

第4回 ゲノムと遺伝子

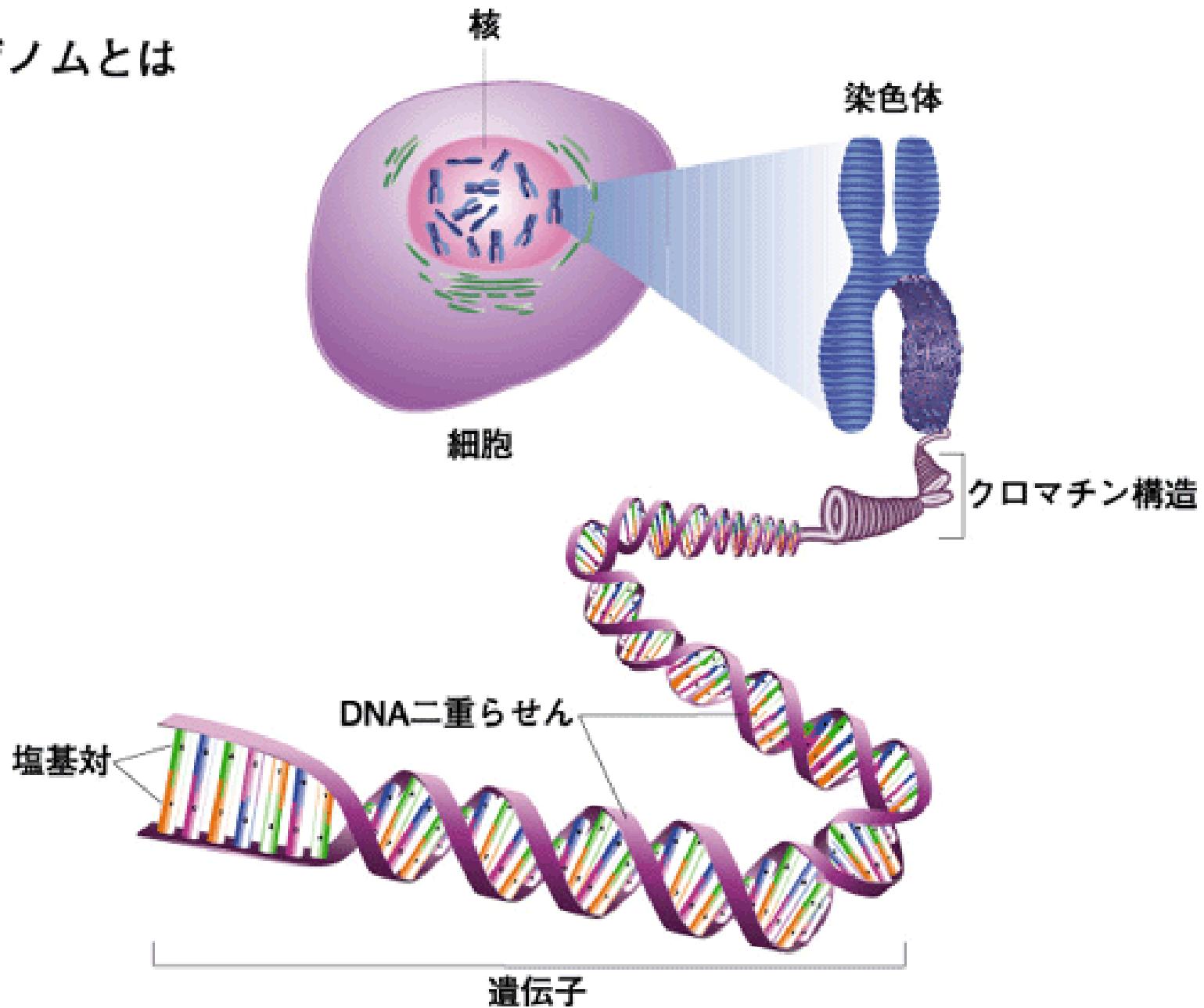
- ゲノムの概念
- 全塩基配列が公開された
ゲノム

分子生命化学教室
荒牧 弘範

ゲノム (genome) とは

- ・ゲノム(genome)という言葉は、遺伝子(gene)と染色体(chromosome)の合成語である。
- ・生物の構成単位である細胞のなかには、染色体と呼ばれる遺伝子をのせた集合体が存在する（ご存知の通り、ヒトでは23対の染色体が存在する）。

