

# 遺伝子の発現と遺伝子工学

---

- 1 . もう少し詳しい遺伝子発現のはなし
  - 1 ) 転写とプロモータ
  - 2 ) 原核生物の場合
  - 3 ) 真核生物の場合
- 2 . 遺伝子工学
  - 1 ) タンパク質より DNA のほうが扱いが簡単
  - 2 ) 鋏 = 制限酵素と糊 = DNAリガーゼ
  - 3 ) プラスミドベクターとライブラリー
  - 4 ) プローブ ( probe ) の利用
  - 5 ) cDNAライブラリー
  - 6 ) DNA塩基配列の決定
  - 7 ) 遺伝子導入動物 ( transgenic organism )

関連するサイトとリンク ( このページへ戻るときはブラウザーの戻るを選んでください )

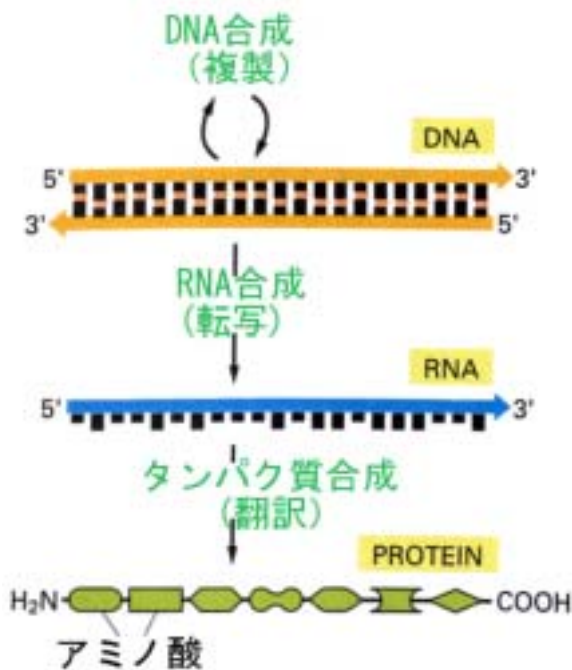
[遺伝学電子博物館\(国立遺伝学研究所作成\)](#)

---

## 1 . もう少し詳しい遺伝子発現のはなし

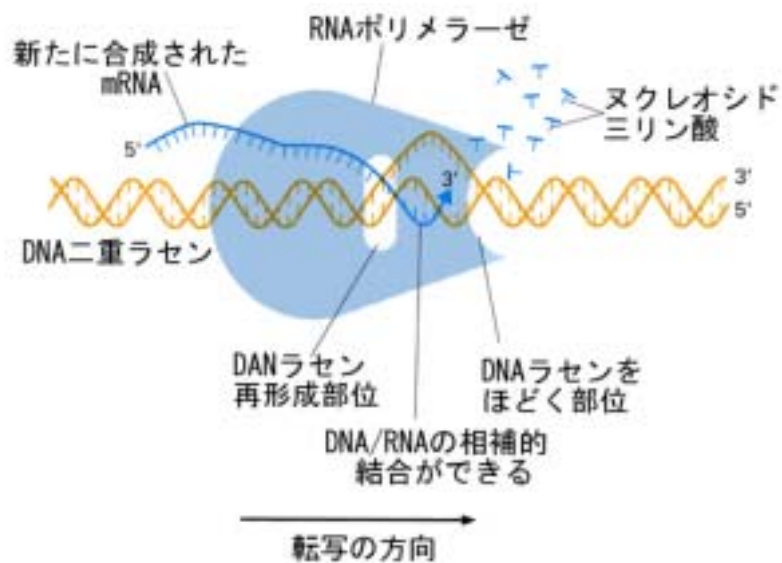
### 1) 転写とプロモータ

「核 - 遺伝情報の倉庫」のページで染色体を構成するDNAの塩基の配列が遺伝の暗号で、コドン ( 3つの塩基 ) が1つのアミノ酸を指定し、したがって塩基の配列によってアミノ酸の一次構造 ( 配列 ) が決定されることを学んだ。DNAからタンパク質への遺伝情報の流れは次の図で表すことができる。



転写によって、DNA の鋳型鎖に相補的なmRNA が作られるが、この反応を触媒する酵素は RNA ポリメラーゼである。

RNA ポリメラーゼは次の模式図が示すように、一方向へ進みながら DNA の二本鎖をほどこき、鋳型鎖となる 3'→5' の塩基の配列を読み取りながら、これと相補的なリボヌクレオシド三リン酸を 1 個ずつ付加して、RNA 鎖を 5'→3' 方向に伸ばしていく。RNA ポリメラーゼの後方では DNA は再び二本鎖のもどりに、合成された mRNA は解離していく。



RNA ポリメラーゼが転写を開始するためには、遺伝子の先頭を認識してそこに強く結合する必要がある。この転写開始のシグナルとなっているのが、プロモーターと呼ぶ DNA の領域である。RNA ポリメラーゼは DNA 上をすべりながら移動しているが、プロモーター領域に出会うとそこで強く結合し、上に述べたように転写開始点から転写を始めることになる。

DNA 上にはさらに、転写の終了を指示する部位も存在する。RNA ポリメラーゼは転写終結部位までくると、そこで DNA を離れ転写が終了する。

遺伝子は、このように転写開始部位と転写終了部位には含まれているのである。

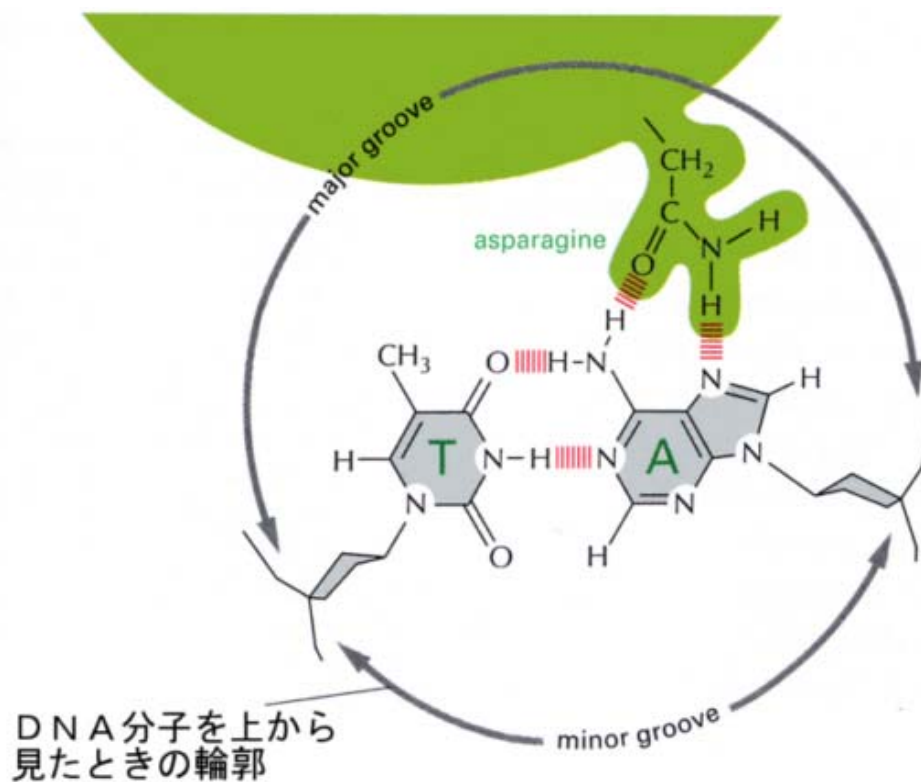
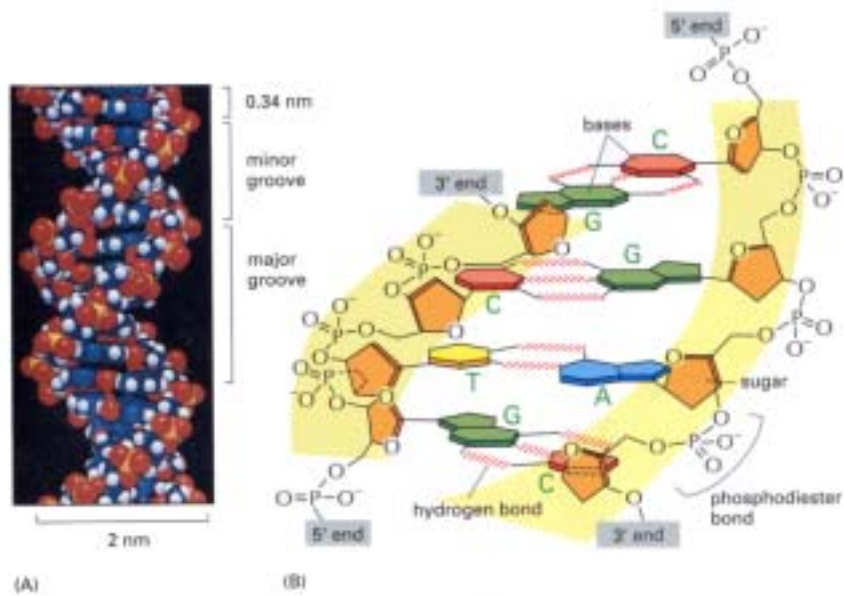
ところで、同じ個体を構成する細胞は同じ遺伝子を少なくとも 1 セットを持っているのだから、RNA ポリメラーゼが区別無く遺伝子の転写を行えば、どの細胞もみな同じ形で同じ働きをすることになる。しかし実際はそんなことはない。今までも見てきたように、各器官を構成している組織は、どれも同じ受精卵に由来するのに、形も働きも千差万別である。それはなぜだろうか。

