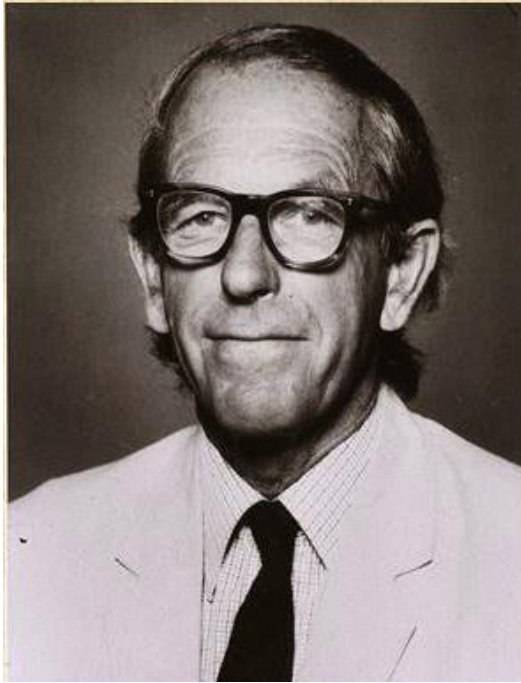


バイオ分析化学特論(14)

1分子シーケンサー

竹中繁織

九州工業大学 物質工学研究系 応用化学部門



Frederick Sanger

1918/8/13～2013/11/19

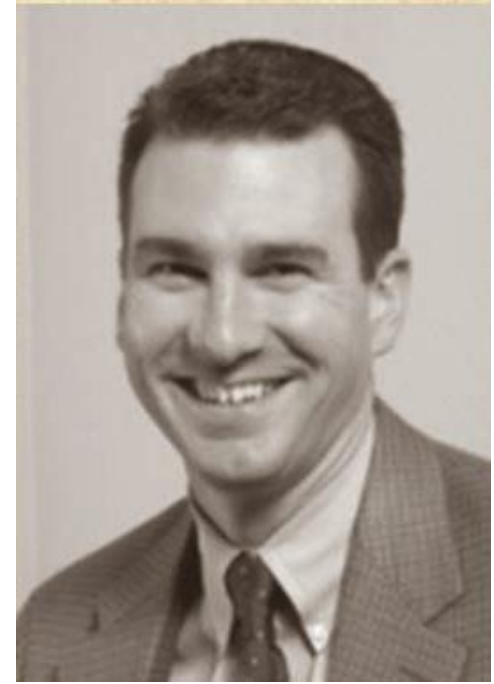
1958&1980ノーベル賞



Walter Gilbert

1932/03/21～

1980ノーベル賞



Allan Maxam

1942/10/28～

化学的分解法 (Maxam-Gilbert法)

(例) 次の13残基のオリゴヌクレオチドの構造決定を行うこととする。

5' ATCGATTCTCGGA 3'

1) まず、このオリゴヌクレオチドの5'末端を³²Pや³³Pで放射標識する(●)。

5' ●-ATCGATTCTCGGA 3'

2) この試料溶液を4等分し、次の4種の化学分解を行う。

切断される塩基	G	G + A	C	C + T
反応液の組成	³² P-DNA DMA緩衝液 ジメチル硫酸	³² P-DNA 水 ギ酸	³² P-DNA NaCl溶液 ヒドラジン	³² P-DNA 水 ヒドラジン
反応条件	20°C, 4.5分	20°C, 10分	20°C, 8分	20°C, 8分

この条件では、特定の塩基のところでランダムに切断が起こり、反応液中に以下のような放射標識断片が生じる。(標識されていない断片もたくさんできるが、検出できないので考えなくてよい)

Gで分解

- ATC
- ATCGATTCTC
- ATCGATTCTCG

Cで分解

- AT
- ATCGATT
- ATCGATTCT

G + Aで分解

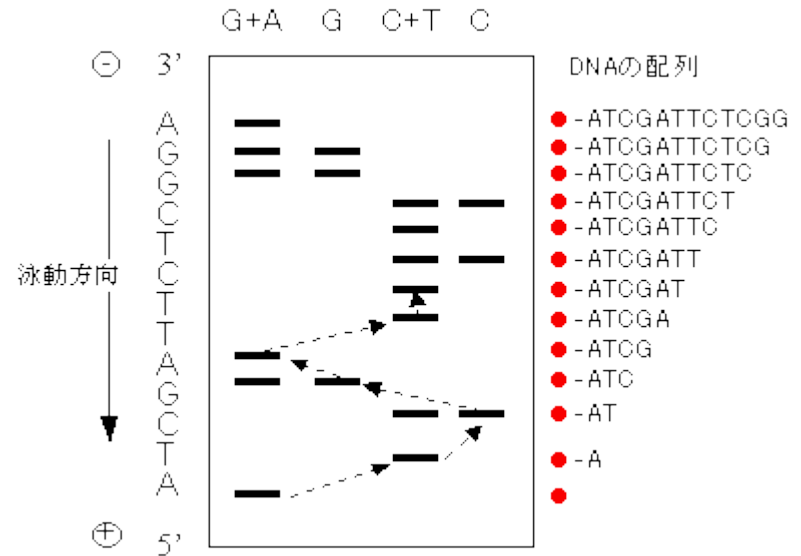
-
- ATC
- ATCG
- ATCGATTCTC
- ATCGATTCTCG
- ATCGATTCTCGG

C + Tで分解

- A
- AT
- ATCGA
- ATCGAT
- ATCGATT
- ATCGATTCT

これら4つの試料をポリアクリルアミドゲル電気泳動にかける。各々のDNA断片はサイズだけで分離される。

4) ゲルのオートラジオグラムを作成する(X線フィルムに露光すると、例の場合、図のようになるであろう。



5) 図の泳動先端(図の下端)のバンドから順に上にバンドの位置をたどっていくことで、目的DNA断片の5'→3'方向の配列を決定できる。(ただし、C+TとCの同じ位置にバンドがある場合はCとなり、C+TにはあるがCにバンドがない場合はTとなる。G+Aについても同様。)

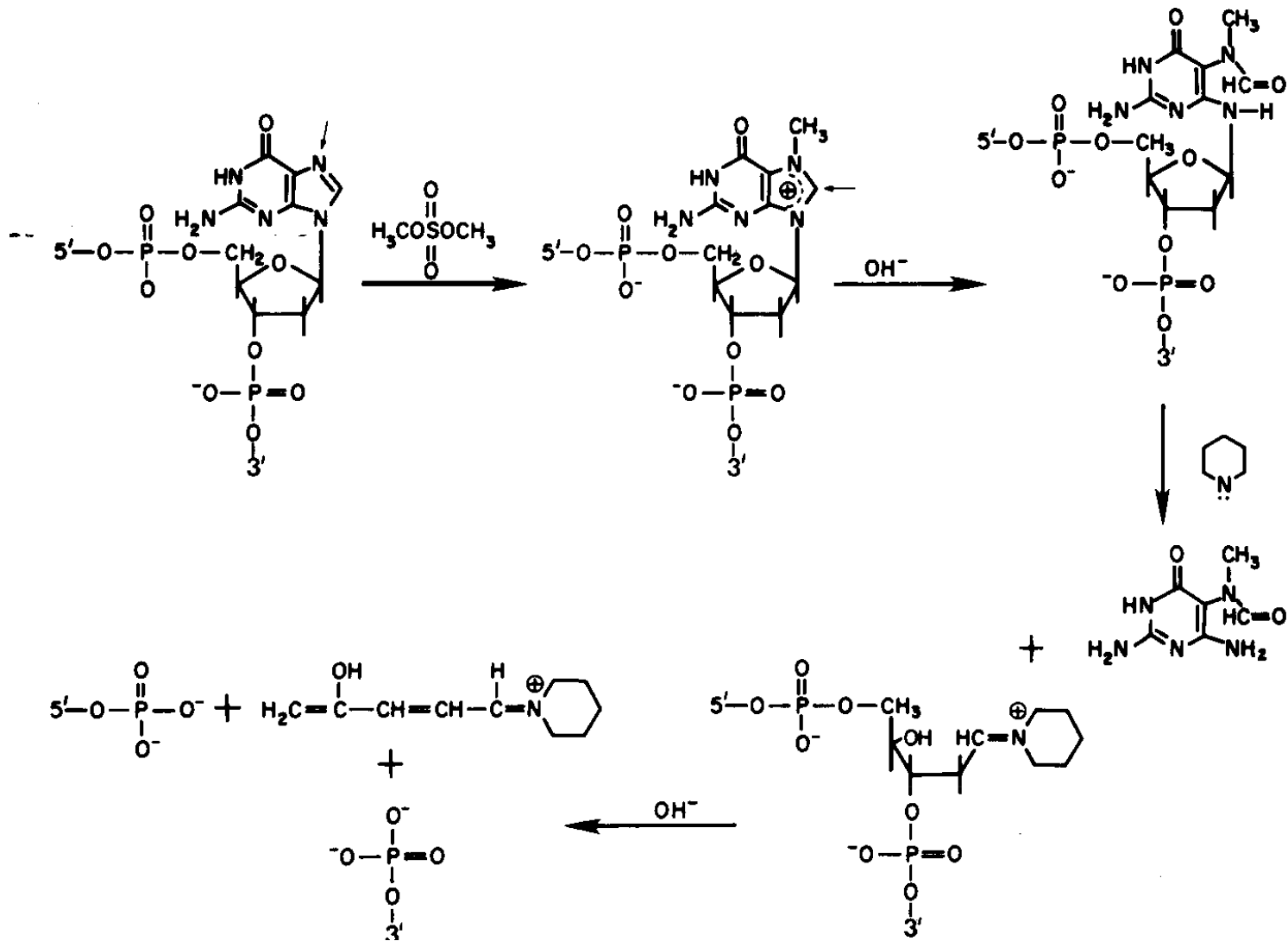


Figure 6. Examples of the detailed chemistry involved in breaking the DNA, Figure 61. above, shows the guanine breakage. The guanines are first methylated with dimethylsulfate. The imidazole ring is opened by treatment with alkali (during the piperidine treatment). Piperidine displaces the base and the triggers two beta eliminations that release both phosphates from the sugar and cleave the DNA strand leaving a 3' and a 5' phosphate.

