



DNS-számítógép

Balló Gábor



Bevezetés

A nukleinsavak az élő szervezetek egyik legfontosabb alkotórészei. Ezekben tárolódnak ugyanis az öröklődéshez, és a fehérjeszintézishez szükséges információk.

Bár a DNS szerkezete már több mint 50 éve felfedezésre került, általános célú információátvitelre és feldolgozásra alkalmas eszközként csak az 1990-es években kezdték tanulmányozni. 1994-ben bebizonyosodott, hogy az DNS-t lehetséges matematikai problémák megoldására is használni. A módszer nyilvánvaló előnyei az adattárolási kapacitás, és a számítás sebessége.

A DNS szál a bináris adatokhoz hasonló formában, négy bázissal (nukleotiddal) kódolja az információt, ezek az adenin, a timin, a citozin, és a guanin, röviden A, T, C, és G. Mivel ezek a bázisok 0.35 nanométerenként helyezkednek el a DNS molekulában, az adatsűrűség körülbelül 20 TB lehet négyzetcentiméterenként.

A szilícium alapú számítógépek az adatokon csak sorosan képesek dolgozni, azonban egy DNS alapú számítógépben minden molekula egyszerre kezelhető. Ez a párhuzamos működés adja a módszer igazi erejét, hiszen így oldhatók meg akár NP nehéz problémák is reális időn belül.

Mindezt látva, eleinte a legfőbb kérdés az volt, hogy vajon képes lehet-e a DNS alapú számítástechnika felváltani a jól megszokott, szilícium alapú rendszereket.

A számítás módjai

A DNS alapú számításokat különböző módokon lehet kategorizálni:

- ☞ Hibridizációs logikán alapuló számítások: Ebben az esetben a számolási feladatok korrekt bázispárok kialakításával való kódolását, és megoldását, valamint az inkorrekt bázispárok elkerülését jelenti.
- ☞ Nukleinsav alapú logikai hálózatok: Itt a nukleinsavakon ható enzimek segítségével végzünk különböző műveleteket (erre látunk később példát).

A számítás folyamata lehet generatív és szubtraktív.

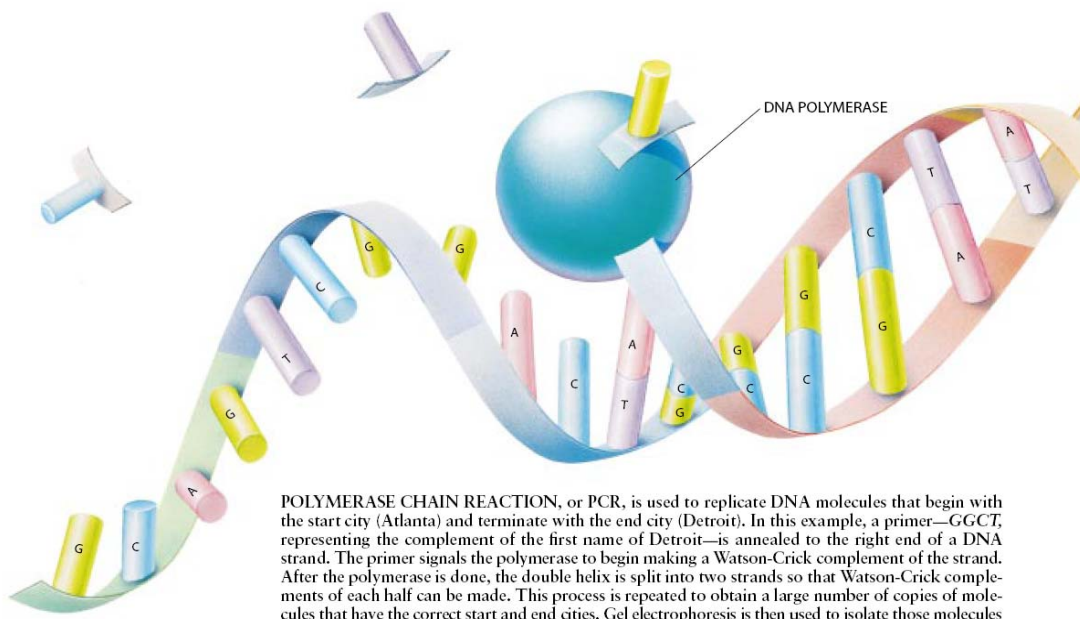
A generatív azt jelenti, hogy a megoldást kódoló bázissorozatokat kisebb oligonukleotid darabokból építjük fel a molekuláris biológia eszközeinek segítségével.