その理由は、それぞれの細胞で**遺伝子の発現 (gene expression)** が調節されていて、特定の遺伝子サブセットしか発現しないからである。

それでは、遺伝子の発現はどのように調節されているのだろうか。

DNA の構造を思い出してほしい。転写したり、転写を調節するためには、何かが DNA に働きかける必要がある。DNA に働きかける分子はやはりタンパク質である。

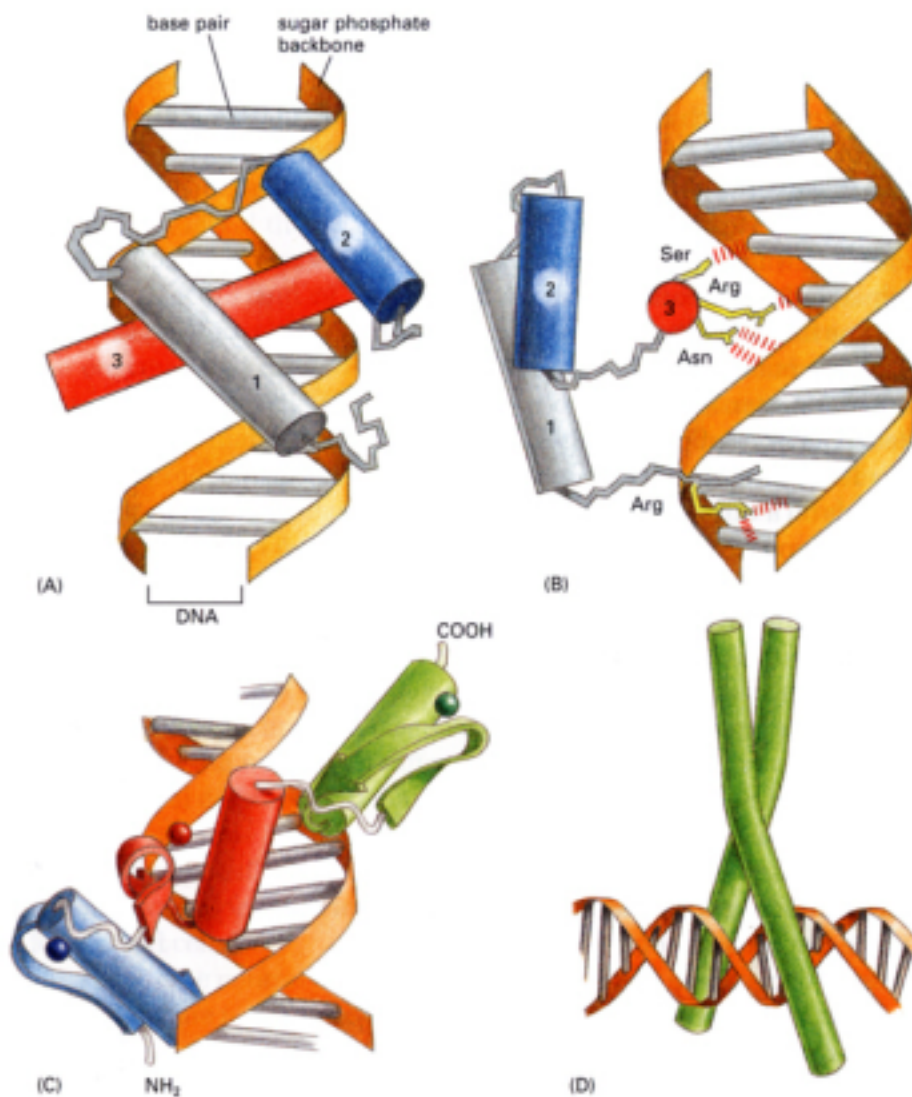
DNA の二重らせんには、大きな溝 (major groove) と小さな溝 (minor groove) が交互にある。この溝と、相補的な水素結合を作っていない塩基部分を認識して、特定のタンパク質が結合することができる。下の図は、断面図なので認識部位を 1 点だけ示してあるが、実際はこのような点が 10 から 20 点あり、結合を強くしている。



RNAポリメラーゼや、これから話をするレプレッサータンパク質もこのような結合を作ってDNAに働きかける。

一方、発現を調節するタンパク質も、このような結合を作ってDNAに働きかける。下の図は、発現調節タンパク質に共通に見られる構造（モチーフ）を、模式的に示したもので

ある。(A)と(B)はホメオドメイン(2年で学習する)、(C)はジンクフィンガー(zinc finger)、(D)はロイシンジッパーである。



さてこれだけの予備知識で、転写がどのように調節されているか、原核生物と真核生物で見ていくことにしよう。

## 2) 原核生物の場合

### 1 - 1 遺伝子の構造

大腸菌 (*Escherichia coli*) の染色体は、真核生物の染色体のような複雑な構造はしていない。1本のDNA分子が環状につながった単純なものである。この染色体はその長さから考えて2000から4000のポリペプチド鎖をコードしていると考えられている。

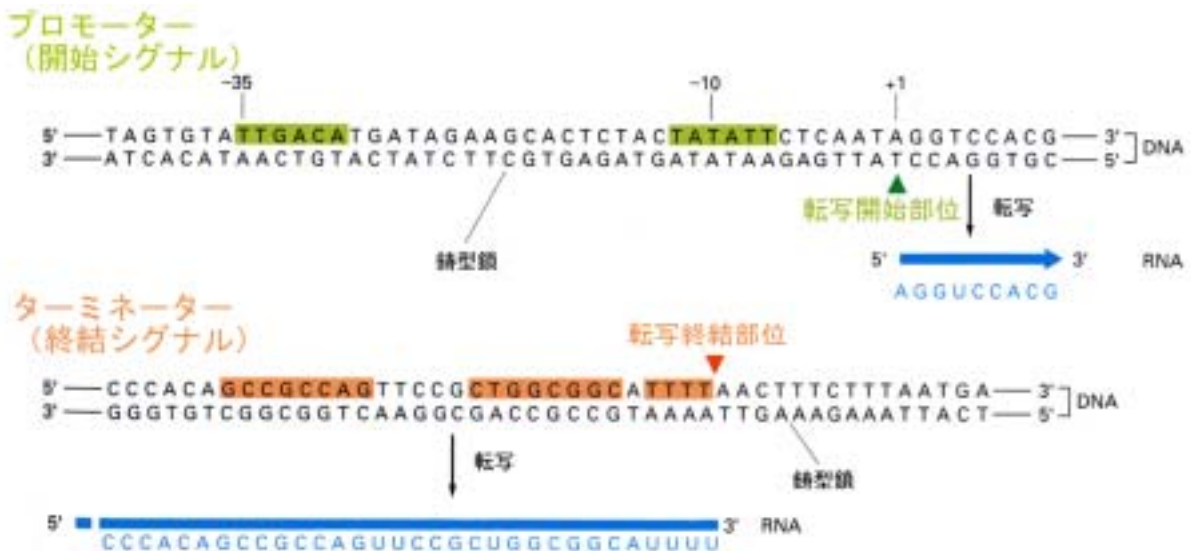
解糖の各段階の酵素や、ATPの生産に働く酵素は、生きて行くためには常に必要なことから、転写はいつも起こっている。このような遺伝子を、**構成的遺伝子( constitutive gene)**と呼んでいる。