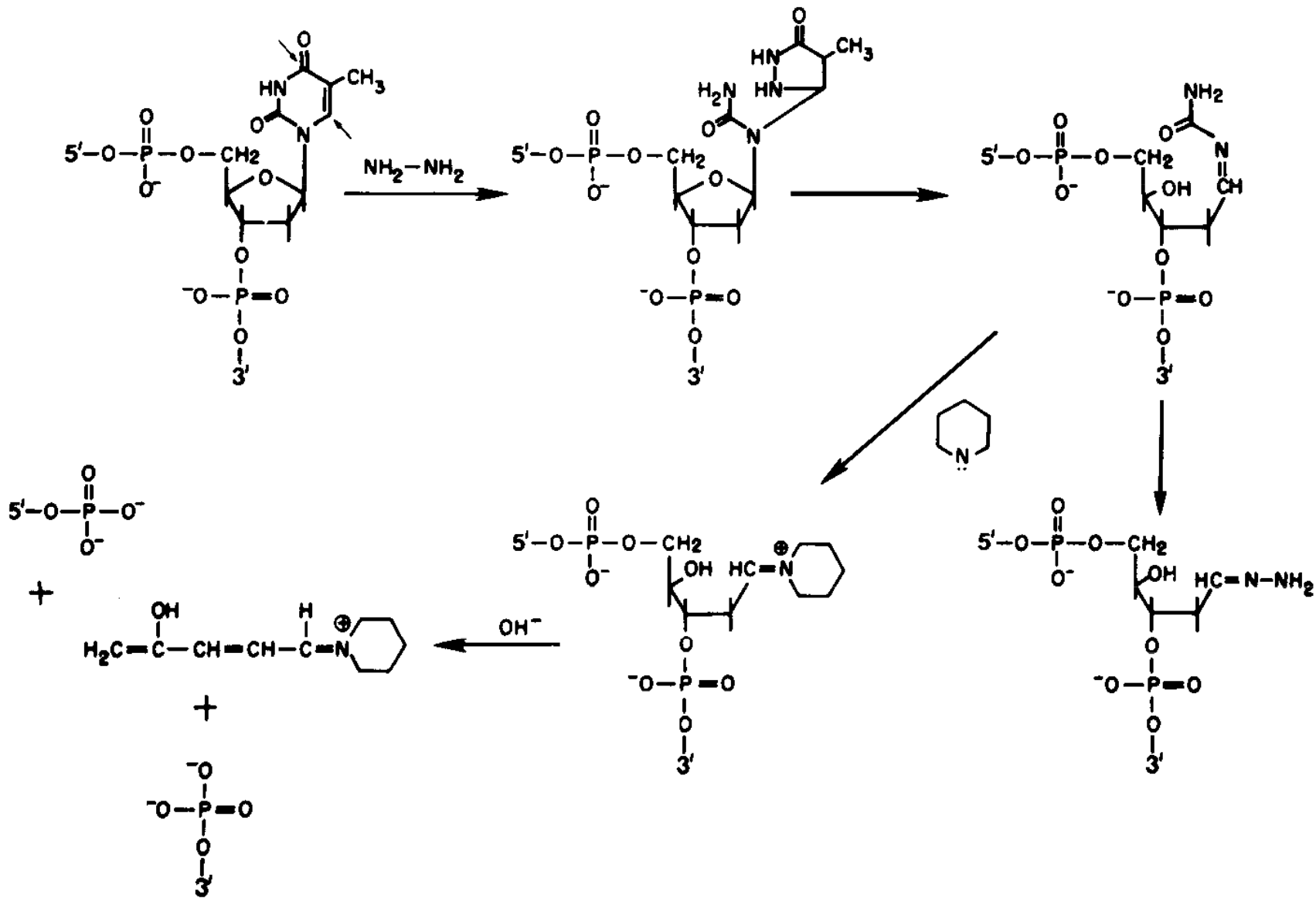


Figure 6. Examples of the detailed chemistry involved in breaking the DNA, Figure 6b. Shows the hydrazine attack on a thymine that breaks the DNA at the pyrimidines.

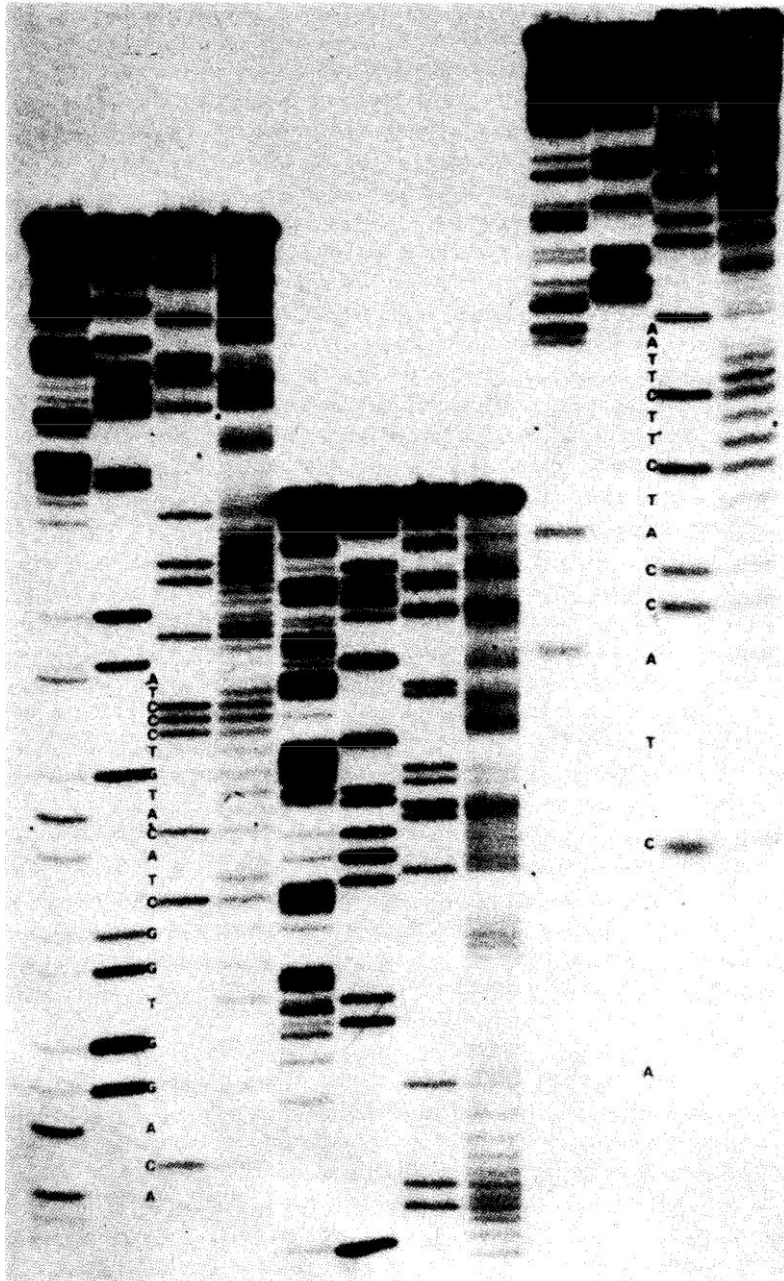


Figure 5. Actual sequencing pattern from the 1978 period. Products of four different chemical reactions, applied to a DNA fragment about 150 bases long, are electrophoresed on a polyacrylamide gel; three loadings produce sets of patterns that have moved different distances down the gel. The four columns correspond to reactions that break the DNA: 1) primarily at the adenines, 2) only at the guanines, 3) at the cytosines but not the thymines, and 4) at both the cytosines and thymines. The very shortest fragments are at the bottom right hand side of the picture and the sequence is read up the gel recognizing first the band in the left hand column corresponding to A, a band in the two! right hand columns corresponding to a C, a band in the far right hand column corresponding to T, a band in the left hand column corresponding to A and so forth. After reading up as far as possible, the sequence continues in the sets of bands at the left hand side of the gel and then still further in the pattern in the center of the gel. From the original photograph the sequence of the entire fragment can be read. The fragment is from the genomic DNA corresponding to the variable region of the lambda light chain of mouse immunoglobulin (8).

