

Der Computer von morgen: DNA statt Silizium?

Computer of the Future: DNA Instead of Silicon?

S. Lorkowski und P. Cullen

Zusammenfassung: Computer auf DNA-Basis, die Nukleinsäuren als Speichermedium und molekularbiologische Techniken für ihre Rechenoperationen verwenden, sind seit einigen Jahren keine reine *Science Fiction* mehr. Im Gegensatz zu herkömmlichen binären Computern auf Siliziumbasis, die als Entscheidungsalternativen nur die Zustände Ein und Aus besitzen, können mit DNA-Computern Rechenoperationen parallel und aufgrund der vier verschiedenen DNA-Bausteine mit vier Entscheidungsmöglichkeiten (G, A, T und C) durchgeführt werden. Diese Fähigkeit ermöglicht DNA-Computern die Lösung mathematischer Probleme, die mit konventionellen Computern nicht zu lösen sind. Mathematische Puzzles, bei denen die Zahl der Lösungswege – und damit die Rechenzeit – exponentiell mit der Zahl der einzubeziehenden Variablen zunimmt, stellen für einen parallel arbeitenden Computer im Gegensatz zu einem binären Computer kein Problem dar, da seine Bearbeitungszeit nur linear mit der Zahl der Variablen zunimmt. Auch wenn es derzeit noch zahlreiche ungelöste technische Schwierigkeiten gibt und die Durchführung der Rechenoperationen sehr arbeitsintensiv ist, sind DNA-Computer dennoch der erste Schritt zu Computern, die Rechenoperationen auf molekularer Ebene ermöglichen.

Schlüsselwörter: DNA-Computer; DNA-Microchips.

Summary: Computers based on the use of nucleic acids as storage media and molecular techniques for computations are no longer science fiction. Compared with standard computers based on silicon chips, which work with binary logic (on and off as possible decisions), DNA computers are parallel processing machines, which can use four possible decisions depending on the four compounds of nucleic acids (G, A, T, and C). This ability of DNA computers make it possible to solve mathematical problems which cannot be solved using conventional computers. Mathematical puzzles with exponentially growing numbers of solutions depending on the linear increase in variables can be solved by DNA computers, because the computation time of a parallel computer compared with a binary computer increases as a linear function of the

number of variables. There are still a lot of technical problems to be solved and the procedures needed for computable operations of a DNA computer are labour- and time-intensive. Nevertheless, DNA computers represent the first step towards computers with the possibility of computing directly with molecules.

Keywords: DNA computer; DNA microchips.

Siliziumchips sind seit mehr als 20 Jahren die Grundlage für die Entwicklung immer leistungsfähigerer Computer. Die Rechenleistung eines Computers zu erhöhen, setzt jedoch voraus, daß die einzelnen Einheiten eines Chips immer weiter verkleinert und immer mehr dieser Bausteine auf einem immer kleinerem Raum untergebracht werden können. Aus physikalischen Gründen (z.B. aufgrund der Quantenunsicherheit und der Wellenlänge der zur Herstellung notwendigen elektromagnetischen Strahlung) kann diese Miniaturisierung jedoch nicht unbegrenzt fortgesetzt werden. Die Möglichkeit, Rechenoperationen auf molekularer Ebene durchzuführen, könnte dieses Problem jedoch lösen und die Computertechnologie revolutionieren.

Nukleinsäuremoleküle, wie DNA und RNA, sind natürliche Informationsträger, die die Anforderungen, Rechenoperationen auf molekularer Ebene durchzuführen, erfüllen können. Computer auf Nukleinsäurebasis können im Gegensatz zu herkömmlichen Rechnern parallel arbeiten und haben vier Entscheidungsmöglichkeiten anstelle der zwei Alternativen Ein und Aus eines binären Systems.

Erste Schritte zur Entwicklung eines Computers auf DNA-Basis wurden bereits unternommen. *Leonard M. Adleman* von der südkalifornischen Universität war einer der ersten Wissenschaftler, der sich mit der Lösung mathematischer Probleme mit Hilfe von Nukleinsäuren und molekularbiologischen Techniken auseinandersetzte [1]. In der Zeitschrift *Science* publizierte er 1994 eine Lösung für eine Variante des mathematischen Puzzles „Problem eines Handlungsreisenden“ („*Hamiltonian path problem*“). Seither ist die Vision eines Computers auf DNA-Basis keine reine *Science Fiction* mehr, sondern eine mögliche, greifbar erscheinende Zukunftsvision.