DNAの遺伝情報をmRNAに転写するためには、すでに述べたようにRNAポリメラーゼがプロモーター部位に結合し、DNA分子をほどこきながら、転写開始部位から終了部位まで、DNAの鋳型鎖の塩基に相補的なヌクレオシド三リン酸を次々と呼び込んで、結合を伸ばして行く必要がある。

RNAポリメラーゼが結合するプロモーター部位には、遺伝子の種類にかかわらず、相同な2つの構造が保存されている。1つは転写開始部位からおよそ6塩基対( base pair, bpと略す)上流の-10 region と名づけられた6塩基対で、もう一つはさらに17bp上流の-35 region と呼ばれる6塩基対の部分である。この2つの塩基配列によってRNAポリメラーゼの結合の仕方が決まり、したがってその進む方向、すなわち転写の方向が決まる。

RNAポリメラーゼは、下に示した転写終結部位までくるとヌクレオシド三リン酸の付加を止め、DNAの鋳型鎖とmRNAを離す。

コード領域は、転写開始部位と転写終結部位にはさまれていることになる。



## 1 - 2 オペロン説

さて、大腸菌をグルコースのかわりにラクトース(グルコースとガラクトースからなる二糖類)を含む培地で培養すると、大腸菌はやがてラクトースを解糖の出発物質として使

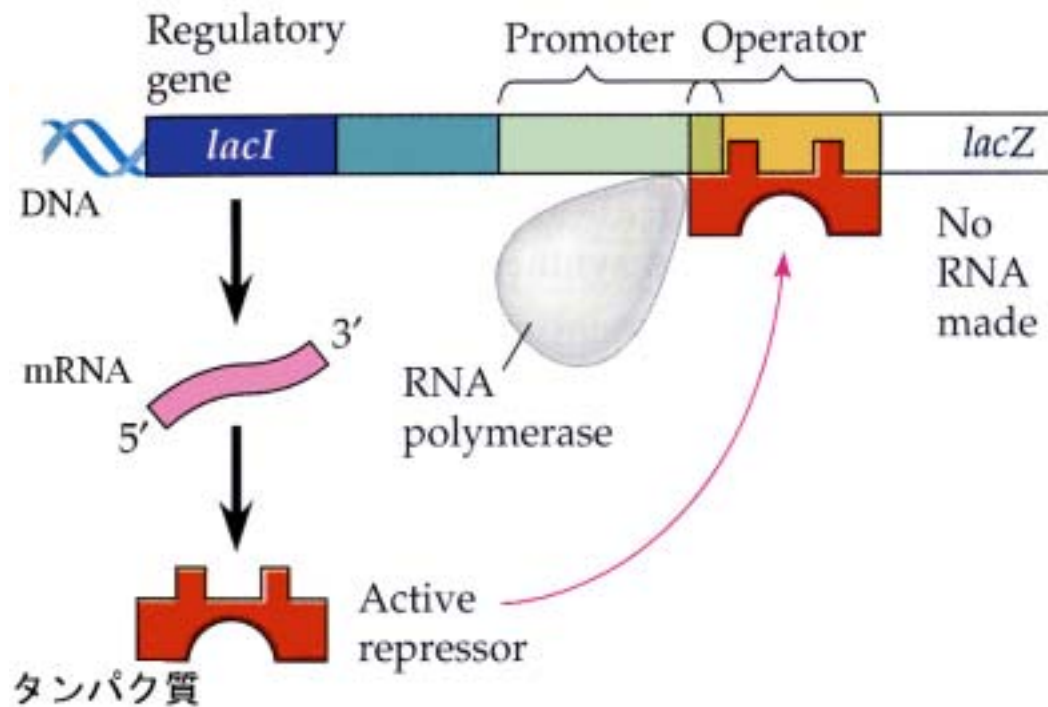
えるようになる。このような状況下では、ふだん、ほんのわずかしかない、ラクトースをグルコースとガラクトースに分解する  $\beta$  ガラクトシダーゼ、ラクトースを細胞内に取り込むパーミアーゼ、それとトランスアセチラーゼの3つの酵素タンパク質が、大腸菌細胞内に出現してくる（誘導、induction）。

突然変異体の研究から、1箇所の変異でこの3つのタンパク質がすべて誘導されなくなることがわかり、3つの遺伝子が一まとまりになっていて、まとめて発現の調節がおこなわれていると考えられた。このような複数の遺伝子のまとまりは**オペロン (operon)** と名づけられた。

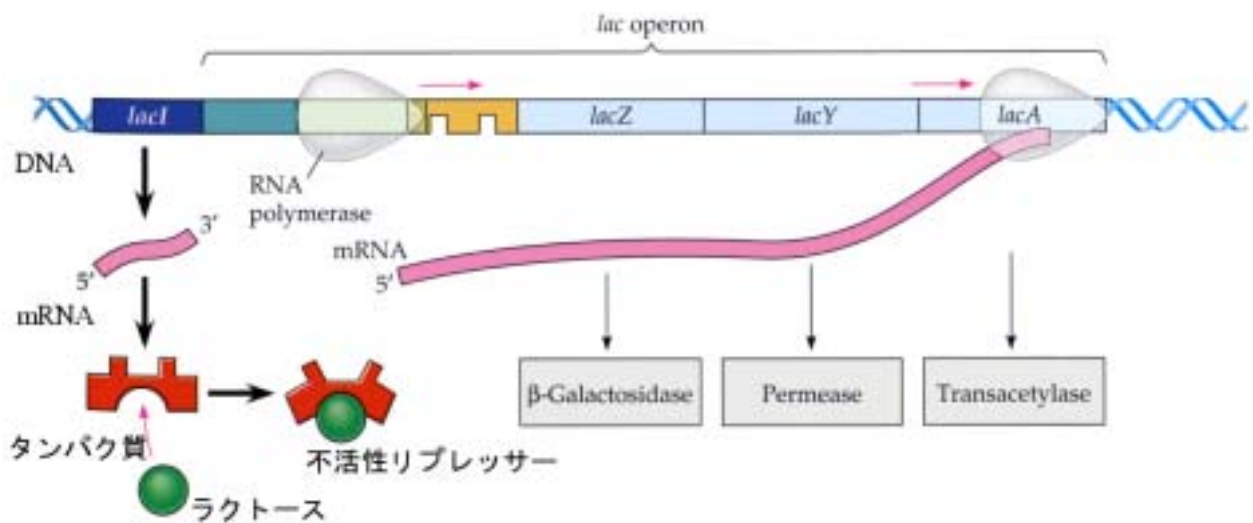
その後の研究によって、ラクトースオペロンは(*lac* オペロン)、上に述べた3つの酵素をコードする**構造遺伝子 (structural gene=protein-coding)** と、そのすぐ上流にあるオペレーター部位を含むプロモーター部位で構成されていることがわかった。RNAポリメラーゼは、プロモーター部位に結合し、その下流にある3つの構造遺伝子をまとめて転写するために、3つの酵素は1本のmRNAから同時に合成される。下の図で *lacZ* は  $\beta$  ガラクシダーゼ、*lacY* はパーミアーゼ、*lacA* はトランスアセチラーゼをコードする遺伝子領域を示す。*lacI* は次に述べるリプレッサータンパク質をコードするリプレッサー遺伝子領域である。



ラクトースのない環境では、ラクトースオペロンのさらに上流にある**調節遺伝子 (リプレッサー遺伝子)** の遺伝情報から作られる**リプレッサータンパク質** が常に合成されていて、プロモーター部位内にあるオペレーター部位と強く結合するために、このリプレッサータンパク質が邪魔になって、プロモーター部位にRNAポリメラーゼが結合できない。そのために上記の3つのタンパク質は合成されない。