DNA SEQUENCING AND GENE STRUCTURE Nobel lecture, 8 December, 1980 by WALTER GILBERT Harvard University, The Biological Laboratories, Cambridge, Massachusetts 02138, USA

【問題】

DNAの構造に関する次の文章について、各問いに答えよ。

DNAは、A, C, G, Tの4種類の塩基を含むヌクレオチドが、直鎖状に結合した物質であり、遺伝情報はDNA上の塩基配列として存在している。DNAの塩基配列を調べる方法の1つの手順を示す。

1. 同じ塩基配列のDNAを多量に集める。
 2. そのDNAの2本鎖のうち、特定の1本鎖の片方の端に印をつける。
 3. 2のDNAを4つの試験管に分けて入れ、Gのところ、GかAのところ、TかCのところ、Cのところ、DNAを切断する。ただし、それぞれのDNAはさまざまな異なる位置の1か所だけで切断されるようにする。
 4. それぞれの試験管のDNAを、ある方法により大きさで分ける。
 5. 印がついたDNAを検出すると、DNAの大きさの違いにより、DNAが小さいものから大きいものへと並ぶ。
- 図は、ある遺伝子の一部の塩基配列を、この方法により調べた結果を示したものである。

問 下線部ウについて、各問いに答えよ。

(1) 図では何個の塩基が決定されたか。

(2) 印に一番近い塩基は何か。

(3) 調べた鎖の塩基配列を答えよ。なお、印に近い側を左に書くこと。

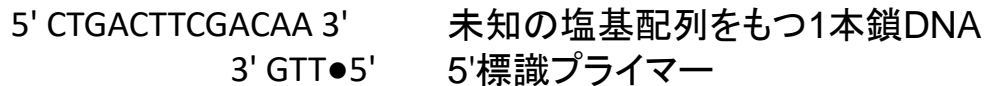


ジデオキシ法 (Sanger*法, 鎖終結法)

* インシュリンの1次構造決定でノーベル賞を受賞した人

大腸菌DNAポリメラーゼIのKlenowフラグメントを用い、配列を決めたい1本鎖DNAの相補的コピーをつくる。この時DNA合成に必要な4種のdNTP(dATP, dGTP, dCTP, dTTP)以外に、DNA阻害剤である**2',3'-ジデオキシヌクレオシド三リン酸(ddNTP)**を少量反応液に加える。ddNTP (terminatorという) が取り込まれると、DNA合成は停止する。反応の停止はランダムな位置で起こるので、鎖の長さの異なる様々な断片が得られ、あとは化学的分解法と同様な考え方で配列を決定する。

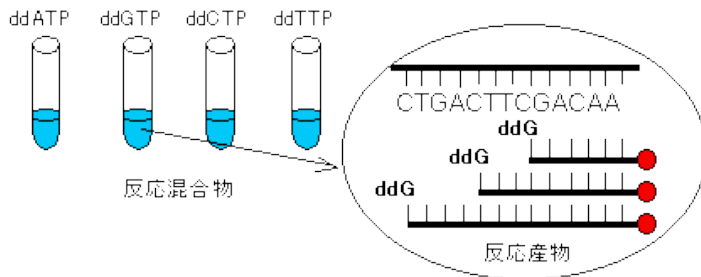
1) まず、配列を決めたい1本鎖DNAの3'末端に相補的な短いオリゴヌクレオチドをプライマーとして準備し、プライマーの5'末端を³³Pや³⁵Sで放射標識するか、または**蛍光色素で標識**する。



2) 配列を決めたい1本鎖DNA溶液を4等分し、標識プライマー、4種のdNTPを加える。これらにterminator(ddNTP)を1種ずつ加える。

3) DNAポリメラーゼIのKlenowフラグメントを加えて複製を開始する。

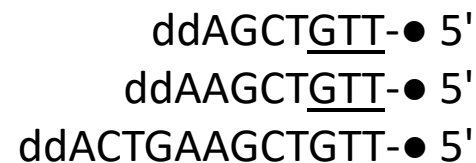
->加えたddNTPのため、鎖伸長はランダムな位置で停止する。



4) これら4つの試料をポリアクリルアミドゲル電気泳動にかける。各々のDNA断片はサイズだけで分離される。

5) ゲルのオートラジオグラムを作成する。

例えばddATPを少量加えた試験管の場合、上の配列では次の3つの標識断片が生成する。下線はプライマー部分を表す。

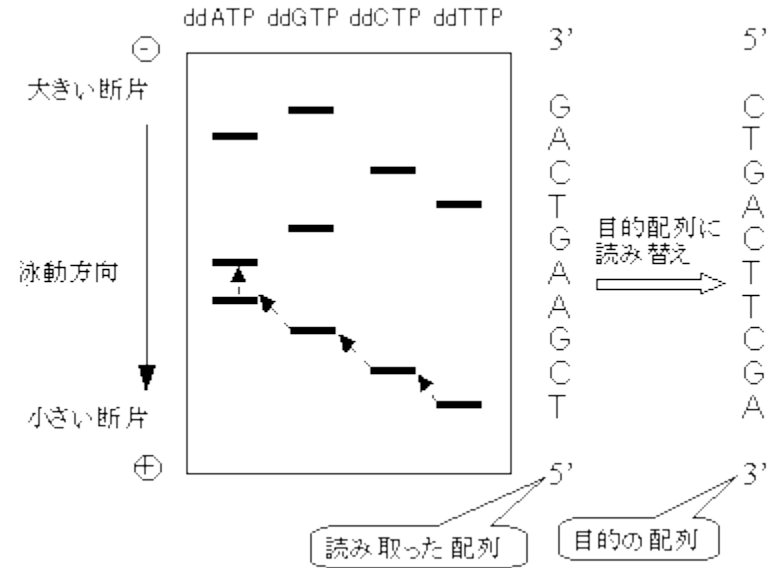


6) 図のように、泳動先端（下端）のバンドから順に上にバンドの位置を辿っていくことで、目的配列の相補鎖の5'→3'方向の配列を決定できる。これを目的配列（鋳型鎖）に読み替えるには、読みとった配列をひっくり返せばよい。

蛍光標識法を用いる塩基配列決定の自動化

Sangerの原法では、実験ごとにプライマーをいちいち放射標識する必要があり、大変面倒である。波長の異なる蛍光発色団でterminatorを標識することにより、迅速で簡便な塩基配列決定法が開発された。

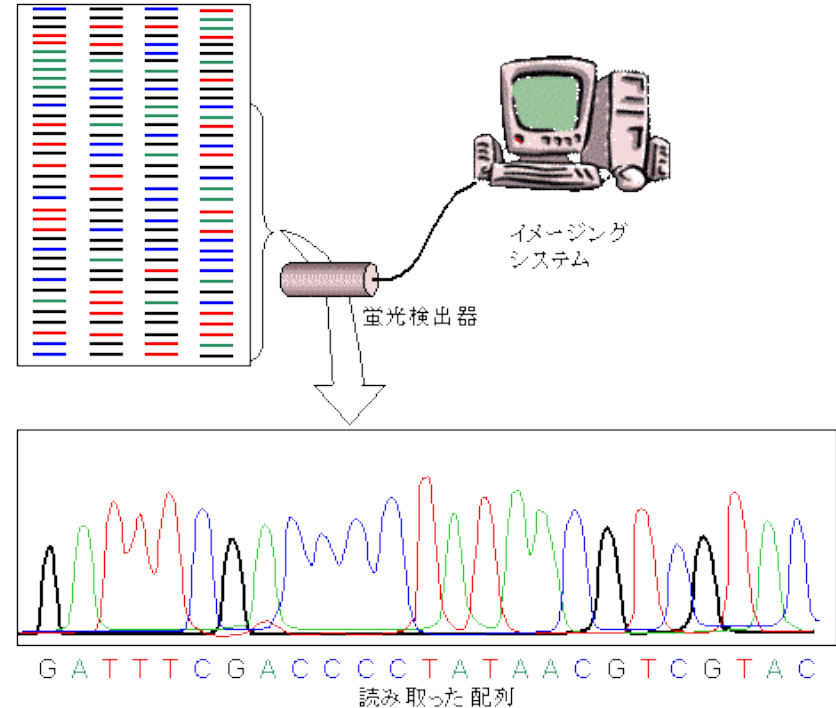
ddATP-● ddTTP-●
ddGTP-● ddCTP-●



Sanger法による塩基配列決定]

1.《 蛍光標識法の手順 》

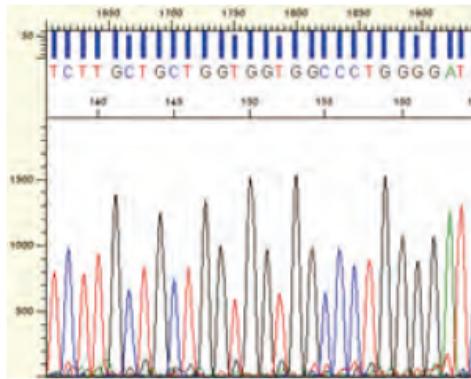
1. 配列を決めたい1本鎖DNA、プライマー、4種のdNTPを含む溶液に、波長の異なる蛍光色素で標識した4種のddNTP混合物(terminator)を加える。
2. DNAポリメラーゼで鎖を伸長させる。加えたterminatorによって、鎖伸長はランダムな位置で停止する。
3. 試料をゲル電気泳動にかける。
4. ゲルにレーザー光を当てて蛍光バンドを検出し、情報をイメージングシステムに転送する(図 8.11)。
5. 検出したバンドの情報をコンピューターで処理し、ピークとして表す。