A szubtraktív módszer pedig az inkorrekt bázissorozatoknak az összes sorozat halmazából való eltávolításán alapszik.

Alapműveletek, eszközök

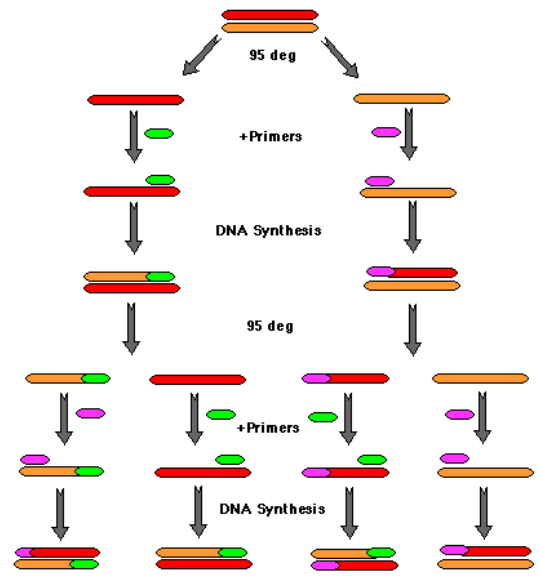
A molekuláris biológia, és így a DNS-számítások eszközei olyan enzimek, amelyek képesek megváltoztatni a DNS láncot, azaz valamilyen műveletet végezni rajta. Néhány példa:

- ☞ Denaturálás: A kettős DNS láncokban melegítés hatására a gyenge H-kötések felbomlanak, és a lánc különálló egyszálú láncokra esik szét.
- ☞ Komplementer láncpárok összetapadása (renaturáció vagy hibridizáció): Ha két olyan DNS lánc találkozik, amelyeken hosszabb komplementer szakaszok vannak, akkor ezek a láncok megfelelő hőmérséklet mellett a komplementer szakaszok mentén összetapadnak hidrogén-kötésekkel, kialakítva egy kettős spirál szerkezetet.
- ☞ Szintetizálás: A DNS láncok mesterséges előállítását jelenti, így kb. 100 nukleotid hosszú szálakat hozhatunk létre.
- ☞ Polimeráz enzimek: Ezek az enzimek képesek egy egyszálú DNS-lánchoz hozzáépíteni a komplementensét. Ehhez azonban szükségük van egy kezdő pozícióra, ahonnan elkezdhetik az építést. Ezeket a pozíciókat nevezzük primereknek. Lényegében olyan kis szakaszok a DNS láncon, ahol már összekötődtek a megfelelő bázispárok. Ezeket a kezdeményeket képesek a polimeráz enzimek folytatni.

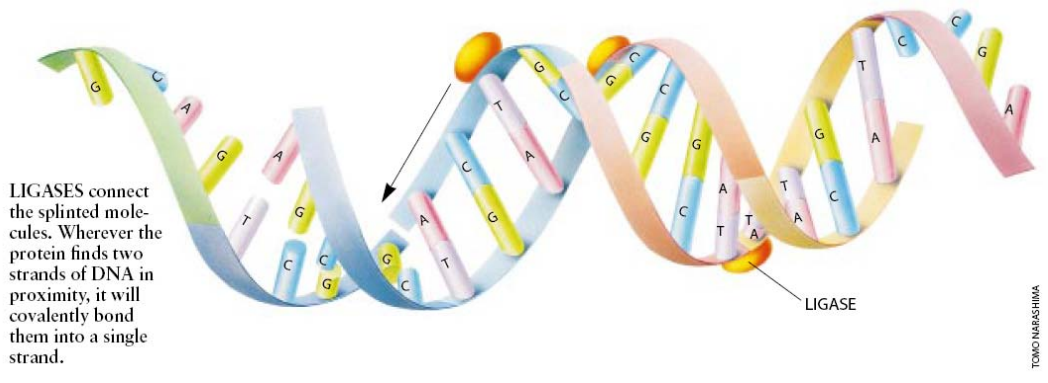


POLYMERASE CHAIN REACTION, or PCR, is used to replicate DNA molecules that begin with the start city (Atlanta) and terminate with the end city (Detroit). In this example, a primer—GGCT, representing the complement of the first name of Detroit—is annealed to the right end of a DNA strand. The primer signals the polymerase to begin making a Watson-Crick complement of the strand. After the polymerase is done, the double helix is split into two strands so that Watson-Crick complements of each half can be made. This process is repeated to obtain a large number of copies of molecules that have the correct start and end cities. Gel electrophoresis is then used to isolate those molecules that have the right sequence length of 24.