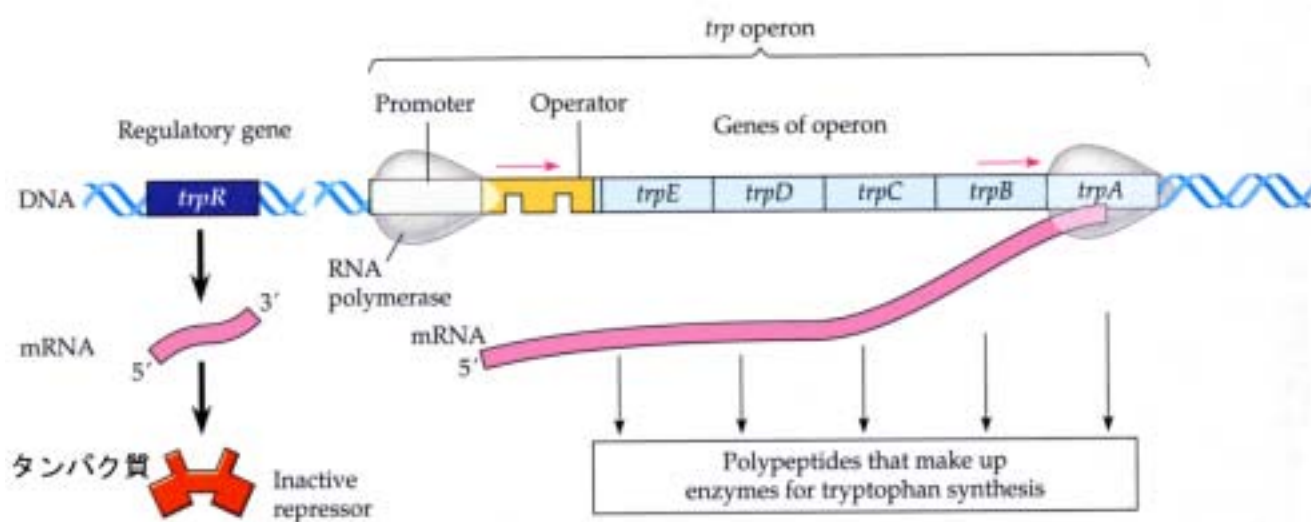
ラクトースが培地に存在すると、少数のラクトースが大腸菌内に入って代謝され、これがリプレッサータンパク質のアロステリック結合部位に結合する。すると、リプレッサータンパク質の構造が変わるため、オペレーター部位からリプレッサータンパク質が外れる。そのため、プロモーター部位にRNAポリメラーゼが結合できるようになり、mRNAの転写が始まる。こうして、ラクトースを解糖に利用できるようになるのである。





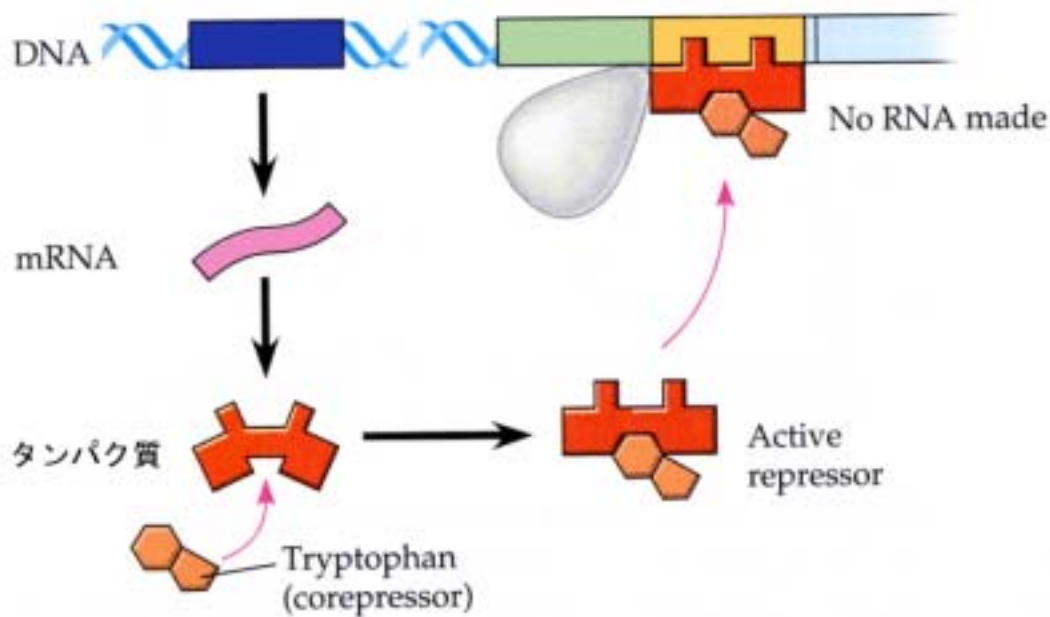
この方式は、いつもはスイッチがオフになっており、必要に応じてスイッチをオンにする方式である。

一方、アミノ酸の一つであるトリプトファンは、いくつかのステップで生合成されるが、それぞれのステップを触媒する酵素はラクトースオペロンのように、ひとまとめのオペロンとなっている。



調節遺伝子の情報から作られるリプレッサーは、ラクトースレプレッサーの場合と違い、このままではオペレーター部位に結合できない。したがって、トリプトファンオペロンはいつも稼働していて、合成酵素群が作られ、トリプトファンが生合成される。

トリプトファンの供給量が需要量を上回ると、トリプトファンがだぶつくので、不必要な生合成を止めた方が無駄がない。実際にトリプトファンはリプレッサーと結合し、その形を変える。するとこのリプレッサータンパク質 - トリプトファン複合体は、オペレータ部位と結合できるようになり、その結果RNAポリメラーゼがプロモーター部位に結合できなくなるために転写が止まる。



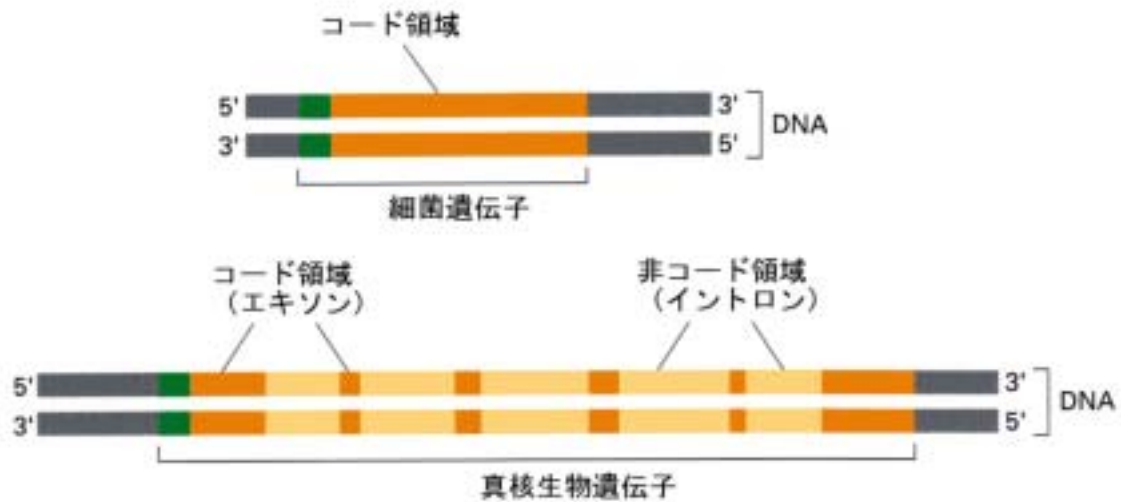
この方式は、いつもはスイッチがオンになっており、必要に応じてスイッチをオフにする方式である。

大腸菌の研究でわかったことは、発現の調節はおもに転写のレベルで行われていること、また染色体にはタンパク質をコードする領域以外に、上に述べたような調節遺伝子 (regulatory gene) があることであった。

### 3) 真核生物の場合

真核生物の染色体は、原核生物のものとは比べるとずっと複雑である。したがって、転写とそれに続く過程もずいぶん異なっている。遺伝子の発現の調節は、転写レベルだけでなく、mRNAの転写後に修飾やプロセッシングを受けることによっても調節される。