[蛍光標識法による塩基配列決定]

《 蛍光標識法の利点 》

1. プライマーを標識する必要がない。
2. 反応は1本の試験管で行える。
3. 電気泳動も1試料1レーンですむため、バンドの物理的なゆがみによる読み取り誤差がない。
また、大量の試料を一度に処理できる。
4. 放射能を使用しないので、いつでも実施でき、また、放射性廃棄物を出さない。
5. オートラジオグラフィーが不要。
6. 自動化ができるので、迅速で簡便。



5'- AGA.....ACGA AGACGCACGACGACCACCGGGACCCCTA -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGGACCCCT -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGGACCCC -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGGACCC -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGGACC -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGGAC -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGGA -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGGG -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCGG -3'
 5'- AGA.....ACGA AGACGCACGACGACCACCG -3'
 5'- AGA.....ACGA AGACGCACGACGACCACC -3'
 5'- AGA.....ACGA AGACGCACGACGACCAC -3'
 5'- AGA.....ACGA AGACGCACGACGACCA -3'
 5'- AGA.....ACGA AGACGCACGACGACC -3'
 5'- AGA.....ACGA AGACGCACGACGAC -3'
 5'- AGA.....ACGA AGACGCACGACGA -3'
 5'- AGA.....ACGA AGACGCACGACG -3'
 5'- AGA.....ACGA AGACGCACGAC -3'
 5'- AGA.....ACGA AGACGCAC -3'
 5'- AGA.....ACGA AGACGAA -3'
 5'- AGA.....ACGA AGACGA -3'
 5'- AGA.....ACGA AGACG -3'
 5'- AGA.....ACGA AGAC -3'
 5'- AGA.....ACGA AGA -3'
 5'- AGA.....ACGA AG -3'
 5'- AGA.....ACGA A -3'

3' 末端の塩基だけが蛍光標識されている。

プライマー配列

電気泳動方向

早く反応が止まった塩基数の少ないヌクレオチド鎖ほど速く移動し、順に蛍光色素の種類を読み取っていく。相補的な塩基が配列になる。

求める塩基配列

5' TCTTGCTGCTGGTGGTGGCCCTGGGGAT 3'

サンガー法の原理

次世代シーケンサー

キャピラリーサンガー法では同時処理できるDNA断片は800～1,000程度なのに対し、次世代シーケンサー(NGS)では数千万のDNA断片を平行処理できます。解析に必要なコストの低減も、技術の進歩とともに進められています。

① DNA合成・光検出によるシーケンシング

目的のDNAを断片化し、DNA断片をガラス表面などに固定して、感度を上げるためにPCR法でDNAを増やし、それぞれクラスター化しておきます。その上でDNA伸長反応を行わせます。DNAが1塩基伸長する毎に塩基に標識した蛍光色素が発光するように工夫して、発光の様子をカメラで撮影して検出します。DNA断片から得られる情報は短いものですが、これらのDNA配列情報を計算機でまとめて解析することにより、ひとつながりのDNA情報を再構築していきます。塩基をつないでいくという化学反応と、蛍光を解析するという光学を組み合わせたものです。DNA増幅などを含めたサンプル調製や、光検出を行うための試薬や検出器を必要とするため、コストが高くなりがちです。

② 1分子のDNAからの検出

DNA1分子を鋳型としてDNA合成を行い、蛍光などで1塩基ごとの合成反応をリアルタイムに検出できる技術も開発されています。DNA合成速度と同じ速度で配列を決定することが出来ます。

③ 光検出以外の検出

光検出を行うための試薬や検出器を必要としない方法として、ナノポアシーケンサーがあります。電圧をかけたナノサイズの穴(ナノポア)をDNA分子が1塩基ずつ通過する際の電流の変化で塩基配列をきめるものです。DNAが通るナノポアの作製には半導体チップ技術が基本的に用いられています。この方法ではDNA分子を正確にナノポアにいれ、速く通過させるかという課題に様々な技術が投入され、低コスト化、スピードアップが可能になっています。

これらの次世代シーケンサーの開発により、すでにヒトのゲノム解析が1,000ドル程度で可能になっています。また、次世代シーケンサーは医療機器としてFDAで認証され、遺伝性疾患の診断に既に使われています。

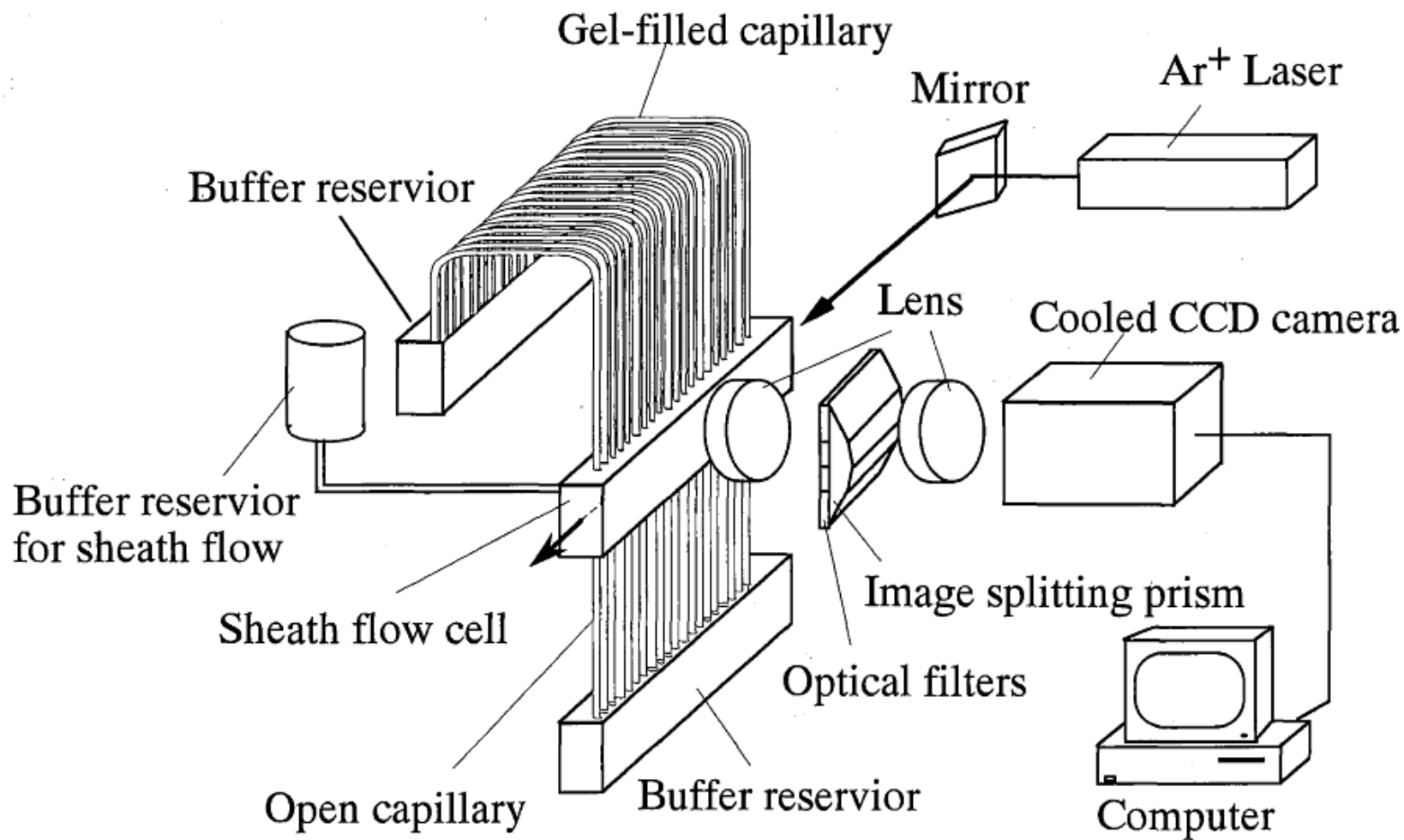


Fig. 1. Schematic diagram of multi-capillary DNA sequencer.

Ar⁺ laser is used for excitation. The fluorescence line image is split into four different color images with image-splitting prism coupled with four optical filters.