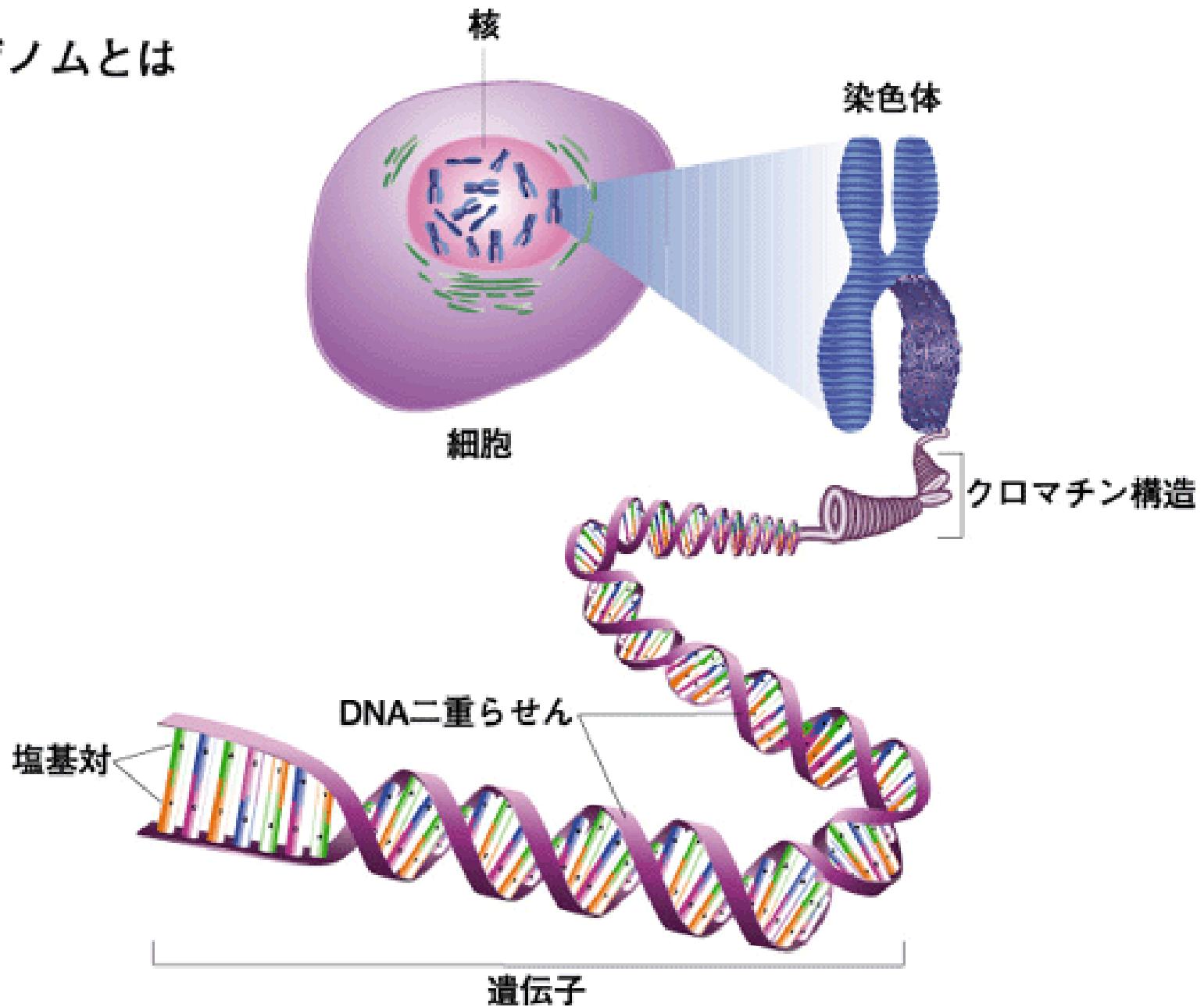
図1 ゲノムとは



ゲノム (genome) とは

- ・一つ一つの染色体には、DNA鎖が、コイル状にぐるぐる巻きになって（これをクロマチン構造と呼びます）収まっている。
- ・このDNA二重らせんのなかに、生物が持つすべての機能や活動をコントロールするための指示が（暗号のように）並んだ部分が含まれており、これを遺伝子と呼ぶ。