真核生物の場合、プロモーターから始まる遺伝子には、タンパク質をコードしていない領域が含まれている。タンパク質をコードしている領域をエクソン (exon) といい、コードしていない領域をイントロン (intron) (介在配列、intervening sequence ととも) と言う。



真核生物の遺伝子の構造を少し詳しく述べよう。次の塩基配列はヒトβ-グロビン遺伝子を示している。

( [Globin Gene Server](#) より )

```

CCCTGTGGAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGA
GGGCAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCT
TACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACC
ATGGTGACCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGC
AAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTATCA
AGGTTACAAGACAGGTTTAAGGAGACCAATAGAACTGGGCATGTGGAGAC
AGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTGCCTATTGGT
CTATTTTCCCACCCTTAGGCTGCTGGTGGTCTACCCTTGACCCAGAGGTT
CTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCC
TAAGGTGAAGGCTCATGGCAAGAAAGTCTCGGTGCCTTTAGTGATGGCCT
GGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCA
CTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGA
CCCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAG
GAAGGGGAGAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGAT
TGCATCAGTGTGGAAGTCTCAGGATCGTTTTAGTTTCTTTTATTTGCTGTT
CATAACAATTGTTTTCTTTGTTAATTCTTGCTTTCTTTTTTTTTCTTCT
CCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAG
GAAATATCTCTGAGATACATTAAGTAACTTAAAAAAAACCTTACACAGTC
TGCCTAGTACACTACTATTTGGAATATATGTGTGCTTATTTGCATATTCAT
AATCTCCCTACTTTATTTTCTTTTATTTTAATTGATACATAATCATTATA

```

CATATTTATGGGTAAAGTGTAATGTTTTAATATGTGTACACATATTGACC  
 AAATCAGGGTAATTTTGCATTTGTAATTTTAAAAATGCTTCTCTTTTA  
 ATATACTTTTTTGTATCTTATTTCTAATACTTTCCCTAATCTCTTCTT  
 TCAGGGCAATAATGATACAATGTATCATGCCTCTTGCACCATTCTAAAGA  
 ATAACAGTGATAATTTCTGGGTAAAGGCAATAGCAATATTTCTGCATATAA  
 ATATTTCTGCATATAAATTGTAAGTATGTAAGAGGTTTCATATTGCTAAT  
 AGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAAG  
 GCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGTAAATCATGTTTCATAC  
 CTCTTATCTTCTCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGG  
 CCCATCACTTTGGCAAAGAATTCACCCACCAGTGCAGGCTGCCTATCAGA  
 AAGTGGTGGCTGGTGGGCTAATGCCCTGGCCACAAGTATCACTAAGCTC  
 GCTTCTTGGCTGTCCAATTTCTATTAAGGTTCCCTTTGTTCCCTAAGTCCA  
 ACTACTAAACTGGGGATATTATGAAGGGCCTGAGCATCTGGATTCTGCC  
 TAATAAA AACATTTATTTTCATTGCAATGATGTATTTAAATTATTTCTGA  
 ATATTTTACTAAAAAGGAATGTGGGAGGTCAGTGCATTTAAACATAAAG  
 AAATGAAGAGCTAGTTCAAACCTTGGGAAAATACACTATATCTTAAACTCC  
 ATGAAAGAAGGTGAGGCTGCAAACAGCTAATGCACATTGGCAACAGCCCTG  
 ATGCCTATGCCTTATTCATCCCTCAGAAAAGGATTCAAGTAGAGGCTTGAT  
 TTGGAGGTTAAAGTTTTGCTATGCTGTATTTTACATTACTTATTGTTTTAG  
 CTGTCCTCATGAATGTCTTTTCACTACCCATTTGCTTATCCTGCATCTCTC  
 AGCCTTGACTCCACTCAGTTCTCTTGCTTAGAGATACCACCTTTCCCCTGA  
 AGTGTTCCTTCCATGTTTTACGGCGAGATGGTTTCTCCTCGCCTGGCCACT  
 CAGCCTTAGTTGTCTCTGTTGTCTTATAGAGGTCTACTTGAAGAAGGAAAA  
 ACAGGG GGCATGGTTTGACT.....

塩基配列は上から下へ、左から右へ並んでいて、最初の段の左端が5'側である。DNA二本鎖のうち、鋳型鎖ではない方の配列を示してある。鋳型鎖を使ってmRNAへ転写されるのだから、作られるmRNAは、上の塩基のうちTをUに変えたものである。遺伝子を一本鎖で示すときはこのような表記の仕方をする。

1) 上2段は**プロモーター領域**で、RNAポリメラーゼが結合し、転写開始部位から転写をはじめめるために必要な領域である。3か所の赤い部分は、特異的なエレメントで、特に最後の配列は**TATAボックス**と呼ばれる。

2) 水色の部分の先頭にある**ACATTTG**は**転写開始点**で、**キャップ配列**とも呼ばれる。これはこの部分が修飾を受けてキャップ構造になるからである。

3) 紺色の A T G が翻訳開始コドンである。転写開始部分から翻訳開始部分まで 50 塩基対離れているが(この値は遺伝子によって異なる)、リーダー配列と呼ばれる。

4) 紫色の部分がエクソンで、合計 3 つある。最初のエクソン 1 は 90 塩基対、すなわち 30 アミノ酸をコードしている。

5) 次はイントロン 1 で、130 塩基対からなり、アミノ酸をコードしていない領域である。

6) 次の紫色部分がエクソン 2 で、222 塩基対からなり、31 から 104 までの 74 アミノ酸をコードしている。

7) 850 塩基対からなる大きなイントロン 2 が次に続く。この部分もグロビタンパク質の構造には全く関係ない。

8) 3 番目の紫色がエクソン 3 で 126 塩基対からなり、105 から 146 までの 42 アミノ酸をコードしている。

9) 終始コドンは T A A で、紺色で示してある。

10) T A A に続く水色の部分は、3'側の転写されるが翻訳されない部分である。この部分には赤色で A A T A A A と示した配列を含んでいるが、この部分に 2 0 0 から 3 0 0 のポリ A テイルが付着する目印となる。転写されると A A U A A A となり、ここからおよそ 2 0 塩基対下流にポリ A が挿入される。ポリ A テイルは、m R N A が分解されるのを阻止し、寿命を長くしている。

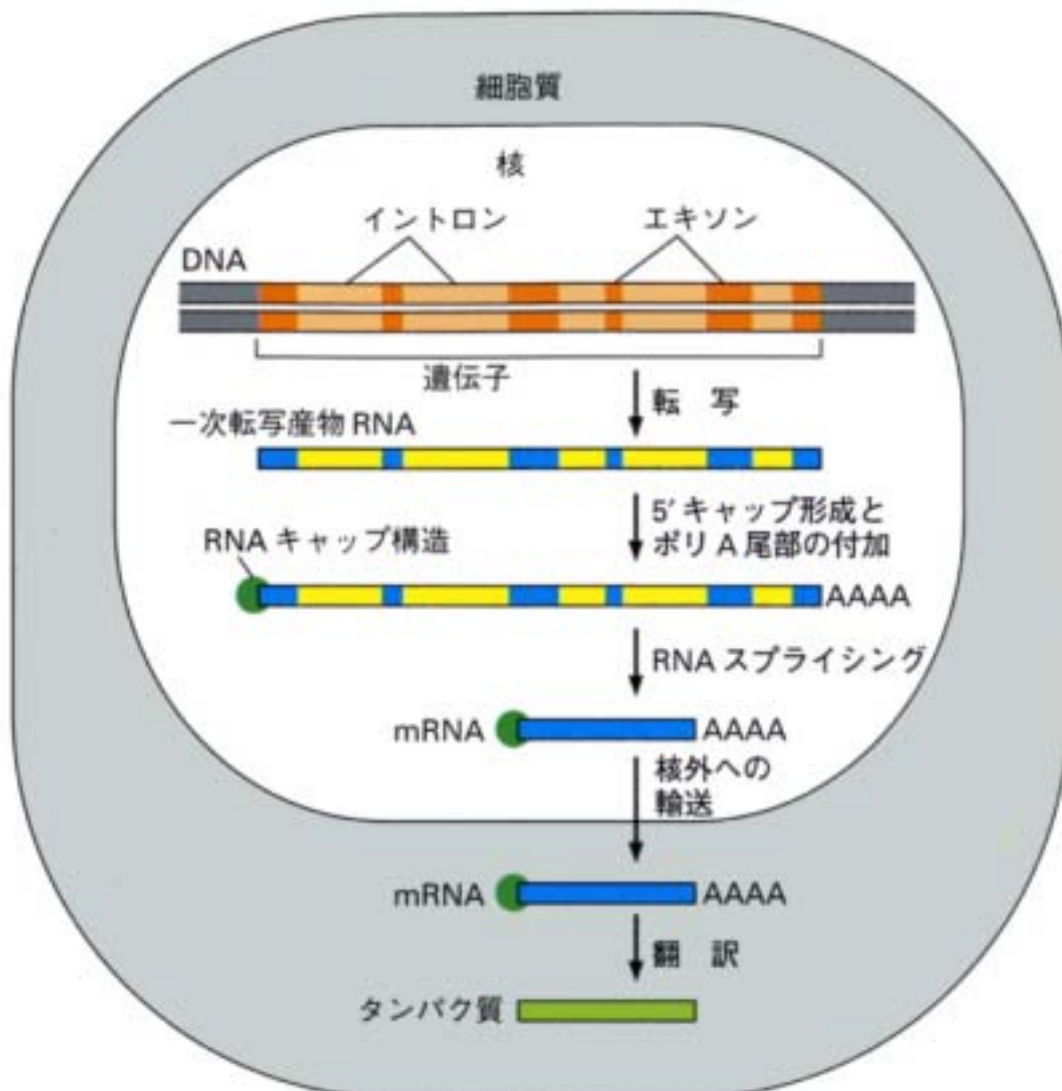
次の図は、遺伝子から、m R N A が転写される様子の模式図である。上の説明と図からわかるように、m R N A は

