文本5.8）和定量（补充图44和45）得到确认。通过捕获上游DPGA的输出并将存储的信号释放到下游DPGA后分离DNA折纸来实现异步（补充图6）。重要的是，我们发现使用WIR4s有效地改善了电路深度，如由11个具有不同地址的双轨门组成的11层电路的性能所示（图3g-i）。

因此，我们认为使用DNA折纸寄存器对DPGA进行物理分离限制了不匹配的DNA链之间的瞬时结合，这极大地减弱了放大电路尺寸下的信号衰减。此外，WIR4通过将长反应途径划分为较短的反应途径来促进反应动力学，从而以类似信号中继的方式缩短时间延迟。鉴于信号衰减取决于反应途径的长度，使用 WIR4 可提高 DIC 的深度可扩展性。

多任务重新配置

在与DNA-UTS建立了DPGA布线之后，我们接下来探索了DPGA的多任务操作的重新配置。我们开发了一个编译器，将自然语言或布尔表达式中的程序转换为布线指令的子集（扩展数据图7和补充文本7）。以两位乘法函数为例，我们将程序编译为一组布线指令，这些指令称为八个双轨门，其中八个WIR1，六个WIR2和四个WIR3（图4a）。DPGA 的路由模式如图 4a 所示。我们成功地在所有16个可能的输入上测试了这个乘法电路，并获得了正确的计算结果（图4b和补充图37）。动态范围由最高OFF状态值和最低ON状态值之间的差异定义，用于量化计算结果的ON/OFF对比度（补充图64c）。值得注意的是，所有四个输出的动态范围都超过0.5（图4b），表明乘法电路在高ON/OFF对比度下正常工作。

o 使用布线指令评估DPGA多任务重新配置的通用性和鲁棒性，我们通过重新配置单个DPGA（补充表6）对103个电路进行了792次计算，该DPGA实现了各种基本电路结构，包括单栅极，级联栅极，栅极的扇入和扇出以及复杂的功能电路。这些DNA回路的大小范围从具有17条链的单个门到一次反应中具有300多条链的11个门。加法是数字计算机的基本操作。我们通过调用DPGA上的七个门来实现一个两位加法算法（图4c）。三个输出端口为所有 16 种情况生成了正确的结果（图 4d）。我们还演示了使用 DPGA 来实现比较操作，这些操作通常用于逻辑语句中以确定变量之间的相等或差异（图 4e）。计算变量a和b之间数值比较的两位比较功能是通过DPGA上涉及11个双轨门的电路实现的，对于a和b的每个可能配对，这些门都正确执行（图4f）。重要的是，我们在反应后2小时内观察到所有792次计算的正确计算结果（补充图64b）。通过在执行各种功能时使用动态范围和矢量接近来检查DPGA重新配置的鲁棒性，我们证明了即使在涉及11个双轨门的情况下也能实现DIC的可靠计算（图4g-i和补充文本6.1）。通过将 11 个门的电路大小作为实际上限，我们估计单个 DPGA 的编程空间将达到 1000 亿个树形结构电路的独特模式，至少比现有方法高两个数量级（MAYA-III22 低于 100,000，IBC23 约为 10 亿;补充文本6.2）。

为了评估使用单DPGA编程进行无差错计算的能力，我们进一步检查了792个实验中生成结果的分布。尽管 2% 的 outH 和 outL 处于未定义的状态（低于 outH 0.4，outL 低于 0.6），但所有双轨结果都导致基于二值化规则的正确布尔值（图 4i）。因此，双轨栅极的电路表现出高水平的噪声容限。我们展示了一个管中含有多达300条DNA链的电路的大规模计算，具有高DNA浓度（ $1\times = 100\text{ nM}$ ），并且直接使用所有杂交链（退火后）而无需进一步纯化。我们认为，顺序置换反应与配置的DPGA中的阈值和扩增相结合有助于提高DIC在单个反应中的可扩展性（补充图63）。