Das mathematische Problem, einen Reiseweg zu vorgegebenen Orten zu finden, ohne einen der Orte mehrmals aufzusuchen, stellt selbst leistungsfähigste

Korrespondenzadresse: Stefan Lorkowski, Institut für Arterioskleroseforschung, Domagkstrasse 3, 48149 Münster. Fax: +49-251-8356-20-5. E-mail: lorkost@uni-muenster.de
Eingegangen: 11. Mai 2000 / Angenommen: 11. Mai 2000

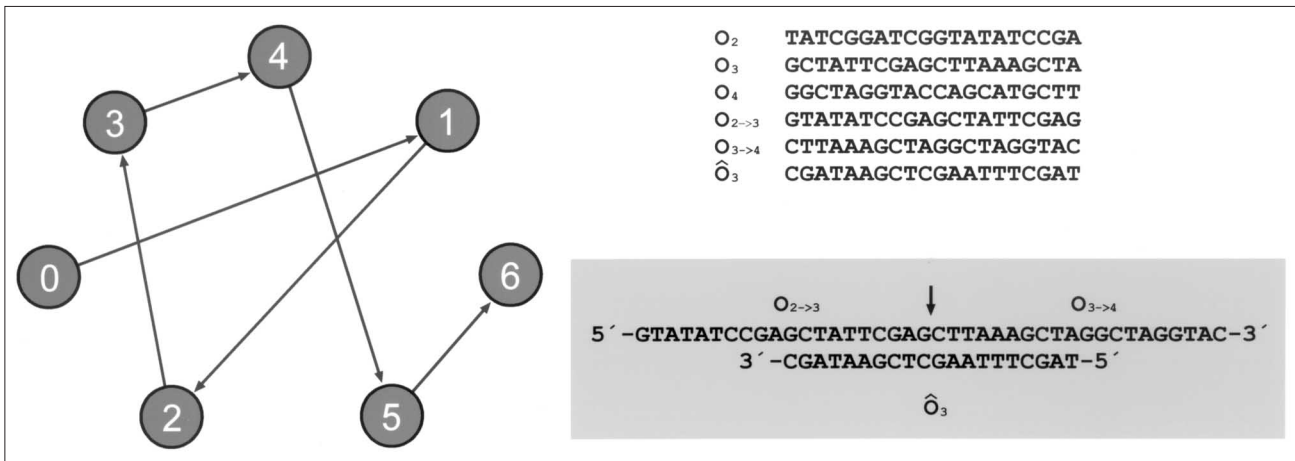


Abbildung 1 Links: Eine mögliche Lösung des Puzzles „Problem eines Handlungsreisenden“ für eine Reiseroute mit sieben Orten ist (mit $i = 0$ als Startpunkt und $i = 6$ als Endpunkt): $0 \rightarrow 1$, $1 \rightarrow 2$, $2 \rightarrow 3$, $3 \rightarrow 4$, $4 \rightarrow 5$, $5 \rightarrow 6$. Rechts: Für jeden der Reiseorte i wird ein Oligonukleotid O_i verwendet (dargestellt sind O_2 , O_3 und O_4 für die Orte 2, 3 und 4). Für die Strecke zwischen zwei Orten werden Oligonukleotide $O_{i \rightarrow j}$ ausgewählt (dargestellt sind die Oligonukleotide für die Strecken zwischen den Orten 2 und 3 sowie den Orten 3 und 4). Die Oligonukleotide O_i der einzelnen Orte überlappen wie dargestellt mit den Oligonukleotiden $O_{i \rightarrow j}$ der angrenzenden Reiseroute, so daß während der Hybridisierung die komplementären Oligonukleotide \hat{O}_i mit den Oligonukleotiden $O_{i \rightarrow j}$ angrenzender Strecken hybridisieren können.

Computer auf eine harte Bewährungsprobe. Allen bekannten Algorithmen, dieses Problem zu lösen, ist zu eigen, daß die Zahl der Lösungswege – und damit die Rechenzeit – exponentiell mit der Zahl der zu erreichenden Orte (den Variablen) zunimmt. Daher ist es selbst für Computer der heutigen Generation schwierig, Lösungen für Probleme dieser Art zu finden. Für einen parallel arbeitenden Computer nimmt die Arbeitszeit jedoch nur linear mit der Zahl der Variablen zu.