1) 転写開始部位から終了部位まで転写される。この配列には、エクソン、イントロンとともに翻訳されない塩基配列を両端に含む。

2) 5'側に 7 - メチルグアノシン (Cap) 構造が、3'側に 100 - 250 個の連続した A (ポリ A テイル) が付加される。

3) イントロン部分がスプライシング酵素によってつまんで切られる (スプライシング、splicing)。

## 真核生物

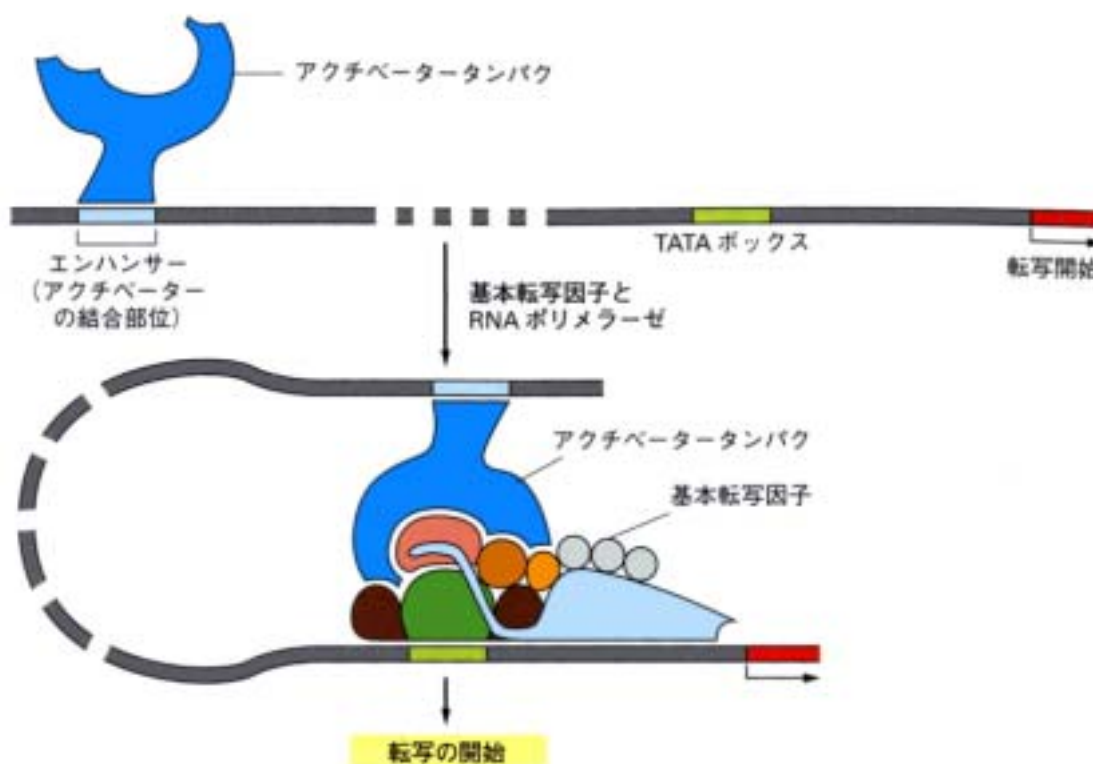


4) 完成した mRNA (成熟 mRNA) が核膜孔からサイトゾールへ出る。

真核生物の場合は、原核生物のオペロンのような構造を取らず、一つの遺伝子は一つのポリペプチド鎖しかコードしていない。したがって上の図で描かれた遺伝子内では、1つのポリペプチド鎖が3つのエクソンに分断されて配置されていることになる。どうしてこうなっているのかはよく解っていない。

真核生物でも、プロモーター領域にRNAポリメラーゼが結合してmRNAの合成が開始されるという点は、原核生物と同じである。RNAポリメラーゼが認識する塩基配列は、読み始めから30bpほど上流にあるTATAボックスと呼ばれる配列である。さらに上流に8 - 12bpからなる複数の上流プロモーター配列が存在する。

さらに、ずっと離れた数 kbp 上流にはエンハンサーと呼ばれる領域があって、ここにアクチベータータンパク質が結合し、転写の速度を高めている。



原核生物のリプレッサーのように、真核生物でも転写の調節をする調節タンパク質がたくさん存在する。これらのタンパク質はDNAと結合して、転写の開始、終了、転写量の調節を行っている。

ヒトのゲノムDNA量（半数体細胞、すなわち精子や卵あたりのDNA量）は、大腸菌のおよそ800倍で、 $3.2 \times 10^9$ の塩基対を含むが、遺伝子の数は5万から10万くらいであろうと考えられている。平均300個のアミノ酸からなるタンパク質を作る遺伝子はコード領域だけで平均 $300 \times 3 = 900$ 塩基対必要だから、10万個の遺伝子のためには $9 \times 10^7$ 塩基対必要となる。しかしこれでも全ゲノムの3%にしかすぎない。イントロンや制御領域がコード領域の10倍あったとしても、たかだか30%にしかすぎない。あとは余分なDNAということになる（冗長、redundancy）。

真核生物のDNAを調べてみると、反復配列と呼ばれる、繰り返しが多く見つかる。単なるジャンクなのだろうという考えや、遺伝子重複で増えてしまったが特に支障がないので保存しているという考え、何か働きがあるという考えなどがあり、なぜそのようなものがあるかよく解っていない。また、このなかには、今は機能を発揮していないが、機能的な遺伝子の配列と非常によく似た塩基配列を持った領域が見つかる（偽遺伝子、pseudogene）。

---

## 2. 遺伝子工学

1970年代の中ごろから始まった組み換えDNAテクニック（recombinant DNA technology）（あるいは応用技術までを含めて遺伝子工学 genetic engineering とも言う）の発展は、いまや生物学の研究に欠かせない技術の一つになった。ここではその方法について概観してみよう。

<http://esg-www.mit.edu:8001/esgbio/rdna/rdnadir.html>

### 1) タンパク質より DNA のほうが扱いが簡単

タンパク質はアミノ酸が一つながりにつながったものである（一次構造、primary structure）。したがって、タンパク質を抽出して、純化し、端から順番にアミノ酸の配列を決めてゆけば、タンパク質の一次構造を解明することができる。

しかしながら、タンパク質は変性しやすかったり、目的とするタンパク質が生体内では少量しか存在しないために、抽出して純化するのが難しかったりする場合が多いので、タンパク質からアミノ酸の配列を決めるのは容易ではない。