DPGA 的并行和串行集成

鉴于DPGA的模块化特性，我们进一步测试了可重构的DPGA是否可以集成到更大规模的电路中，以实现实际可访问的计算能力。原则上，可分割成独立部分的算法可以通过并行DPGA电路实现。作为概念验证，我们并行使用了三个不同配置的DPGA实体，将平方根函数操作到一个十进制数字（图5a, b）。在该算法中，一个十进制数字由使用二进制编码十进制码的四位表示。平方根r用定点二进制表示，整数位用两位表示，十进制数字用四位表示。该电路由三个并联配置的DPGA进行分割和实现（图5c）。我们发现，对于所有实验，整个DPGA电路在2小时内计算出结果（图5d和补充图41）。值得注意的是，这个DPGA电路涉及21个双轨门，超过300条DNA链，其中最大的子电路包含130条链。

由于DPGA可以与WIR4集成，因此配置的DPGA可以串行组装以实现高深度电路。为了证明这一点，我们通过条件分支实现了有限状态机（图5e）。从星期一到星期日的星期

几被分配了数字1-7。计算代表第二天的数字的状态机由两个级联DPGA实现，一个作为条件电路来决定输入是否为7，另一个作为处理电路（图5f）。使用基于聚乙二醇（PEG）沉淀的折纸纯化和磁场介导的隔离通过WIR4进行信号传输（补充图51和52）。两种方法都产生了正确的计算结果，在所有七种输入组合下具有几乎相同的性能（图5g和补充图52），这表明DPGA的可扩展性是固有的，并且与解决方案传输方法无关。因此，并行和串行连接DPGA的潜力为DPGA网络的构建铺平了道路，以实现包括任意数字操作在内的复杂算法。

DPGA 的模块化设计允许将多个 DPGA 集成到一个网络中。接下来，我们开发了一个 DPGA网络来求解表达式 $x^2 - bx + c = 0$ 的二次方程，其中b和c是两位整数（图5h）。较大的解 x 可以通过二次方程的根公式计算： $x = \frac{b \pm \sqrt{b^2 - 4c}}{2}$ 。该算法是通过包含 30 个双轨门的 11 层 DIC 实现的。在我们的设计中，电路被划分并映射到具有五个级联DPGA的网络，其中中间值 $s = b^2 - 4c$ 和 $r = \sqrt{s}$ 。该函数通过两个函数实验得到证实： $x^2 - 3x = 0$ 和 $x^2 + x - 2 = 0$ （图 5i, j）。比较WIR4介导的DPGA网络和直接添加理想输入的单个DPGA，我们发现，在存在WIR4的情况下，下游DPGA尽管运行速度降低了大约三倍，但与理想值的偏差很小（图5k和扩展数据图9）。因此，DPGA网络的发展尽管以时间消耗为代价，但仍有可能实现用于高复杂操作的大规模集成电路。

基于 DPGA 的非线性分类器

已经开发了用于鉴定疾病相关生物标志物的分子电路^{15,16}，但通常是线性的，因此在处理来自多个生物标志物的信息时受到限制，而这些生物标志物可能并不总是线性可分离的。我们试图通过利用DPGA的高可积性来构建非线性分类器来解决这个问题。我们针对基于miRNA表达水平的癌症分类器，并使用来自癌症基因组图谱（TCGA）的肾透明细胞癌（KIRC）的miRNA表达数据训练，开发具有mir-200c，mir-204和mir-887三个miRNA输入的决策树回路（图6a和补充文本5.12）。分类器是通过集成DPGA和模数转换器（ADC）来实现的（图6b），后者将模拟单轨miRNA信号转换为双轨DNA-UTS，然后由DPGA进一步处理（图6c）。ADC产生的DNA信号通过WIR4传输到DPGA。以mir-204为模型，我们验证了ADC转换的测试（图6d, e）。

在使用合成 miRNA 分子(18 KIRC 和5个健康样品，在图6f 中由彩色点表示)的非线性分类器的实验测试中，在从 ADC 接受 DNA-UTS 后，DPGA 在2小时内计算了23个测试样品的正确输出 H 和输出 L (图6g 和扩展数据图10)。对分类器的双轨输出的分析表明，分子计算、模型预测和 TCGA 标记的实际疾病状态之间是一致的(图6h)。尽管这些分类器测试涉及 Tris-EDTA (TE)缓冲液中的合成 miRNA，但也可以在细胞培养物中用 DPGA 进行实验(补充图61)，并且 ADC 接收来自含血清溶液的 miRNA 信号(补充图62)，表明应该可以将 DPGA 直接与细胞和体液接口。由于 DPGA 的可编程性和可扩展性允许它们与