Adleman verwendete zur Lösung dieses Puzzles DNA-Oligonukleotide O_i aus 20 Basen mit einer zufällig ausgesuchten Sequenz. Diese Oligonukleotide O_i repräsentierten die einzelnen Orte i der Reise. Anschließend wurden Oligonukleotide $O_{i \rightarrow j}$ mit 20 Basen konstruiert, die den Wegen zwischen zwei Reisepunkten entsprachen. Diese Oligonukleotide $O_{i \rightarrow j}$ waren derart gestaltet, daß die ersten zehn Basen ihres 5'-Endes identisch mit den letzten zehn Basen des 3'-Endes der Oligonukleotide O_i waren und die letzten zehn Basen ihres 3'-Endes mit den ersten zehn Basen des 5'-Endes der Oligonukleotide O_j übereinstimmten (vgl. Abb. 1). Anschließend wurden die zu den Oligonukleotiden O_i komplementären Oligonukleotide \hat{O}_i mit den Oligonukleotiden $O_{i \rightarrow j}$ gemischt. Die Hybridisierungsprodukte (vgl. Abb. 1) wurden miteinander ligiert und die entstandenen Ligationsprodukte repräsentierten mögliche Reiserouten des Handlungsreisenden.

In einem nachfolgenden Schritt wurden die Ligationsprodukte in einer PCR amplifiziert, um die Lösungen des Puzzles zu selektieren, die nicht den Ort 0 als Startpunkt und den Ort 6 als Endpunkt der Reiseroute aufwiesen. Dazu wurde der Primer für den Ort 0 (Oligonukleotid O_0) und der zur Sequenz des Oligonukleotids des Ortes 6 komplementäre Primer \hat{O}_6 ver-

wendet. Die PCR-Produkte wurden in einem Agarosegel separiert. Banden mit der korrekten Größe (eine Reiseroute mit sieben Koordinaten entspricht z.B. einem PCR-Produkt von 140 bp) wurden eluiert. Anschließend wurde das Eluat schrittweise mit einem Biotin/Avidin-Beadsystem aufgereinigt. Dazu wurden die Oligonukleotide \hat{O}_i an Beads gekoppelt. Das Eluat wurde sukzessive mit den sieben Fraktionen der gekoppelten Oligonukleotide \hat{O}_i aufgereinigt, so daß am Ende nur noch PCR-Produkte im Eluat verblieben, die alle Sequenzabschnitte O_i enthielten.

Abschließend wurde das Eluat in i PCR-Ansätzen amplifiziert, wobei in einer einzelnen Reaktion der Primer O_0 und ein Primer \hat{O}_i verwendet wurde. Die PCR-Produkte wurden erneut in einem Agarosegel aufgetrennt. Im aufgeführten Beispiel mit einer Reiseroute zu sieben Orten müssen in der Amplifikation ausschließlich Produkte mit einer Größe von 40, 60, 80, 100, 120 und 140 Basen entstehen. Die Lösung des Problems kann durch eine Sequenzierung der größten Bande bestimmt werden.

Diese Publikation hat zahlreiche Wissenschaftler zu neuen Ansätzen und lebhaften Diskussionen angeregt. *Eric B. Baum* stellte bereits 1995 Überlegungen an, wie ein Speichermedium bzw. eine Datenbank auf DNA-Basis konkret zu entwerfen wäre [2]. *Richard J. Lipton* von der Universität Princeton, USA, konnte nicht nur eine Lösung des mathematischen „SAT-Problems“ mit Hilfe eines auf DNA basierenden Algorithmus finden [3], sondern konnte unter der Mitarbeit von *Dan Boneh* und *Christopher Dunworth* eine theoretische Methode entwickeln, das sogenannte „data encryption standard system“ (DES), ein Kodierungsverfahren, das von der *National Security Agency* entwickelt worden ist, dekodieren zu können. Der Lösungsansatz, der aufgrund der 2^{56} möglichen Lösungen

von einem binär arbeitenden Computer nicht zu bearbeiten ist, könnte mit einem auf DNA basierenden parallelen Computer berechnet werden. *Guarnieri et al.* stellten 1996 einen Algorithmus vor, mit dem sie mit Hilfe von DNA-Molekülen zwei rationale, nicht negative binäre Zahlen addieren konnten [5]. 1997 stellte eine Arbeitsgruppe um *Albert Libchaber* eine Lösungsstrategie für eine allgemeinere Variante des von *Adleman* bearbeiteten Problems vor, die in kürzerer Zeit die Bearbeitung von mehr als sieben Variablen ermöglichte [6].