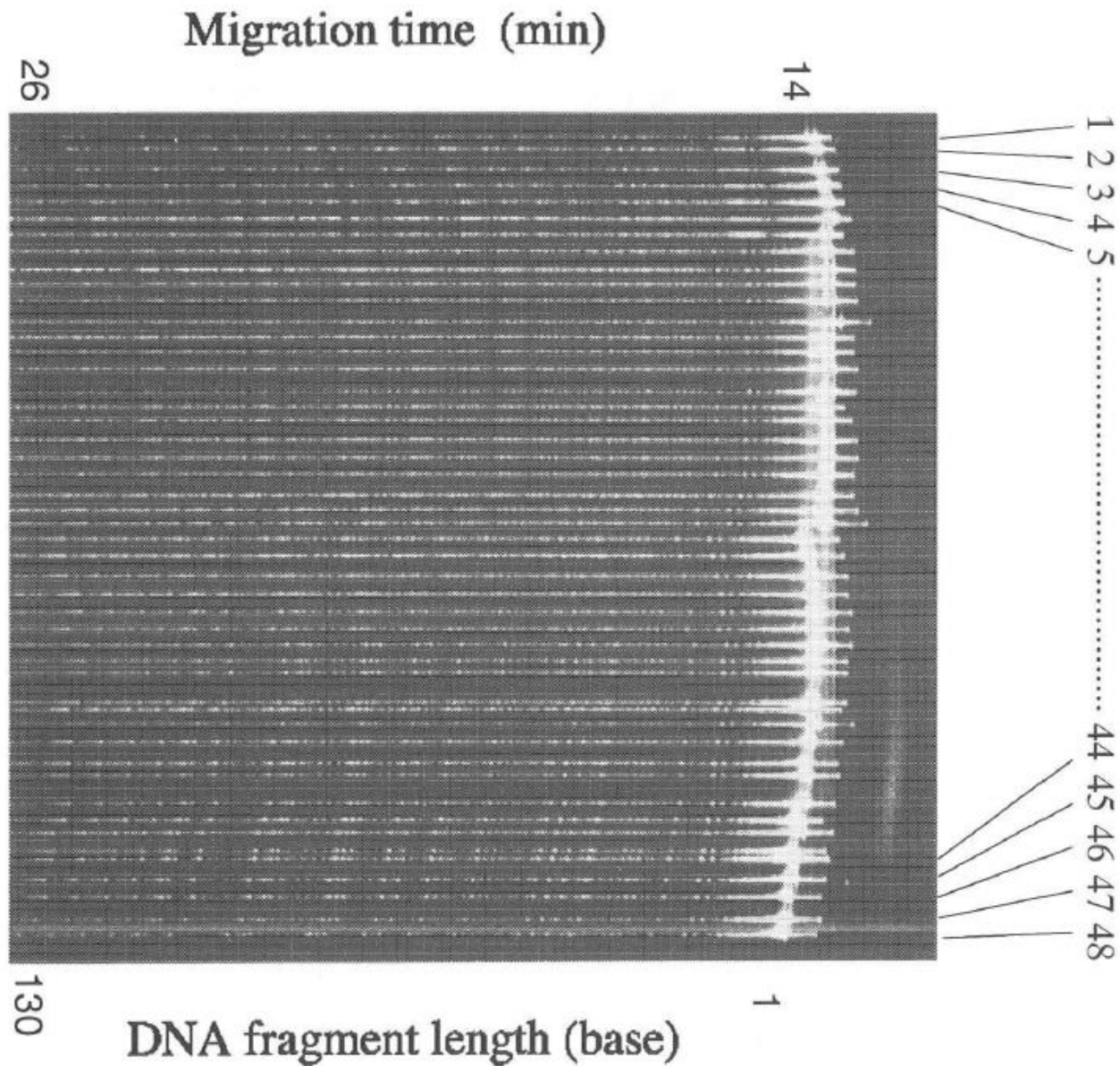


Fig. 3. Image of 48 gel-filled capillary array.

The image represents one channel (green) of DNA sequencing using a 48-capillary array.

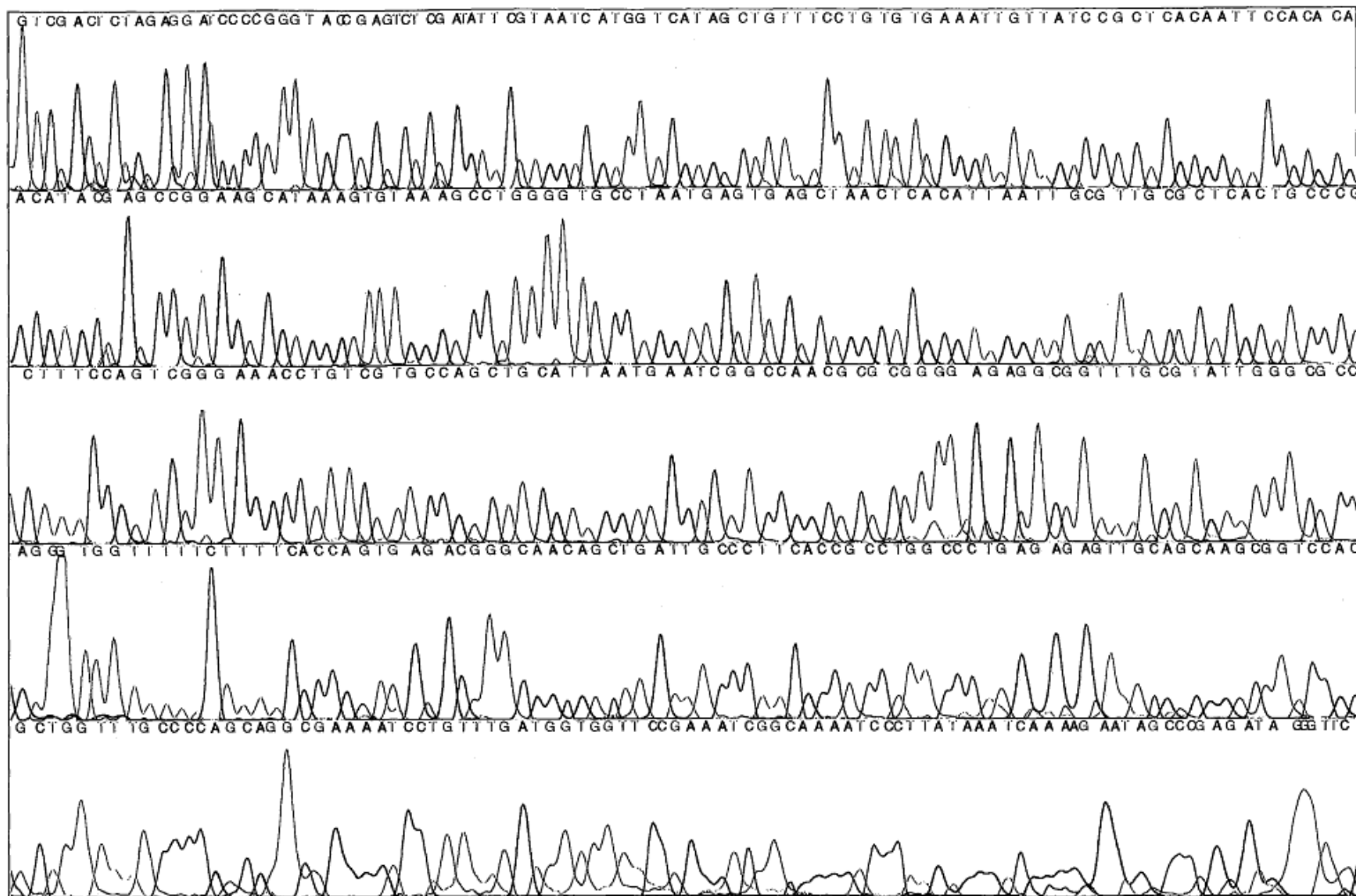


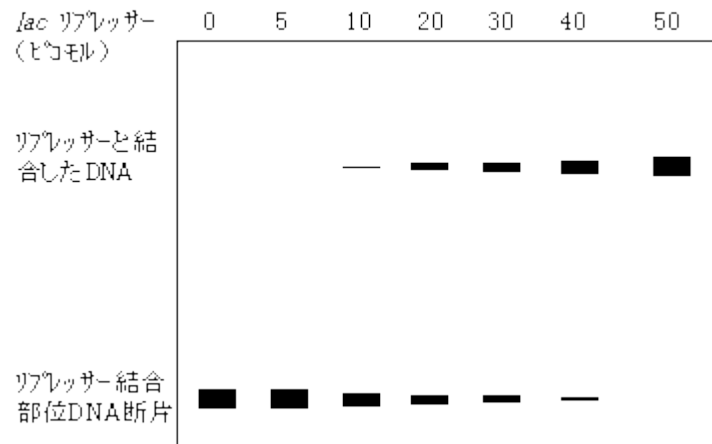
Fig. 4. Analyzed four-color DNA sequencing data of M13mp18 DNA from one capillary in a 48-capillary array.

The sample was injected by electrophoresing it at 25 V/cm for 20 s and separated at 100 V/cm in 4% T, 5%C polyacrylamide gel-filled capillaries (75 μ m I.D. \times 25 cm).