アミノ酸の配列を決めているのは、DNAの遺伝情報であり、そこから転写されたmRNAの塩基の配列である。遺伝の暗号が解った時点で、塩基の配列が決まれば暗号をもとに戻してアミノ酸の配列を推定できるようになった。

そこで、DNAのタンパク質をコードしている部分を切り出したり、mRNAを取り出してそれをDNAに置き換えて、大腸菌などを使ってDNAを増量し、塩基の配列を決める方法が開発された。

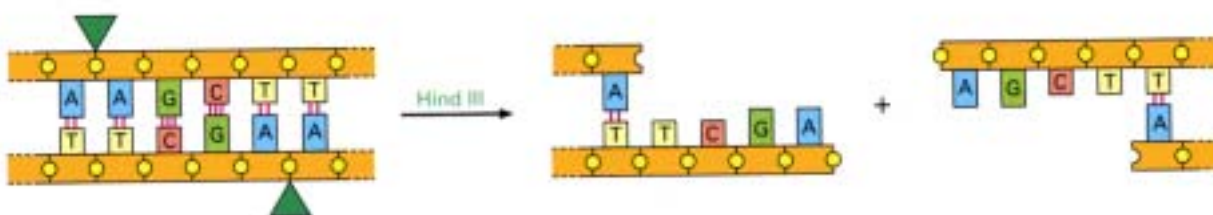
### 2) 鋏 = 制限酵素（restriction enzyme）と糊 = DNAリガーゼ

染色体DNAは連続した大きな分子である。これをまるごと取り扱うには大きすぎる。タンパク質のように適当な大きさの単位の方が扱いやすい。幸いなことに、DNA上では、



調節領域やコード領域が不連続（飛び飛びに存在する）であることが解り、うまく切り出せば取り扱い得ることが解った。

ちょうどその頃、塩基の配列の特定の場所でDNAを切断する酵素が細菌で見つかった。下の図は HindIII という制限酵素が、5'-AAGCTT-3'の回文構造をした塩基配列を認識して、楔形の位置でDNAを切断する模式図である。



このような切断の仕方をするので、両端に糊代ができることになる。

上の図の切断と反対のはたらきをするのが、DNAリガーゼである。糊付けしたいDNA分子の両方に5'-AAGCTT-3'という配列があれば、DNAリガーゼを使ってDNA分子をつなぎ合わせることができる。

DNA分子中に別のDNA分子断片を組み込むためには、同じ制限酵素で切断して糊代を両方のDNAにつくり、DNAリガーゼで糊付けすればよい（組み換え）。

認識する塩基配列の数や組み合わせの異なる制限酵素が、多数発見されている。複数の制限酵素を使って染色体DNAを切断し、電気泳動装置にかけて大きさで分け、制限酵素切断箇所によるDNA断片の配列地図を作ることができる。

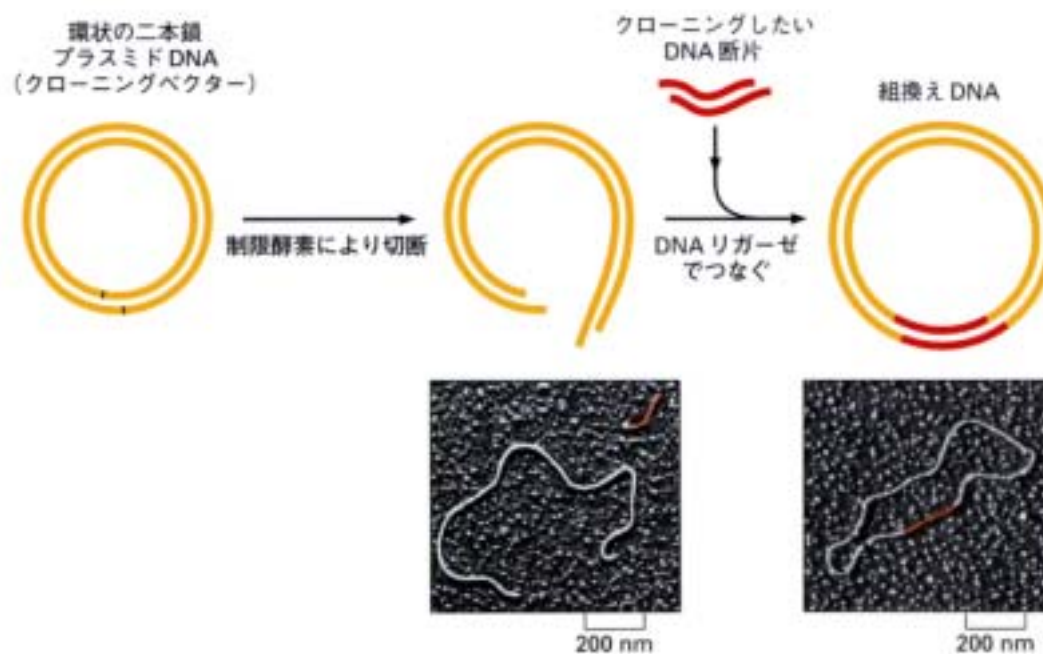
こうして、DNAを適当の大きさに切断したり、つなぎ合わせたりすることができるようになった。

### 3) プラスミドベクターとライブラリー

こうして、染色体DNAから適当な大きさのDNA断片を手に入れることができたなら、これを大腸菌のプラスミドに、上に述べた銚と糊を使って組み込み、大腸菌に導入することができる。

プラスミドとは、宿主染色体とは独立に存在して、自律的に増殖し、子孫に伝えられて行くDNA分子をいう（核外遺伝子）。DNA断片を大腸菌内に運び込む、このようなプラ

スミドをベクター (vector、運び屋) と呼ぶ。



薬剤耐性遺伝子を持ったプラスミドをベクターとして使い、その薬剤を含む培地で大腸菌を培養すれば、プラスミドが導入された大腸菌だけが生き残る。

こうして作成した大腸菌クローンの集合をライブラリーという。今まで述べた方法では、染色体ゲノムから出発しているから、ゲノム DNA ライブラリー (genomic DNA library) である。

#### 4) プローブ (probe) の利用

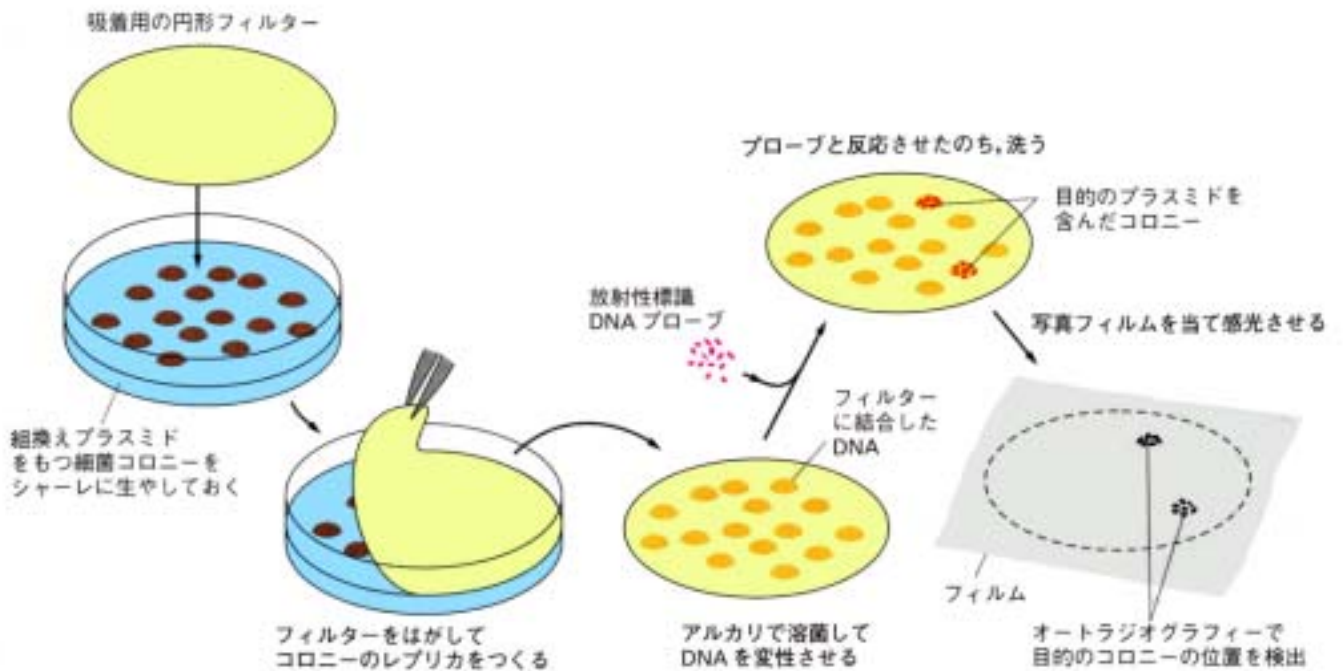
本当に読みたい本 (知りたいと思っている遺伝子を含むクローン) をライブラリーから探すためには、内容を照合する必要がある。そのために利用されるのがプローブである。