Die Entwicklung von DNA-Computern übt auch derzeit wieder Faszination aus. Zwei Veröffentlichungen in diesem Frühjahr in den Zeitschriften *Nature* und *Proceedings of the National Academy of Sciences* beweisen dies. Eine Arbeitsgruppe um *Laura F. Landweber* und *Richard J. Lipton* versuchte mit Hilfe eines neuen Ansatzes, ein fundamentales Schachproblem zu lösen [7]. Das sogenannte Springerproblem beinhaltet die Frage, in wie vielen Positionen Springer auf einem Schachbrett platziert werden können, ohne sich gegenseitig angreifen zu können. Auch bei diesem mathematischen Problem nimmt die Zahl der Lösungen exponentiell mit der Zahl der Variablen, in diesem Fall der Zahl der Felder auf dem Schachbrett, zu.

Da das Springerproblem mit einem komplexen mathematischen Ansatz nur schwierig zu lösen ist, konstruierte *Laura Landweber* mit ihrem Team ein Rechnersystem aus 1024 Nukleinsäuresträngen mit verschiedenen Kombinationen von RNA-Basen. Jeder RNA-Strang repräsentierte eine mögliche Anordnung der Springer auf dem Schachbrett. Um das Rechenexempel zu vereinfachen, reduzierten die Wissenschaftler das Spielfeld auf neun Felder (3 × 3 Felder). Dadurch reduziert sich die Zahl der möglichen Kombinationen der beiden Springer zueinander von einigen Millionen bei einem konventionellen Spielfeld auf 512 mögliche Anordnungen auf dem kleinen Spielfeld.

Die Arbeitsgruppe stellte zunächst eine RNA-Bibliothek her, die RNA-Moleküle enthielt, die alle möglichen Anordnungen der Springer auf dem Schachbrett repräsentierten. Ein Teil dieser Bibliothek, wurde mit DNA-Molekülen hybridisiert, die der Anwesenheit des Springers auf dem Feld *a* des 3 × 3-Schachfeldes entsprachen (vgl. Abb. 2). Gleichzeitig wurden die durch die Anwesenheit eines Springers auf dem Feld *a* notwendigen Abwesenheiten der Springer auf den Feldern *f* und *h* durch Zugabe von entsprechenden DNA-Molekülen bedingt. Nach einer Behandlung mit Ribonuklease H zum Verdau der entstandenen RNA/DNA-Hybride waren somit die Anordnungen *a/f* und *a/h* der Springer als falsche Lösungen eliminiert.

In einem zweiten Ansatz wurden DNA-Moleküle, die der Abwesenheit eines Springers auf dem Feld *a* entsprachen, mit einem Teil der RNA-Bibliothek hybridisiert. Nach der Behandlung mit RNaseH verblieben in diesem Reaktionsansatz alle möglichen Positionen des ersten Springers auf den acht restlichen Positionen. Die beiden Reaktionsansätze wurden miteinander vermischt, damit wieder alle verbleibenden Posi-

tionen der Springer in einem Ansatz vereint waren. Um die verbleibenden möglichen Anordnungen zu eliminieren, mußten diese Reaktionen nacheinander für die Felder *b*, *c*, *d* und *f* mit dem verbliebenen kombinierten Ansatz wiederholt werden. Nach einer anschließenden Amplifikation der verbliebenen Nukleinsäuren wurden die PCR-Produkte in Gelen elektrophoretisch aufgetrennt. Daraus resultierten spezifische Bandenmuster, die die Lösungen des Problems kodierten und direkt ausgelesen werden konnten. Von den 94 möglichen Lösungen konnten *Landweber et al.* mit ihrem Verfahren 42 finden, wobei ein einziger Fehler zu verzeichnen war.

Auch eine vollständige Lösung des Springerproblems ist mit dem von *Landweber et al.* vorgestellten RNA/DNA-Computer möglich, sofern der Ansatz in einem größeren Maßstab durchgeführt werden würde. Dies ist notwendig, da nicht alle Anordnungen der Springer mit gleicher Wahrscheinlichkeit auftreten und daher in einigen Versuchsansätzen (= Rechenoperationen) unterrepräsentiert sind.