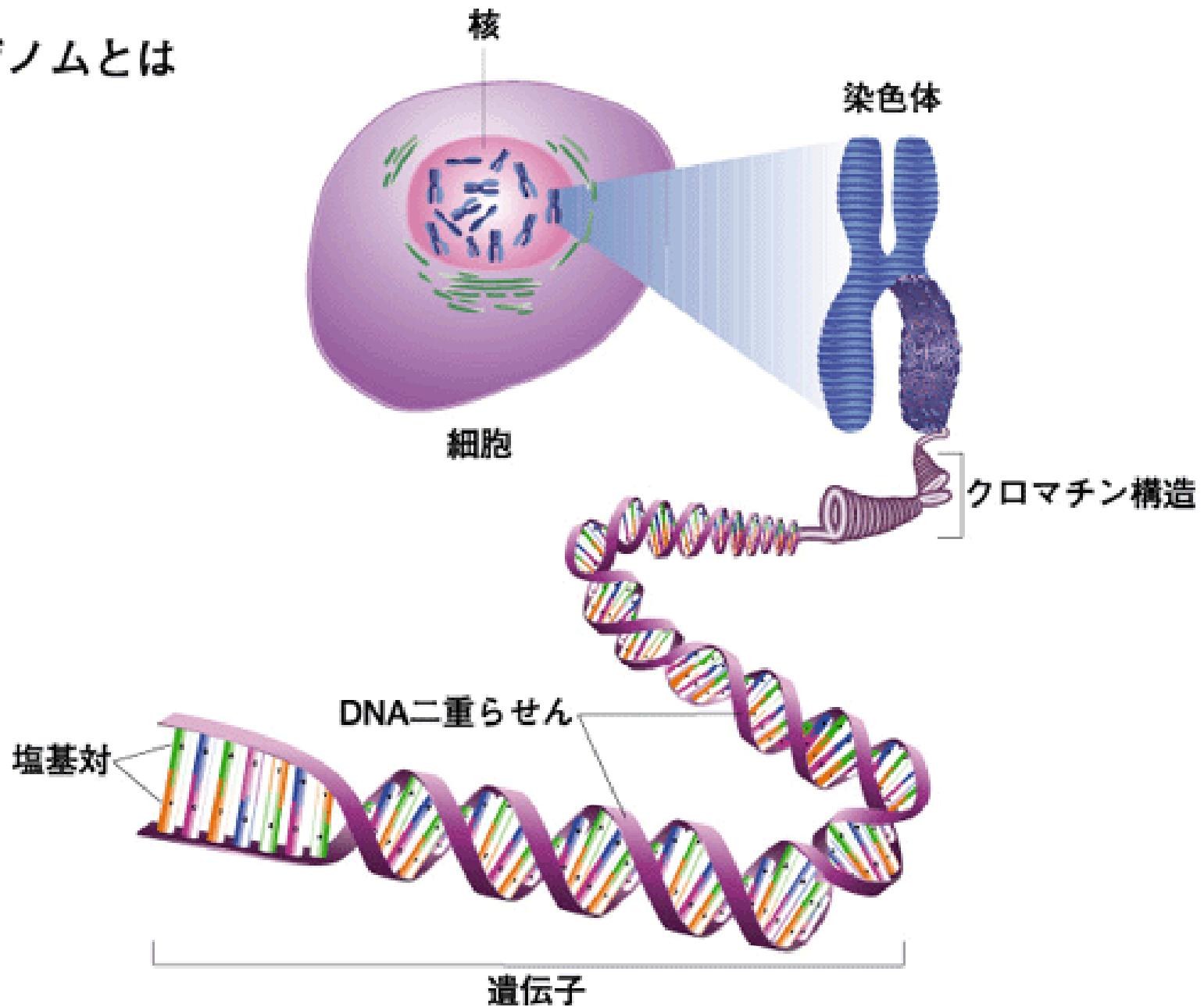
図1 ゲノムとは



ゲノム (genome) とは

- また、DNA二重らせんの内の10%弱が遺伝子部分に相当すると言われている。最近、ヒトの場合、この遺伝子が約2万2千個存在すると言われている。
- ゲノムとは、上に述べた染色体から遺伝子にいたる一つのシステムをセットにして呼んでいる。