転写制御因子などのDNA結合タンパク質の検出

1) ゲルシフト法(電気泳動移動度シフト解析: electrophoretic mobility shift assay, EMSA)

タンパク質(転写制御因子)が結合すると、分子サイズが大きくなり電気泳動におけるDNAの移動度が低下することを利用した転写制御因子の検出法。32Pで標識したDNA断片と転写制御因子を混ぜ、ゲル電気泳動にかける。オートラジオグラフィーでDNAの位置を見ると、因子の結合したDNAはゆっくり動くので、通常のコムラよりも遅れて移動するバンドとして検出される。DNAのほんの一部に結合があっても検出でき、DNA結合反応の定量的解析ができる。

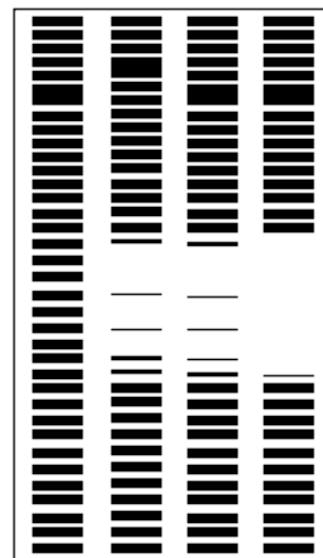


DNase I フットプリント法

DNA分解酵素の一種であるDNase I (デオキシリボヌクレアーゼ I) は裸のDNAを分解するが、タンパク質と結合したDNAは分解しない。この性質を利用して、転写制御因子が結合するDNAの領域を決定する方法である。一端を32Pで標識したDNA断片と転写制御因子を混ぜ、少量のDNase Iを低温または短時間作用させて部分的にDNAを分解する。電気泳動とオートラジオグラフィーでDNAの断片を調べる。標識末端からのDNA長に従ってはしご状のバンドが見えるが、転写制御因子が結合した部分ではDNAが切れないために断片が生じず、はしご状のバンドの一部が足跡(footprint)のように抜ける。同じDNA断片をMaxam-Gilbert法で処理したものと比較することで、抜け落ちた領域が決定できる。

結合因子の量

- + ++ +++



結合因子で保護されている領域

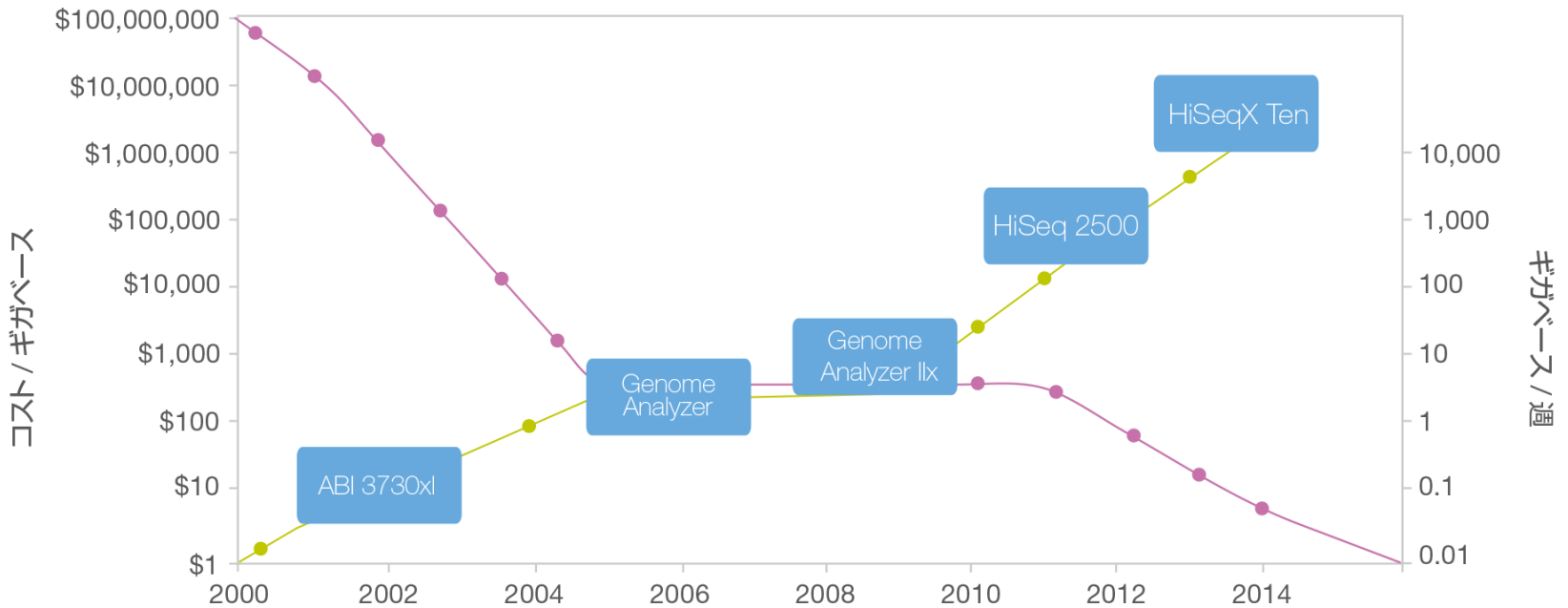


図1: 2000年以降のシーケンスコストおよびデータアウトプット 2000年以降のシーケンスにおける、データアウトプットの急激な増加に伴うコストの低下。グラフの両側のY軸は対数軸です。

1年間にシーケンスできるヒトゲノム数

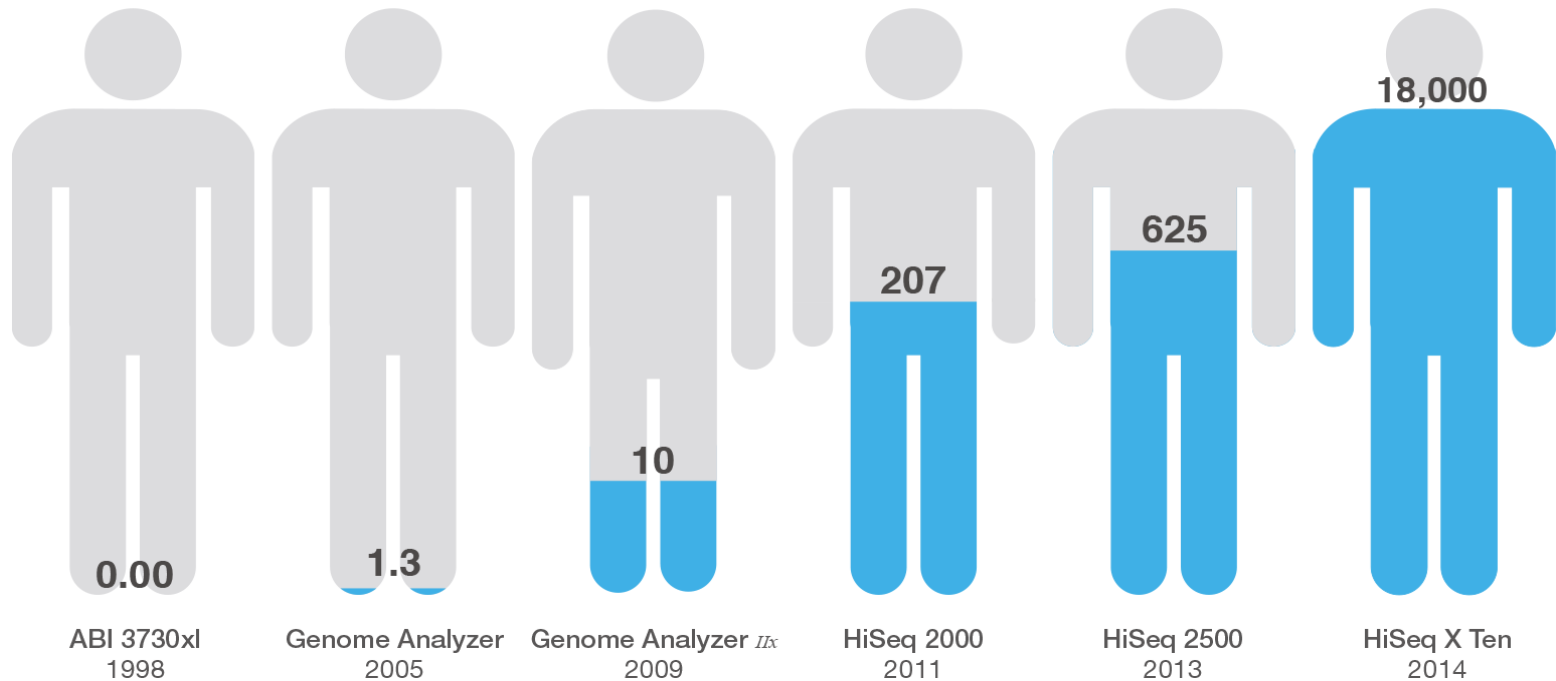


図2: 数十年間に及ぶヒトゲノムのシーケンス 1990年代以降、32億個のヒトゲノム塩基すべてをシーケンスするキャパシティ(30倍のカバレッジ)は、指数関数的に増加し、2005年には、イルミナGenome Analyzer Systemの導入により、年間1.3サンプル分のヒトゲノムをシーケンスすることができるようになりました。ほぼ10年後、イルミナHiSeq X Tenシステムを用いてシーケンスを行えるヒトゲノムの数は年間18,000サンプル分のヒトゲノムにまで増加しました。