Ein Forschungsteam um *Lloyd M. Smith* an der Universität Wisconsin/Madison, USA, konnte einen ersten DNA-Computer herstellen, der mit Nukleinsäuren arbeitet, die auf einem festen Substrat immobilisiert sind (8). Dies könnte der erste Schritt zu einem DNA-Computer sein, der mit elektronischen Komponenten zur Eingabe der Rechenoperationen und Ausgabe der Rechenergebnisse verbunden werden kann.

Lloyd M. Smith und seine Mitarbeiter konstruierten einen DNA-Computer, der sich mit dem sogenannten „*satisfiability problem*“ (SAT-Problem) auseinandersetzte. Dieses stellt prinzipiell das gleiche mathematische Puzzle dar, wie das Springerproblem, nur in einer abstrakteren Form. Analog zu der RNA-Bibliothek, die von *Laura Landweber* und ihren Kollegen verwendet

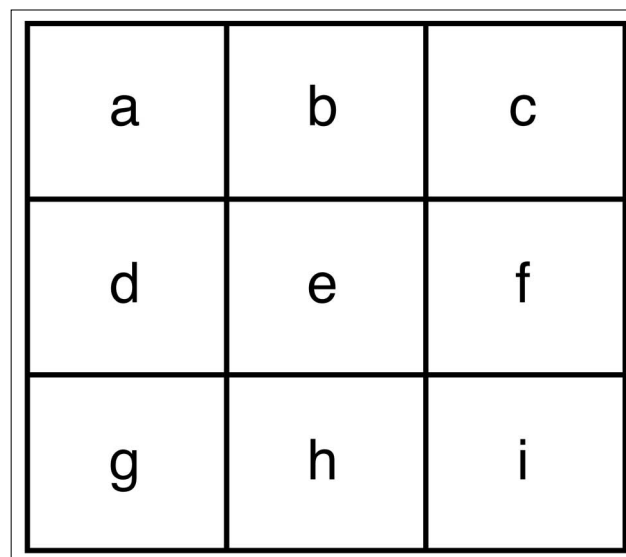


Abbildung 2 Schematische Darstellung des Schachbretts zur Veranschaulichung der Lösung des Springerproblems mit Hilfe eines DNA/RNA-Computers nach *Faulhammer et al.* (7)

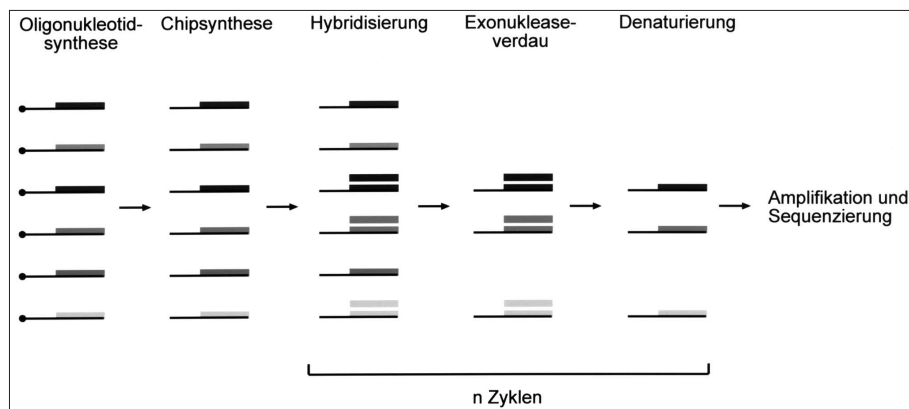


Abbildung 3 Schematischer Ablauf einer Rechenoperation mit einem an einem Substrat immobilisierten DNA-Computer nach Liu et al. (8)

wurde, konstruierten *Smith et al.* eine Bibliothek aus DNA-Oligonukleotiden, die die möglichen Antworten – sowohl falsche als auch richtige – ihres Problems darstellten. Die Oligonukleotide wurden als Proben analog zu einem DNA-Mikroarray zur Sequenz- oder Expressionsanalyse auf einer Goldoberfläche aufgebracht. Zur Eliminierung der falschen Antworten wurden Oligonukleotide synthetisiert, die richtige Antworten repräsentierten und komplementär zu den entsprechenden DNA-Molekülen in der DNA-Bibliothek waren.