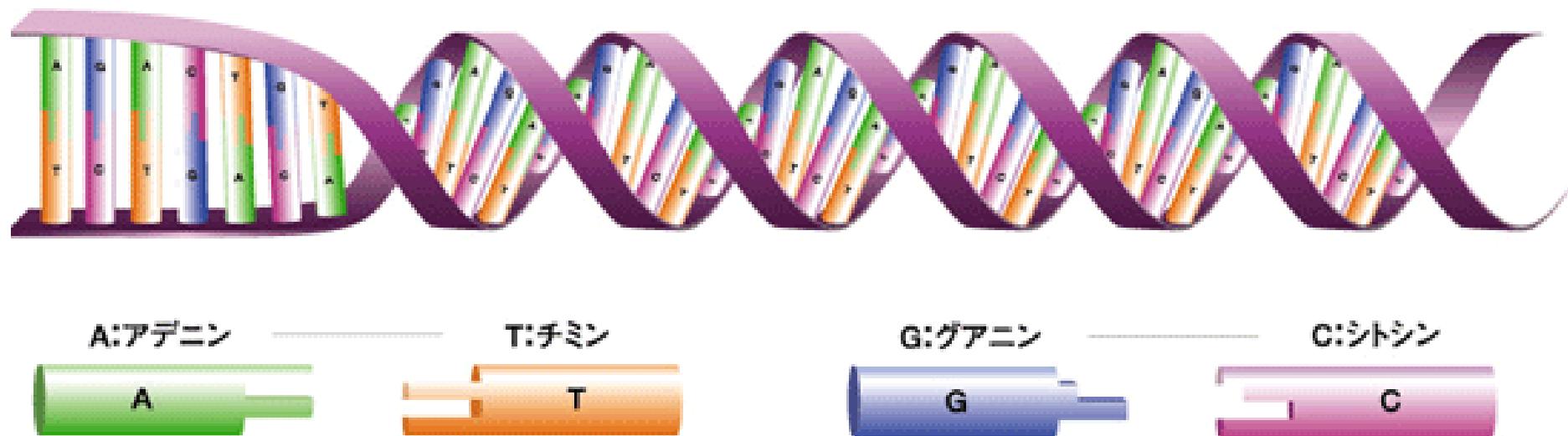
図1 ゲノムとは



遺伝子上の暗号とその役割

- DNA二重らせんには、遺伝情報が暗号のように保存されている。暗号として使われている文字は4種類の塩基（A：アデニン、G：グアニン、C：シトシン、T：チミン）の対からできており、A-TまたはG-Cの塩基対が存在する。
- ヒトで、この塩基対の数は約30億個であることがわかっている。