不同类型的 ADC 相结合，我们预计集成的 DPGA 将能够处理来自 miRNA 以外的生物标志物的信息，并且可能证明对于临床和诊断应用是有用的。

Discussion

我们的通用 DNA 计算系统通过使用 DNA-UTS 实现了 DPGA 中有效的信号传输，该系统在电子/光子集成电路中起着电子或光子的作用。由此产生的可编程 DIC，而不是简单地实现一个特定的算法，可以执行包含2000多个寡核苷酸的指令集的大规模计算路径，并达到通用计算的复杂度阈值。各种应用，如基于非线性分类的诊断，可以很容易地通过简单地耦合 DPGA 与特定的 ADC 的目标。

DPGA 电路可以通过可编程方式生成以执行各种功能(扩展数据图2)。DPGA 的可编程性是通过使用具有统一输入/输出接口的双轨门来实现的，这使得每个门都能被寻址，类似于 FPGA。门通过 DPGA 和 FPGA 中的地址信息进行连接。接线指令的顺序依赖于上游/下游门的地址，也独立于它们的逻辑功能。另外值得注意的是，DPGA 中的双轨门可以被多次调用(例如，09门在方程求解电路中被使用了5次)，允许电路重新配置而不需要重新设计 DNA 序列。电路的可重用性在功能上等同于 FPGA 中的物理重组。作为 DPGA 的计算单元，这些双轨门的指令与传统的编译语法兼容，为可编程通用 DNA 计算机的软件开发提供了巨大的潜力。特别是，因为所有逻辑门的序列都是正交的，所以 DNA 链可以通过高通量 DNA 合成来大量生产35(补充图29)。

介导时空分离电路之间信号传输的 DNA 折纸寄存器是 DPGA 级集成的关键。它们可以直接在液相电路中传输信号，以集成和调节级联式 dpGA 的异步执行，而牺牲中间数据的手动变速器。DNA 折纸寄存器的使用极大地抑制了限制电路大小的瞬态结合和限制电路深度的信号衰减，从而促进了多层 DPGA 的高阶集成并增加了可以实现的程序复杂性。例如，DPGA 网络与多达5个级联 DPGA，30个逻辑门和大约500个参与链的大规模集成相当于11层逻辑门和级联 SDR 的30个步骤的可实现的电路深度(扩展数据图10)，并突破可编程 DNA 电路的当前限制(补充图57)。尽管基于 DNA 的寄存器设计需要手动干预，但它允许引入具有高复杂性和可扩展性的全 DNA 电路，而不需要其他生物分子组分(例如酶22，光/电子传感器36或微流体29)。

我们相信 DIC 的可编程性和可积分性可以扩展到几个领域。首先，我们的 DPGA 编译器为 DNA 计算的发展提供了一个高度通用和用户友好的接口，它可以自动编译目标函数并生成相应的 DNA 指令，而不需要知道它们的内部实现。其次，DPGA 结合数据寄存器的层次结构可以移植到其他分子反应体系中。例如，可以产生和简化 DNA 链的 DNA 处理酶不是使用 SDR 进行分子实现，而是为实现反馈数字电路提供新的工具5。第三，目前 DPGA 之间传输所需的手动转移可能会进一步自动化，例如，通过将每个 DPGA 封装在

囊泡隔室内²⁹或将每个 DPGA 定位在 DNA 折纸模板上²⁷。磁场辅助方法在微流体系统中可能是可积分的^[36]，并且通过将 DPGA 与基于液滴的微流体耦合可以完全自动执行 DIC ^[37]。

Methods

准备双轨闸门和布线说明书

将 DNA 寡核苷酸溶解在 $1 \times$ TE 缓冲液(无核酸酶, pH 8.0, Sigma-Aldrich)中, 通过监测它们在 260nm 处的吸收并在 -20°C 下储存, 用紫外/可见光谱法定量。用荧光染料或猝灭剂标记的寡核苷酸溶解在去离子水(Milli-Q)中, 并在 -20°C 的去离子水中保存。单链杂交制备相应的双链 DNA。为了抑制由合成和定量误差引起的潜在泄漏, 我们对那些注定要释放的物质使用较低浓度的链。通过将所需的链与 TE 缓冲液($1 \times$ TE: 40mM Tris 碱基, 20mM 乙酸, 调整至 pH8.0的2mM EDTA)中的相应终浓度混合, 制备杂交结构, 12.5 mM MgCl_2 。最终浓度见补充表1。