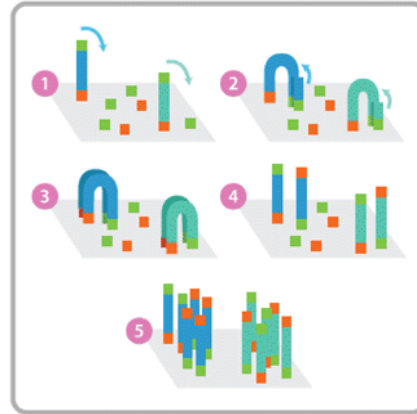
次世代シーケンシング

① ライブラリー調製

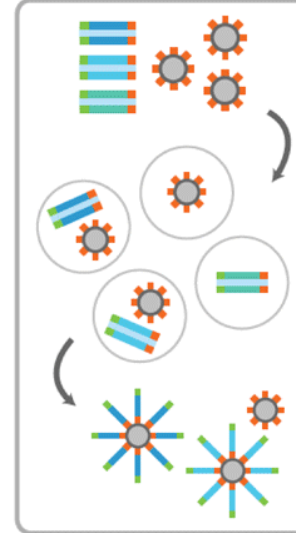


② クローンの増幅

ブリッジPCR

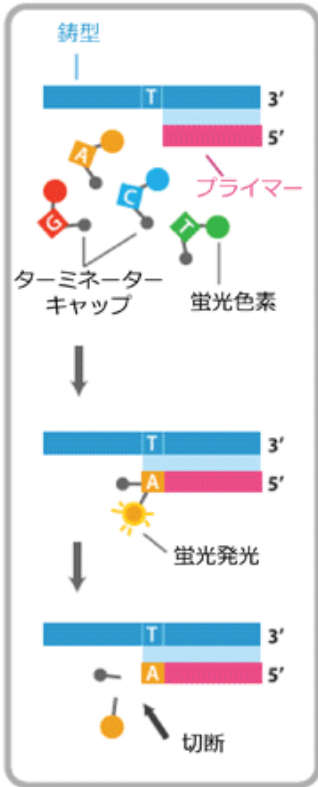


エマルジョンPCR

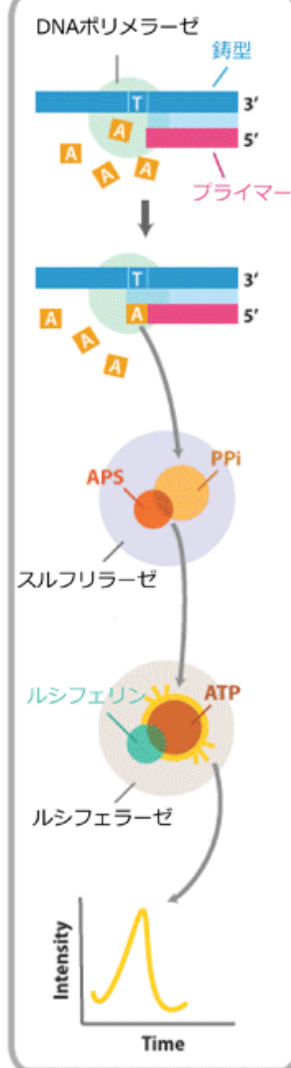


③ 環状アレイシーケンシング

合成によるシーケンシング

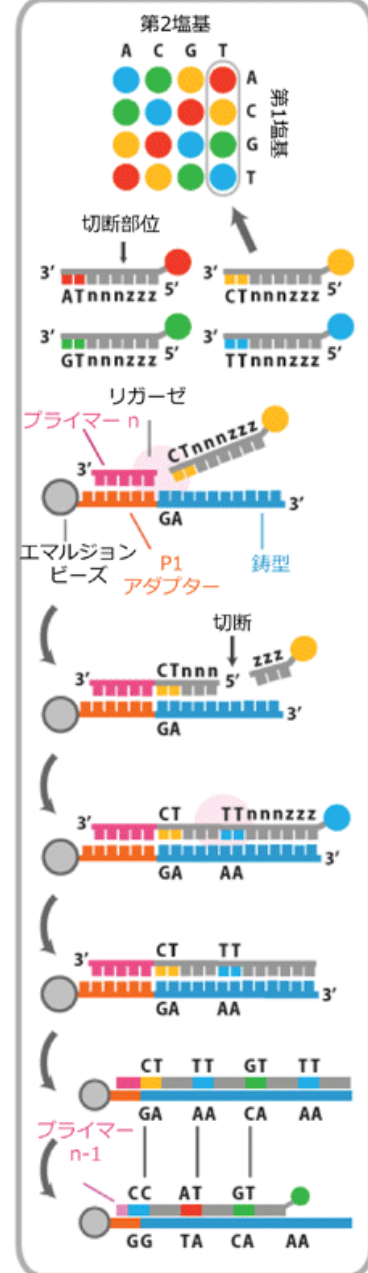


パイロシーケンシング



③ 環状アレイシーケンシング

ライゲーションによる
シーケンシング



パイロシーケンス法の原理

パイロシーケンス法は、Luciferase の発光反応を利用した合成による配列解読 (sequencing by synthesis) 原理に基づいています。

ステップ 1

一本鎖 PCR 産物にシーケンシングプライマーをハイブリダイズした後 (図 1)、必要な酵素、基質類を自動分注します*。

ステップ 2

dNTP をターゲット配列に対応する順序で 1 種類ずつ添加すると、DNA Polymerase による塩基伸長反応に伴って取り込まれたヌクレオチド量に比例した PPi (ピロリン酸) が遊離します (図 2)。

ステップ 3

遊離した PPi と反応して Sulfurylase が ATP を産生します。この ATP が Luciferase 発光をカスケードします。光は CCD に検出され、ピーク波形 (パイログラム: Pyrogram) として観察されます。各ピーク (光シグナル) の高さは、取り込まれたヌクレオチド数 (遊離した PPi 量) に比例します (図 2)。

ステップ 4

dNTP 分解酵素の一つである Apyrase が、反応に寄与しなかった dNTP および ATP を分解します (図 2)。

ステップ 5

この過程の繰り返しにより相補的な DNA 鎖が形成され、得られるパイログラムの発光ピークから塩基配列が決定できます (図 3)。

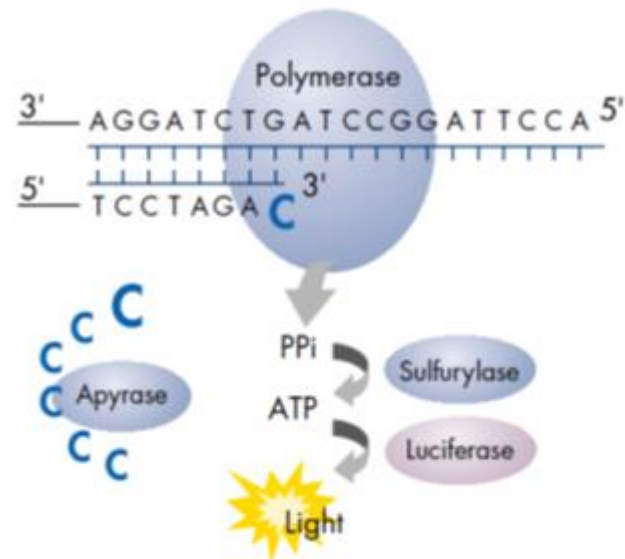
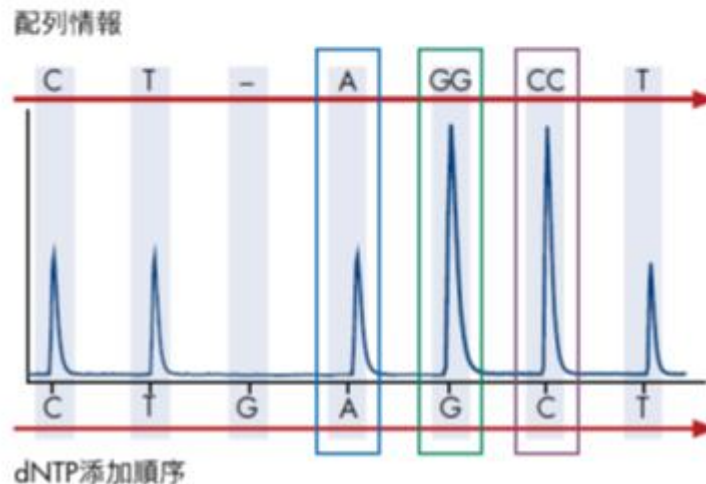