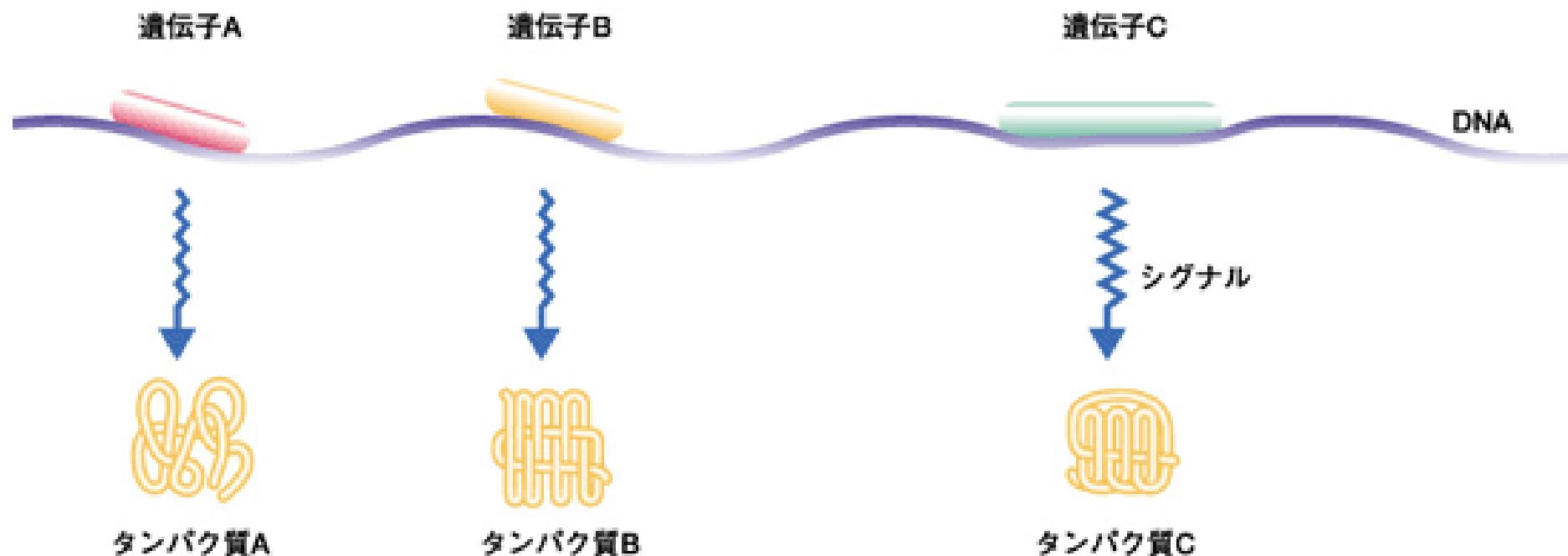
図2 遺伝子上の暗号とその役割



遺伝子上の暗号とその役割

- 生物の機能・活動に関与する営みは、すべてさまざま機能を持ったタンパク質（例えば、ホルモン、酵素、神経伝達物質など）によって行われているが、この生産が個々の遺伝子によってコントロールされている。
- 一つの遺伝子は数万～十数万個の塩基対からなっており、これによって、どの細胞で、どのタイミングで、どのようなタンパク質を、どの程度作るかの情報伝達が行われている。