Zunächst wurde der DNA-Array mit einer Lösung von DNA-Molekülen hybridisiert, die einen Teil der richtigen Lösungen in Form von komplementären Oligonukleotiden enthält. Die Teillösungen, die sich aus diesem ersten Kriterium ergaben, wurden durch Ausbildung von DNA/DNA-Hybriden detektiert. Mit Hilfe einer einzelstrangspezifischen Exonuklease wurden die verbliebenen einzelsträngigen Nucleinsäuren auf dem Chip, die falsche Lösungen repräsentierten, gespalten und entfernt. Durch Erhitzen wurden die restlichen Doppelstränge in die Einzelstränge denaturiert. Der Array wurde erneut mit einer Lösung aus DNA-Molekülen hybridisiert, der DNA-Oligonukleotide enthielt, die Lösungen anhand eines zweiten Kriteriums repräsentierten. Auch hier wurden die verbliebenen Einzelstränge gespalten. Der Array wurde nach der erneuten Denaturierung für weitere Rechenzyklen verwendet, bis der Chip mit allen DNA-Molekülen, die aufgrund weiterer Kriterien Lösungen des Problems darstellten, hybridisiert worden war. Nach der letzten Denaturierung befanden sich nur noch DNA-Moleküle auf dem Chip, die allen Kriterien entsprochen hatten, also in jedem Hybridisierungszyklus (Rechenzyklus) einen komplementären DNA-Strang gefunden hatten. Die verbliebenen Nucleinsäuren (die Lösungen des Puzzles) konnten amplifiziert und anschließend durch Sequenzierung identifiziert werden.

Das größte Problem derartiger DNA-Computer ist laut *Lloyd Smith* die Synthese der zahlreichen Oligo-

nukleotide, die bereits für einfache „Rechenoperationen“ benötigt werden. Doch bereits heute besteht die Möglichkeit, die Synthese großer Mengen von Oligonukleotiden automatisiert durchzuführen. Auch die Herstellung von DNA-Arrays und die nachfolgenden Arbeiten sind automatisierbar.

Auch wenn die Leistungsfähigkeit von DNA-Computern noch weit von der Leistung heutiger Rechner mit Siliziumtechnologie entfernt ist, stellen die bislang publizierten „Rechenmaschinen“ doch interessante und vielversprechende Ansätze dar.

Der große Vorteil von Nucleinsäuren als Speichermedium ist zweifelsohne ihre hohe Informationsdichte, die dem eines herkömmlichen Siliziumchips und derzeit gebräuchlichen Speichermedien bei weitem überlegen ist. Ein Gramm purer DNA enthält annähernd so viele Informationen wie 10^{12} herkömmliche CD-ROMs. Durch die Entwicklung neuer künstlicher DNA-Bausteine könnte die Leistungsfähigkeit der DNA-Computer weiter gesteigert werden, da ihnen dann mehr als vier Entscheidungsalternativen für Rechenoperationen zur Verfügung stünden.

Die Zukunft wird daher zeigen müssen, ob eine verbesserte und vereinfachte Chemie entwickelt werden kann, die leistungs- und konkurrenzfähige Computer auf DNA-Basis – und damit Rechenoperationen auf molekularer Ebene – möglich machen wird.

Literatur

1. Adleman LM. Molecular Computation of Solutions to Combinatorial Problems. *Science* 1994; 266:1021-4.
2. Baum EB. Building an Associative Memory Vastly Larger Than the Brain. *Science* 1995; 268:583-5.
3. Guarnieri F, Fliss M, Bancroft C. Making DNA Add. *Science* 1996; 273:220-3.
4. Ouyang Q, Kaplan PD, Liu S, Libchaber A. DNA Solution of the Maximal Clique Problem. *Science* 1997; 278: 446-9.
5. Faulhammer D, Cukras AR, Lipton RJ, Landweber LF. Molecular Computation: RNA Solutions to Chess Problems. *Proceedings of the National Academy of Sciences* 2000; 97(4):1385-9.
6. Liu Q, Wang L, Frutos AG, Condon AE, Corn RM, Smith LM. DNA Computing on Surfaces. *Nature* 2000; 403:175-9.