☞ Sokszorozás polimeráz láncreakcióval (PCR): Ezzel a módszerrel kis mennyiségű DNS molekulát sokszorosíthatunk. Lényege, hogy előbb melegítéssel szétválasztjuk a kettős DNS láncokat egyszálú részekre, majd a kívánt bázissorozatokhoz kötő primereket keverünk az oldatba. Ez után polimeráz enzimmel felépítjük a szálak komplementerét.

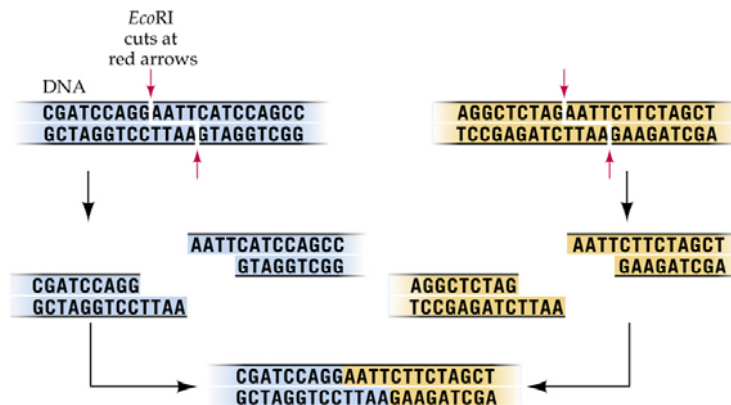


☞ Ligáz enzimek: A ligáz enzimek két különálló láncot kötnek össze kovalens kötéssel. A sejtek ezt az enzimet a roncsolódott DNS-darabok újbóli összeillesztésére használják.

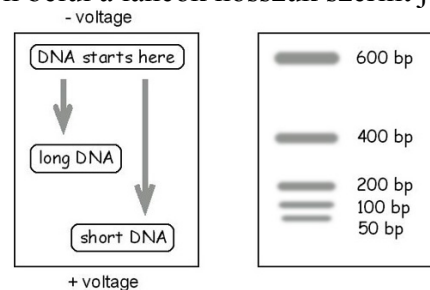


LIGASES connect the splinted molecules. Wherever the protein finds two strands of DNA in proximity, it will covalently bond them into a single strand.

- ☞ Nukleáz enzimek: Ezek a nukleinsavak szétvágására, darabolására képesek. Például, az EcoRI restriktációs endonukleáz a G bázisok mellett vágja el a DNS láncot. Ezeket az enzimeket a baktériumok használják vírusok ellen. A saját DNS-ük védve van, de a baktériumba bejutó vírusoké feldarabolódik.



- ☞ Gél-elektroforézis: Ez egy eljárás, amivel a DNS láncokat a hosszuk szerint lehet szétválasztani. A molekulákat egy gél tetejére tesszük, amibe áramot vezetünk. Ennek hatására a negatív töltésű DNS molekulák elindulnak az anód irányába. Mivel a rövidebbek gyorsabban haladnak, mint a hosszabbak, rövid időn belül a láncok hosszuk szerint jól elkülöníthetőek lesznek.



Hamilton utak keresése

Kérdés, hogy ezek a műveletek, és a DNS, mint információhordozó, elegendőek-e egy univerzális számítógép építéséhez?