図 2. パイロシーケンス反応の原理*

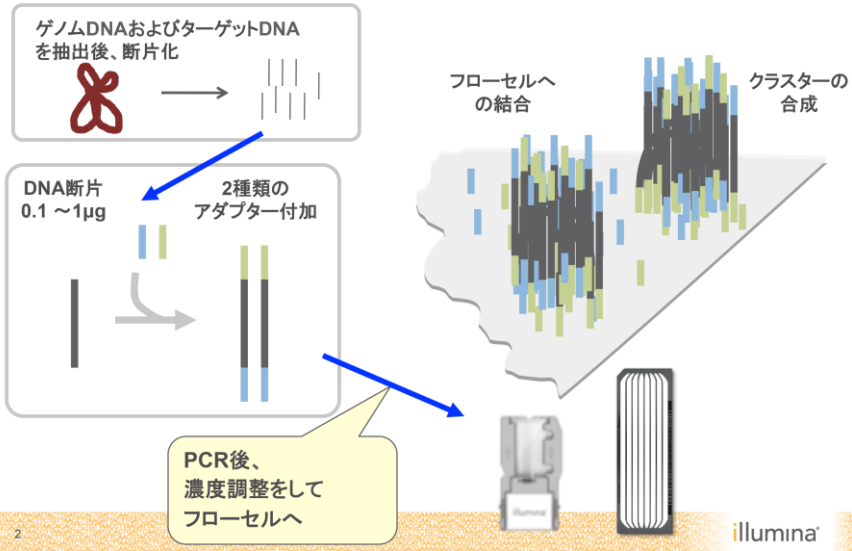


* 自動分注される酵素と基質

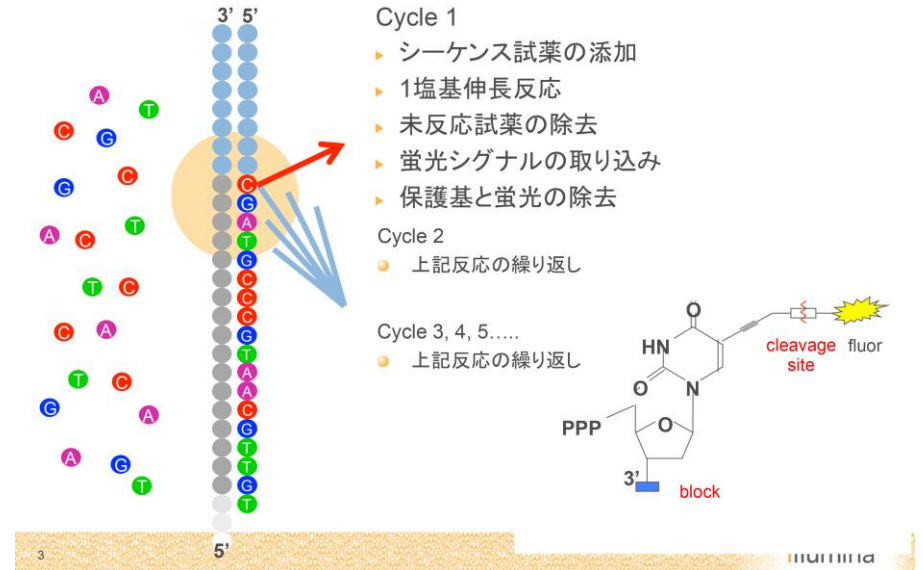
酵素: Polymerase、Sulfurylase、Apyrase、Luciferase

基質: APS (adenosine 5' phosphosulfate)、Luciferin

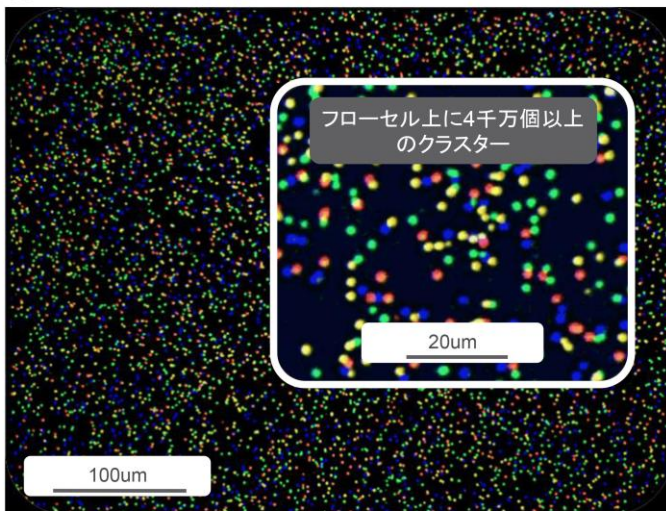
サンプル調製とDNA分子の増幅(自動化)



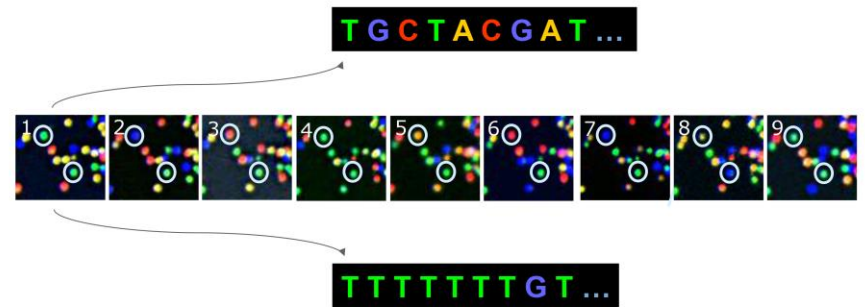
可逆的ターミネーター法を用いた1塩基伸長反応



画像イメージの取り込み



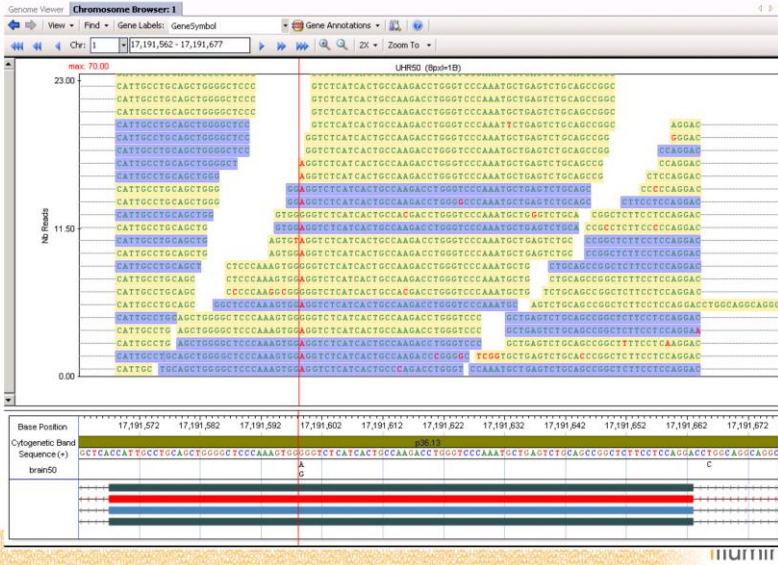
画像蛍光シグナルから塩基への変換



1塩基伸長反応ごとに蛍光イメージをとる
→ 蛍光の色から各塩基を決定する

リファレンス(既知のゲノム配列)へのマッピングと

シングルリード法・ペアエンド法



シングルエンド

- ▶ DNA断片の片側を100bp読み取り



ペアエンド

- ▶ DNA断片の両端を100bp読み取り
- ▶ DNA断片の長さにより、さらに2つに区分
 - 200-500bp: ペアエンド法(イルミナシステムのみ)
 - 2-5kb: メイトペア法



ペアエンド法
200-400bp



メイトペア法
2-5kb

• 両端を読むことで、データ量が2倍に
 • アライメント時の効率アップ

ペアエンド(両端読み)とは？

- ▶ 同一のDNA断片の両端を読む
- ▶ 片側を読んだ後、DNA断片をひっくり返してもう一方の端を読む

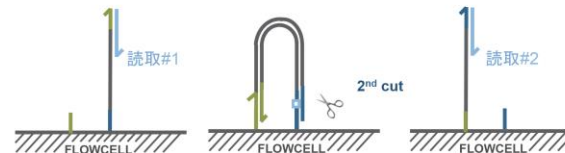


表.1 入手可能なパイロシーケンシングを利用したNGS機器の技術概要

NGSシステム	GS Junior	GS Junior Plus	GS FLX+ システム	
			GS FLX Titanium XL+	GS FLX Titanium XLR70
リード長	400bp	700bp	700bp (up to 1,000bp)	450bp (up to 600bp)
処理量	35Mb	70Mb	700Mb	450Mb
実行あたりのリード数	100,000 ショットガン, 70,000 増副産物	100,000 ショットガン, 70,000 増幅産物	1,000,000 ショットガン	1,000,000 ショットガン, 700,000 増幅産物
精度	99% at 400bp	99% at 700bp	99.997%	99.995%
実行時間	10 時間	18時間	23時間	10時間

表.2 入手可能な合成によるシーケンシングを利用する次世代シーケンシング機器の技術詳細

NGSシステム	MiSeq	NextSeq 500		HiSeq 2500		HiSeq 3000	HiSeq 4000
実行モード	N/A	中出力	高出力	迅速な実行	高出力	N/A	N/A
実行ごとのフローセル数	1	1	1	1 or 2	1 or 2	1	1 or 2
出力範囲	0.3-15 Gb	20-39 Gb	30-120 Gb	10-300 Gb	50-1000 Gb	125-750 Gb	125-1500 Gb
実行時間	5-55 hrs	15-26 hrs	12-30 hrs	7-60 hrs	<1-6 days	<1-3.5 days	<1-3.5 days
フローセルごとのリード数	25,000,000	130,000,000	400,000,000	300,000,000	2,000,000,000	2,500,000,000	2,500,000,000
最大リード長	2 x 300bp	2 x 150bp	2 x 150bp	2 x 250bp	2 x 125bp	2 x 150bp	2 x 150bp

- [1953年](#) DNAの二重らせん構造の発見
- [1972年](#) リコンビナントDNAの技術が開発された。これ以前はバクテリオファージかウイルスのDNAしかDNAシーケンシングのサンプルには使えなかった。
- [1975年](#) 最初の完全なゲノム配列としてバクテリオファージφX174の全ゲノムが解読された。
- [1977年](#) マキサムとギルバートによる"DNA sequencing by chemical degradation"(分解によるシーケンシング:化学分解法)が発表された。サンガーはこれとは別に"DNA sequencing by enzymatic synthesis"(合成によるDNAシーケンシング:酵素法)を発表した。
- [1980年](#) サンガーとギルバートがノーベル化学賞受賞
- [1986年](#) カリフォルニア工科大のLeroy E. Hoodの研究室とスミスが半自動のDNAシーケンサーを発表
- [1987年](#) アプライドバイオシステムズ社が世界初の自動DNAシーケンサーABI 370を発売。
- [1995年](#) Richard Mathies らが蛍光色素によるシーケンシング技術を発表。
- [1998年](#) ワシントン大のPhil Green と Brent Ewing が配列解析ソフト[phred](#)を発表。