接线指令的阈值和放大器复合物是通过在浓度比为1:1.2的情况下分别退火来制备的, 最终浓度见补充表2。通过将链与猝灭剂和荧光团混合制备记者, 其中前者比后者多50%, 分别达到最终浓度15和10 μM 。所有实验和退火配合物的缓冲液为 TE 和12.5 mM Mg^{2+} 。混合股通过加热到 95°C 退火2分钟, 在保持在 4°C 之前以每6s 0.1°C 的速度缓慢冷却到室温。杂化分子储存在 4°C 以备进一步使用。

DNA 反应动力学的荧光测量

使用 Synergy H1杂交多模式阅读器(BioTek)和 Corning 96孔黑色测定板进行荧光测量。所有的电路计算都在仪器上并行进行。根据实验的整体持续时间, 每1或2分钟收集一次荧光动力学数据。激发(发射)波长为510纳米(540纳米)的染料 TET 和640纳米(670纳米)的染料 Cy5。一般来说, 除了输入链以外的所有电路元件都与12.5 mM MgCl_2 混合在 TE 缓冲液中。实验在96孔黑色测定板(Corning)中进行, 每孔98 μl 反应混合物用于所有实验。初始值记录为基线。然后暂停实验, 加入2 μl 输入链, 随后通过摇动混合。然后在混合阅读器中更换平板, 实验继续进行。对于单个 OR, AND, NOT 和 XOR 门, 两层和三层电路, 扇入和扇出电路, 四层电路和阈值复合物, 以40nM ($0.4 \times$)的标准浓度以100nM ($1 \times$)的标准浓度进行实验, 对于全减法器, 两位乘法, 两位加法, 两位比较操作和平方根操作, 标准浓度为60nM ($0.6 \times$)。对于布线指令中的放大器, 使用 $2 \times$ 杂化分子和 $10 \times$ 燃料。对于双轨逻辑电路, 使用两个不同的荧光团——猝灭对 TET-BHQ2和 Cy5-BHQ2

——在500nM (5 ×)的标准浓度下读取输出轨迹。输入的标准浓度为200nM (2 ×)。对于每个输入组合，我们同时记录两个分别代表高信号和低信号的荧光通道。将每对实验的输出轨迹结合到一个单一的图中，可以同时观察不同输入组合产生的输出。在整个反应过程中，温度保持在18 °C。

DNA 折纸寄存器的制备与操作

根据 Rothmund 的方法³⁸，从短纤维和输出结合链和 M13m18ssDNA (NEB)组装90 × 60nm²矩形 DNA 折纸结构。在补充图49,54和55中，主链和输出结合链的位置和序列用不同的颜色表示。在100μl 系统中，我们使用50nM 短链，100nM 输出结合链和10nM M13 ssDNA 形成折纸。在热循环(Bio-Rad)中，将 DNA 折纸退火并在1 × Tris-醋酸 -DNA (TAE)-Mg²⁺ + 缓冲液(Tris 40mM，醋酸20mM，EDTA 2mM 和乙酸镁12.5 mM，pH8.0)中组装，具有以下设置：在95 °C 孵育2分钟，在12s/0.1 °C 下缓慢冷却至60 °C，在60 °C 孵育12分钟，在12s/0.1 °C 下缓慢冷却至25 °C，然后在4 °C 保留长达24小时。组装的矩形 DNA 折纸结构使用 PEG 沉淀从过量的短纤维和输出结合链中分离出来³⁹。简而言之，将20mM MgCl₂的 DNA 折纸结构与含有15% PEG8000w/v (no. MW: 8000, Sigma)，5mM Tris，1mM EDTA 和505mM NaCl。将溶液在1.5 ml 离心管中混合，并在离心机(Eppendorf)中以12,000 rpm 和4 °C 旋转15分钟。用移液管去除上清液。将沉淀溶解于20μl 1 × TAE-Mg²⁺ + 缓冲液中，并在40 °C 和400rpm 下温育过夜。纯化的矩形 DNA 折纸的浓度用微量紫外可见光谱仪(NanoDrop)定量。我们在室温下孵育2小时，然后在4 °C 下储存，直到进一步使用，在折纸中加入5倍多的块状链。