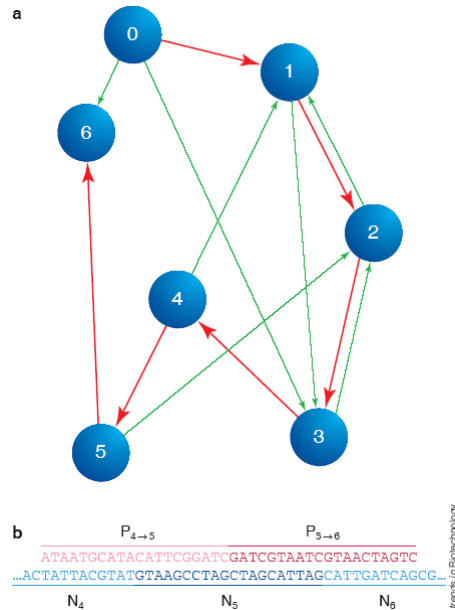
Leonard Adleman professzor úgy gondolta, igen. Ő alkotta meg az első DNS-számítógépet 1994-ben az Egyesült Államokban, hogy feltételezését bizonyítsa. Egy olyan problémát szeretett volna megoldani, amely jól tükrözi a DNS alapú számítások előnyeit is, így a Hamilton utak keresésének problémáját (vagy más néven utazó ügynök problémát) választotta. Ez a feladat lényegében arról szól, hogy egy irányított gráf összes csúcsát úgy kell végigjárnunk, hogy mindegyiket pontosan egyszer érintjük. Mivel a feladat NP-nehéz, azaz a megoldásához szükséges idő exponenciálisan növekszik a gráf csúcsainak számával, Adleman alkalmasnak vélte az új számítási módszer erejének kipróbálására. Egy ilyen feladat megoldása 50 csúcs esetén még a mai leggyorsabb számítógépekkel is évekig tartana, hiszen nincs rá gyors algoritmus.

Adleman professzor 7 csúcsra oldotta meg a problémát a következő lépésekben:

1. Véletlenszerű utak generálása.
2. Azoknak az utaknak a kiválasztása, melyek a megadott kezdőcsúcsból indulnak, és a megadott végcsúcsban végződnek.
3. Annak vizsgálata, hogy mely utak haladnak át pontosan 7 csúcson.
4. Minden csúcsra azoknak az utaknak a kiválasztása, melyek áthaladnak az adott csúcson.
5. Ha ezek után maradnak utak, akkor azok a megoldások.

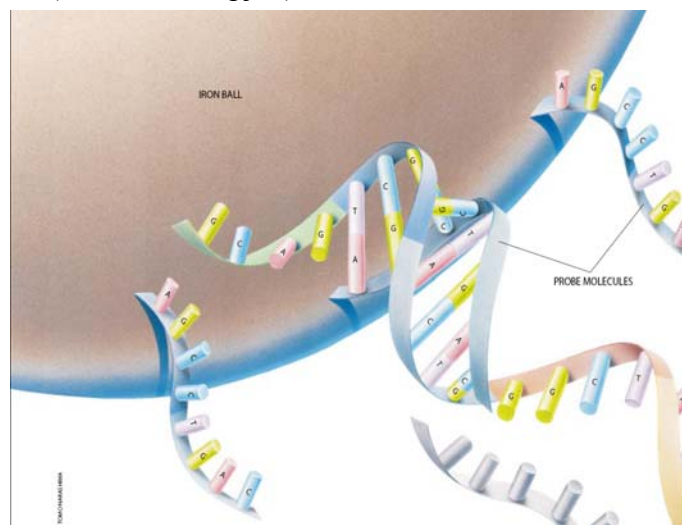
Ez ugyan nem egy tökéletes algoritmus, de ha elég sok utat generálunk, akkor jó eséllyel kapunk helyes eredményt.

Adleman a gráf minden csúcsához hozzárendelt egy 20 nukleotid hosszúságú egyszeres DNS láncot. Ezek a 20 hosszú láncok olyan 10 hosszúságú láncokból épülnek fel, amelyek egyediek, és komplementer láncaik is különböznek mindegyiktől. Hasonlóképpen az élekhez is 20 hosszú egyszeres láncokat rendelt, ezeket a csúcsok láncainak segítségével hozta létre olyan módon, hogy ha két csúcs közt vezet él, akkor a második csúcs első tíz nukleotidjának komplementer láncához kapcsoljuk az első csúcs második tíz nukleotidjának megfelelő komplementer láncot. Így ha két csúcs között él van, létre tudnak hozni egy olyan kétszálú molekulát, amiben az egyik szál a csúcsokat tartalmazza, a másik pedig az éleket.