図3 遺伝子上の暗号とその役割



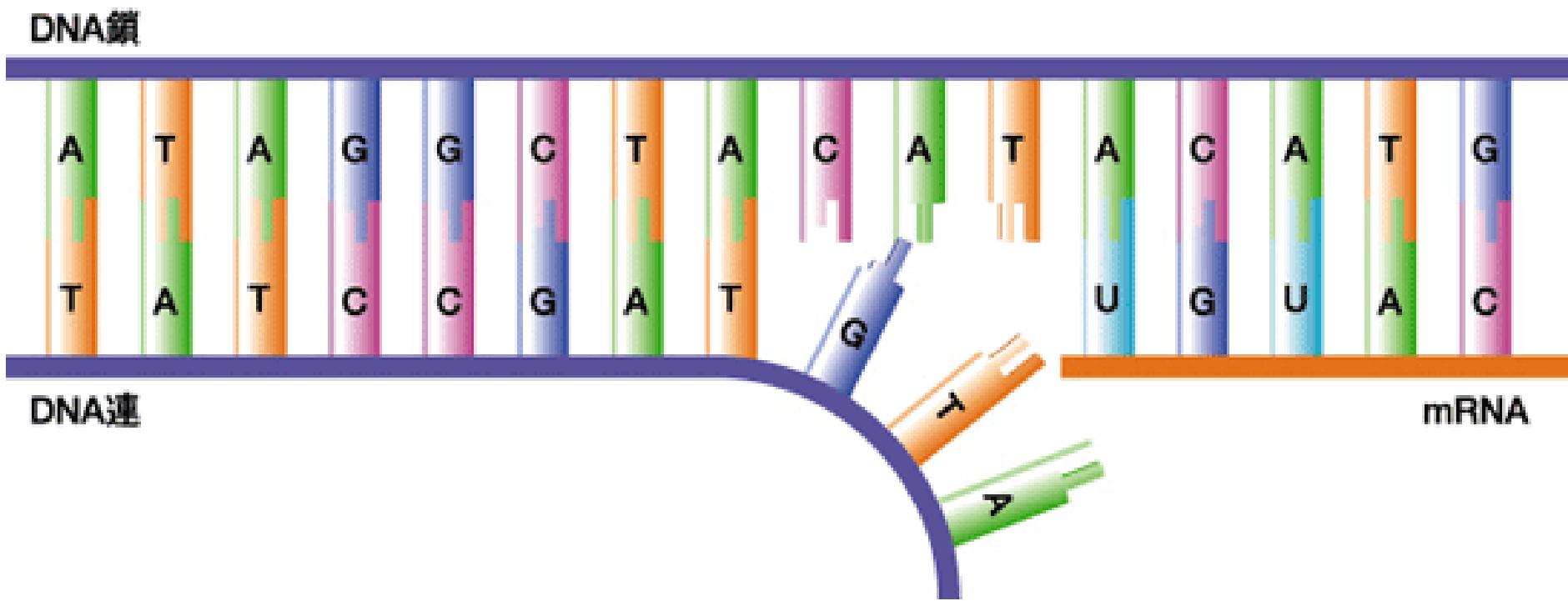
遺伝子による情報=あるタンパク質を $\begin{pmatrix} \text{どの細胞で} \\ \text{どのタイミングで} \\ \text{どれくらい} \end{pmatrix}$ つくるかをコードしている

遺伝子情報の伝達の仕組み

- ・二重らせんのリボンに埋め込まれた遺伝子情報が、どのように伝わって、実際にタンパク質の生成が始まるのか？
- ・タンパク質は約20種類あるアミノ酸が繋がって出来ている物質であるから、どんなアミノ酸をどの順で、どのくらい繋げるかという指示が遺伝子情報として伝達されるので、ここでは、DNAおよびRNAによる転写→プロセシング→移動→翻訳といったプロセスが存在する。