用 TET-BHQ2记录器记录自由状态输出的浓度，以测试 DNA 折纸记录器的书写和读取效率。(1)将输出链(100nM)与 DNA 折纸(10nM)一起温育，具有一个含有21个位点的结合区域和500nM 报道基因在室温下孵育2小时。(2)将输出链与 DNA 折纸一起孵育2h，然后在室温下加入1μM 的检索和报告基因，再孵育2h。(3)将输出链与报告基因孵育2小时，测定各反应的荧光强度。

级联 DPGA 的执行

将上游 DPGA 的所有电路元件混合并温育2小时以完成反应。对于基于聚乙二醇沉淀的转移，将 DNA 折纸寄存器加入到最终浓度为10nM 并再孵育1小时。然后使用 PEG 沉淀纯化折纸，然后用 NanoDrop 定量，并加入到下游 DPGA 的反应体系中，最终浓度为10nM。WIR4中相应的 WIR2也被加入并温育2小时，然后使用相同的信号传递操作将其输出转移到下一个 DPGA。对于磁场介导的转移，将生物素修饰的 DNA 折纸寄存器与链霉亲和素修饰的磁珠(New England Biolabs，4mg ml⁻¹)温育以形成 MB 折纸寄存器。在

上游 DPGA 孵育后，将 MB 折纸寄存器加入到10nM 的最终浓度。在温育1小时后，将 DPGA 输出写入 DNA 折纸寄存器，施加磁场并去除上游 DPGA 溶液。然后加入回收链以回收储存的输出，输出转移到下游 DPGA 反应系统。接收到输出信号后，将下游 DPGA 孵化2小时，并使用相同的信号传输操作将其输出转移到下一个 DPGA。用荧光记录仪记录反应动力学。

11层电路的执行和结果

该11层电路由11个具有不同地址的双轨门组成，便于用一个 DPGA 实现电路或分成多个 DPGA。首先，我们从左到右增加电路深度，并在一个单一的反应系统中执行每个计算。我们发现，当电路深度达到五层时，计算速度显著降低，信号下降到泄漏水平(图3h)。此外，双轨结果随着电路层的衰减。计算结果显著偏离正常的 ON 或 OFF 状态在五层深度(图3i)。接下来，我们将电路分成三个子电路，每个子电路用一个配置的 DPGA 实现，其中的信息通过 WIR4传输。我们发现，当电路深度增加到11层时，结果仍然在正确的范围内(图3i 和补充图48)。进一步的分析表明，WIR4，特别是内部 DNA 折纸寄存器，主要通过限制瞬态结合和时间延迟来改善可实现的电路大小和深度(扩展数据图8和补充图47)。

数据归一化与双轨结果计算

所有数据在绘制时都从原始荧光水平标准化为输出信号的相对浓度，尽管仪器性能、电路功能和分子实施存在差异，但有助于对数据进行定量分析。微板读数器支持多达96个平行动力学实验，在这些平行实验中，由仪器引起的荧光读数的差异可以忽略不计。每组并行实验针对的是同一电路，但输入不同。最小水平(输出0)由时间 $t = 0$ 时所有测试数据点的最小值确定。对于给定的荧光团，并行实验具有至少一个增加的输出信号(即最大 ON 完成水平)；最大水平(输出1)由最高信号的最后五个数据点的平均值确定。图1e 所示的荧光数据以这种方式标准化。对应于标准浓度($1 \times$)的荧光水平是从平板上记者产生的最高信号中获得的。当以 $t = 0$ 连接任何输入链/链时，可忽略的浓度($0 \times$)对应于反应混合物的背景荧光，所述输入链/链是从平板上记者产生的最低信号的第一次测量获得的。所有在单个平板上的实验都归一化在一起，这样就可以直接比较不同输入模式下电路的输出。我们使用同一电路的两个不同的记录器读取高信号输出 H 和低信号输出 L，从而得到的输出 H 和输出 L 的信号分别归一化。在正常化后，使用这个方程计算双轨结果：

$$\text{Dual-rail result} = \frac{\text{out}_{11} - \text{out}_1 + 1}{2}.$$

Error flags were calculated using equation:

$$\text{Error flag} = \frac{||\text{out}_{11} - \text{out}_1| - 1|}{2} + \frac{||\text{out}_{11} + \text{out}_1| - 1|}{2},$$