Miután így az összes lehetséges útvonal kombinációt sikerült létrehozni, már csak ki kellett választania azokat, amelyek megfelelnek a feltételeknek, azaz hogy minden csúcs pontosan egyszer szerepel bennük. Erre a PCR módszert használta. Hogy a megfelelő láncokat szaporítsa, primerként először a kezdő csúcs nevét alkotó bázispárokat használta, majd pedig a célcsúcsét. Ebből következően a molekulák között a 2. feltételnek megfelelőek lettek többségben.

Ez után gél-elektroforézissel megvizsgálta, hogy mely molekulák állnak pontosan a megfelelő számú bázisból (a hosszuk alapján).



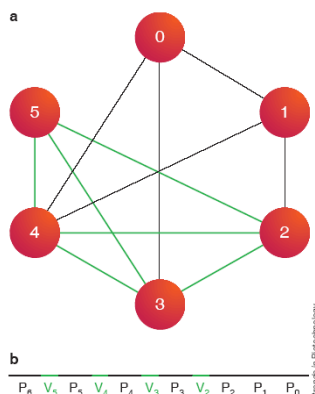
Végül azt, hogy a láncban szerepel-e egy adott csúcs kódja, kis fémgömbök segítségével vizsgálta meg. Ezekhez hozzá lehet kapcsolni a kérdéses csúcs kódjának komplementerét, majd az

oldatba helyezve őket, azok a láncok, amelyekben az adott csúcs szerepel, hozzákapcsolódnak a gömbhöz, így egy mágnessel ki lehet őket venni. Ezt a folyamatot ismételve eljuthatunk a megoldásig.

Adleman a kísérlet végén a maradék DNS láncokat megvizsgálva arra jutott, hogy azok valóban a megoldást kódolják.

Egyéb feladatok

Egy másik probléma megoldására is használhatók DNS molekulák, ez a szintén NP-teljes, maximális teljes-részgráf keresés problémája. Teljes-részgráfnak nevezzük azt a részgráfot, melynek minden csúcsából vezet él az összes többi csúcsba. A legnagyobb ilyen részgráf a maximális teljes-részgráf.



A gráf csúcsait ebben az esetben is oligonukleotidokkal kódoljuk, amelyek különböző módon való összekötődései csúcsok halmazait, a lehetséges teljes-részgráfokat jelölik. Ezek közül ki kell választani azokat, amelyek helytelen, nem létező élet tartalmazznak, vagy túl rövidek, ekkor a maradék a megoldást fogja tartalmazni.

Próbálkoztak még további problémák megoldásával is, ilyen pl. a SAT probléma, melyben egy logikai kifejezésről kell eldönteni, hogy kielégíthető-e. Ennek egy speciális esetét RNS segítségével tanulmányozták, ami jól mutatja, hogy a DNS-en kívül más hasonló molekulák is használhatók számításokban.

Fizikai akadályok

A számítások nukleinsavakkal való kivitelezése ugyan gyors, de vannak hátrányai is.

Az összes lehetséges megoldás előállításához nagyon sok memóriát, azaz DNS láncot igényel. Ez azt jelenti, hogy a feladat komplexitása ugyanúgy exponenciális növekedést mutat, ám ez most nem a számítási időben, hanem a szükséges DNS mennyiségben mutatkozik meg. Hamilton utak keresése esetén például 200 csúcs vizsgálatához szükséges DNS tömege már meghaladná a Föld tömegét.

A módszer másik hátránya a pontossága, és a hibák gyakori előfordulása. Ezt ugyanis a folyamatban résztvevő enzimek nagyban meghatározzák. Adleman kísérletében ez még nem okozott hibát, mivel ő viszonylag kevés molekulával számolt. Egy általános célú számítógépnek azonban nagyságrendekkel többet kellene számolnia, ami megnövelné a hibák számát is.

Alternatív lehetőségek

Mindezt látva, ma már belátható, hogy a DNS-számítógép a közeljövőben nem válthatja föl a megszokott szilícium alapú rendszereket. Ez viszont nem jelenti azt, hogy ne lenne használható, a sejteken belüli alkalmazás továbbra is reális lehetőség.