既知のタンパク質の遺伝子を探すためには、そのアミノ酸配列から mRNA の塩基の配列を推定する。なるべく対応するコドンの少ないアミノ酸 (メチオニンやトリプトファンやチロシンなど) が続く部分が選ばれる。

こうして人工的に相補的 DNA を合成し、放射性同位元素で標識してプローブを作成する。このプローブを使えば、目的のタンパク質をコードする DNA 領域 (あらかじめ処理を施し一本鎖にする) と水素結合を作るので、目的とする DNA を同定することができる。



目的とする DNA のプローブを使えば、ライブラリーのどのクローンに目的のタンパク質をコードする DNA が含まれているかを知ることができる。



## 5) cDNAライブラリー

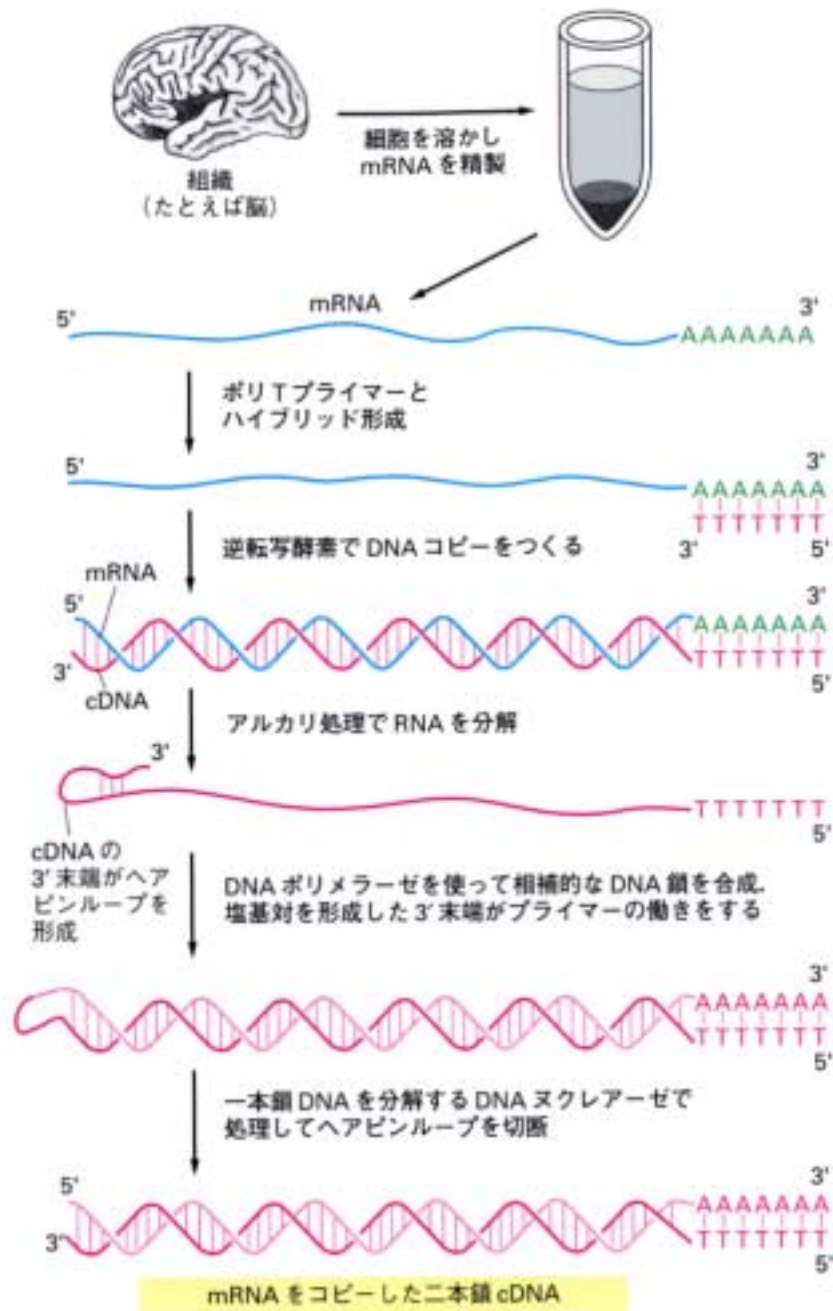
これまで述べたのは、染色体からDNAを切り出し、それを大腸菌へ導入し、目的とするDNA断片を含むクローンを探し出す方法だった。

しかしこれはショットガン法とも言われるように、無駄も多いし、目的のものを探し出すのも大変である。もう一つのライブラリーの作り方は、mRNAから作る方法である。

特定の臓器や発生の特定の時期に、特定のタンパク質が多量に合成されることが解っていれば、そこからmRNAを取り出し、相補的DNAを逆転写酵素(reverse transcriptase)を使って合成することができる。これは一本鎖なので、DNAポリメラーゼを使ってこれを二本鎖にし、両端に糊代を付けてプラスミドへ組み込むことができる。

あとは上に述べた方法と同じである。こうしてcDNAライブラリー(cDNA library)を作ることができる。

cDNAライブラリーがゲノムDNAライブラリーと大きく異なる点は、cDNAライブラリーが染色体の遺伝子の転写された領域、しかもエクソン領域のみのライブラリーであることである。



## 6) DNA塩基配列の決定

こうして、目的とするDNAを含む大腸菌クローンが特定できたら、大腸菌を培養して大量に増やし、塩基の配列の決定に回すことができる。最近では、PCR (polymerase chain reaction) 法を使って、少量のDNA断片から比較的簡単に量を増やすことができるようになった。

塩基配列を決定する方法はいくつか開発されているが、ここでは教科書に書かれた方法を説明する。目的とする一本鎖のDNA断片とDNAポリメラーゼ、DNA合成の原料となるdATP、dTTP、dGTP、dCTP、それとラジオアイソトープ(RI)で標識したddATPをインキュベートして、DNA合成を行う。すると合成されるDNAの鎖の中に、dATPのかわりにddATPが原料としてランダムに取り込まれてしまい、そこでそれ以上の鎖の伸長が止まってしまう。

4つの塩基でこの操作を繰り返し、4つ並べてポリアクリルアミドゲル電気泳動(PAGE)を行うと、鎖の短いものほど原点から遠くまで泳動する。4本のカラムをまとめてX線写真フィルムに載せて感光させ現像すると、RI標識ddヌクレオチドのあるところがバンドとして現れてくる。

下から順番に4つのカラムのバンドをたどって行くと、この並び方が元のDNAの塩基の配列の3'→5'の相補的配列になる。最後に、制限酵素による地図をもとに、元の遺伝子の全体像を再構築する。

塩基の配列が決定されると、遺伝子データベースに登録する。ここには世界中から集められた遺伝子の塩基配列が登録されていて、コンピュータネットワークを介して照合することができる。自分の見つけた遺伝子がどの遺伝子の塩基配列と似ているかなども、すぐに調べることができる。

#### 7) 遺伝子導入動物 (transgenic organism)

遺伝子DNAを動物に導入することも可能となった。マウスに成長ホルモン遺伝子を導入し、スーパーマウスを作ったというニュースを聞いたことがあるだろう。

<http://www.ultranet.com/~jkimball/BiologyPages/R/RecombinantDNA.html>