其中 | ... | 表示内部内容的绝对值。

原子力显微镜成像

在多模 VIII 原子力显微镜(Bruker, Inc.)上用峰力模式对 DNA 矩形折纸进行成像。在使用峰值力流体尖端扫描样品之前,将大约30 μ l TAE-Mg2 + 缓冲液加入到液体细胞中以渗透尖端。为了在 DNA 折纸寄存器上写入输出链,将该输出链和输出链以1:10摩尔比混合并温育1小时,然后 PEG8000纯化以除去未结合的链。将回收链(100 \times)加入到培养液中培养4小时,以取代存储的输出链。使用链霉亲和素(Sigma)可视化 DNA 折纸寄存器上生物素修饰的输出链的储存过程,其在写入之前,写入之后和数据检索之后与链霉亲和素以1:10的比例混合并在室温下温育30分钟。最后,将5 μ l 混合溶液沉积在新鲜切割的云母表面上并温育3分钟。用 TAE-Mg2 + 缓冲液洗涤样品5次后进行 AFM 成像。

分子动力学模拟

采用粗粒度模型 oxDNA40,通过分子模拟获得了泄漏机理的数据。OxDNA 是自顶向下参数化的,并将每个核苷酸描述为具有六种各向异性相互作用的位点:排除体积,堆叠,交叉堆叠,氢键,骨架连接和静电排斥。在这里,我们使用了更新的 oxDNA2力场和显式的静电学⁴¹。

六个模拟系统的初始结构最初以 PDB 格式获得,然后使用 TacoxDNA webserver⁴²导出到 oxDNA 格式。结构分两步放松。蒙特卡罗模拟使用了 DNA 松弛力场和进一步松弛使用最大骨架力选项在一个分子动力学模拟与 DNA 2力场。在这两个过程中,基于预期设计的相互陷阱被应用来加强对预期设计的松弛。然后释放外力和骨干力限制,并使用相同的力场进行生产模拟运行。为了缩短仿真时间,在间隙为0、1和2 nt 的系统中引入了互陷效应。监测 T (21)和 A (51)(距离1), C (10)和 G (62)(距离2)和 C (0)和 G (72)(距离3)之间的距离(补充图16)。

所有模拟系统的模拟都是在300K 的正则 NVT 集合中,在40.87 nm 的周期性立方盒中使用类似 Anderson 的恒温器进行的。整合的模拟时间步长为15.15 fs (0.005 oxDNA 时间单位),分子动力学步长设定在3 \times 10⁸至2 \times 10⁹之间,足以研究每个系统的泄漏机制。

为了加速扩散动力学和改善取样，粒子的平移扩散系数被设置为2.5，比实验快了大约两个数量级。类安德森恒温器的牛顿步长为103，每 1×10^4 步保存一次构型用于分析。本文所有模拟的盐条件设定为单价 NaCl 浓度为1M NaCl⁴⁴。使用一套 oxDNA 分析脚本 (`distance.py`，`bond_analysis.py` 和 `contact_map.py`)⁴⁵来详细探索泄漏机制。

决策树训练和验证

我们使用来自 TCGA KIRC 数据集的 miRNA-seq 数据来训练决策树分类器。我们将521个 KIRC 和71个健康样本随机分为训练集和测试集，比例为8:2。决策树分类器以每百万读数的碱基2对数的 miRNA 表达值进行训练，并获得每个 miRNA 的阈值(训练集包括414个 KIRC 和59个健康样品)，以使用 Python 的 Scikit-learn (Sklearn)对癌症和健康组进行分类。该分类器使用一个测试集(包括107个 KIRC 和12个健康样本)进行验证。将数据库中 mi204、200c 和887的表达值缩放到0-100nM。DPGA 利用这些 miRNA 输入执行概念验证的非线性分类。