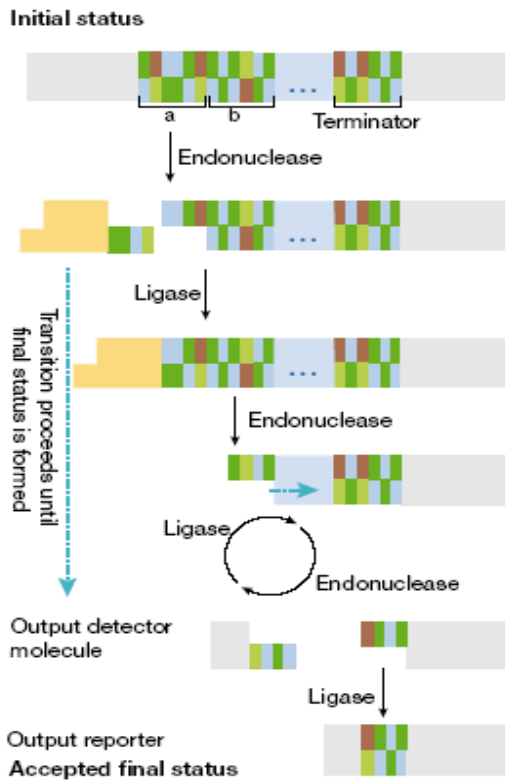
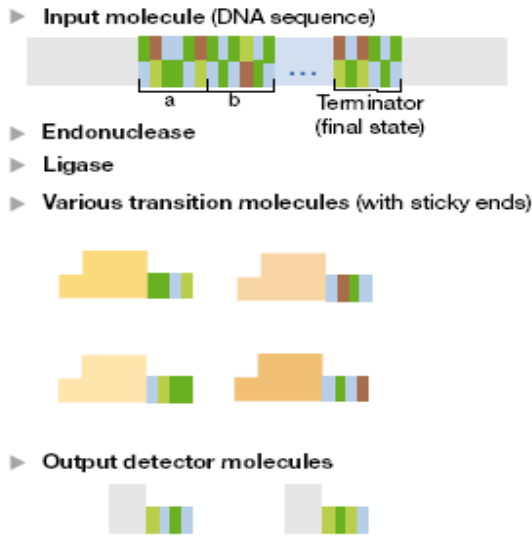
Erre egy példa az Ehud Shapiro által fejlesztett, biomolekulákból épített véges automata,

ami a Turing-géphez hasonló elven, egy véges szimbólumsorozaton végez műveleteket. A gép a bemenet, és az aktuális állapot függvényében, az átviteli szabályoknak megfelelően változtatja az állapotát, amíg végállapotba nem kerül, vagyis amíg az összes bemenetet fel nem dolgozta.

A gép restrikciós endonukleázt és ligázt használ az állapotok megváltoztatására, a bemenet, és az átviteli szabályokat pedig kettős DNS láncon tárolja. Ilyen módon, Shapiro célja további fejlesztésekkel egy teljesen működőképes molekuláris Turing-gép létrehozása. Ez jelentős eredmény lenne, hiszen a Turing-gép bármilyen matematikai műveletet képes elvégezni.

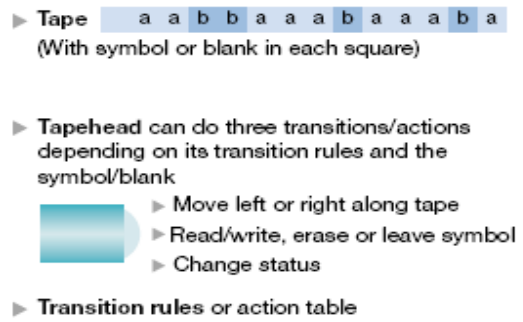
Ezzel a technológiával olyan molekuláris gépeket lehetne létrehozni, melyek képesek felismerni a sejtekben zajló folyamatokat, és célmolekulákat készíthetnek ezek befolyásolására. Nagy jelentősége lenne ennek a betegségek elleni harcban, mivel a sérült sejteket fel lehetne ismerni, és ki lehetne javítani.

Molecular finite automaton



(Design of a molecular finite automaton Benenson *et al.*, 2001, *Nature*)

Turing machine



Example

Status	Symbol	Transition rules
Status 0	a	Move right, change to status 1
Status 0	b	Move right, ...
Status 1	a	..., ...
⋮		..., Halt