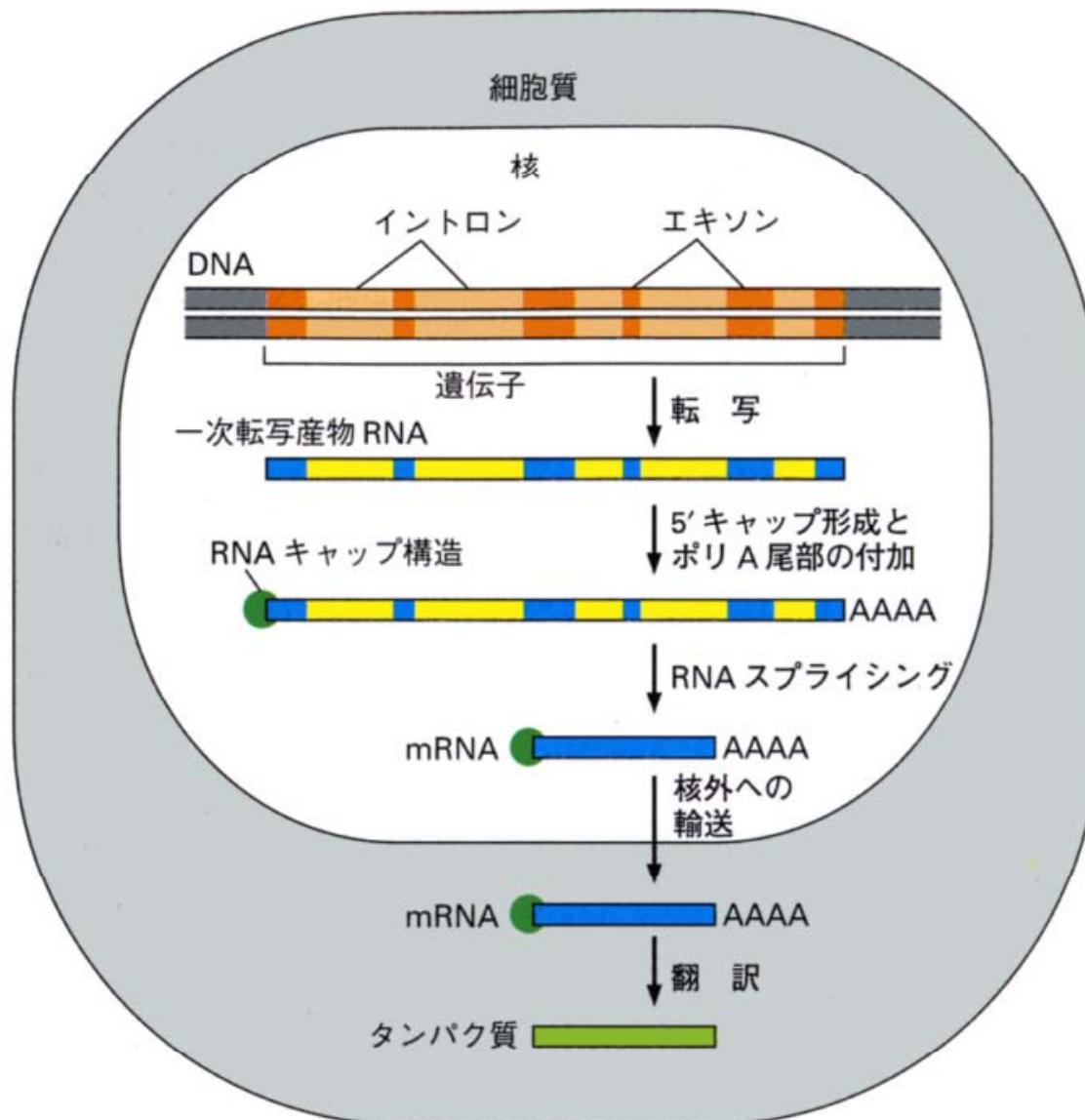
1. 転写

- 二重らせん鎖の形で存在するDNAは、情報伝達を行う際、このリボンが1本ずつに分かれ（塩基対が離れます）、その上に並んでいる塩基に対応する塩基（AであればU,TであればA, GであればC, CであればGが）がmRNAとよばれる1本のリボンの上に「転写」される。
- DNAに含まれる塩基がA, G, C, Tであるのに対して、RNAでは、含まれる塩基はA, G, C, U（ウラシル）となり、DNAでのTのかわりにUになっている。

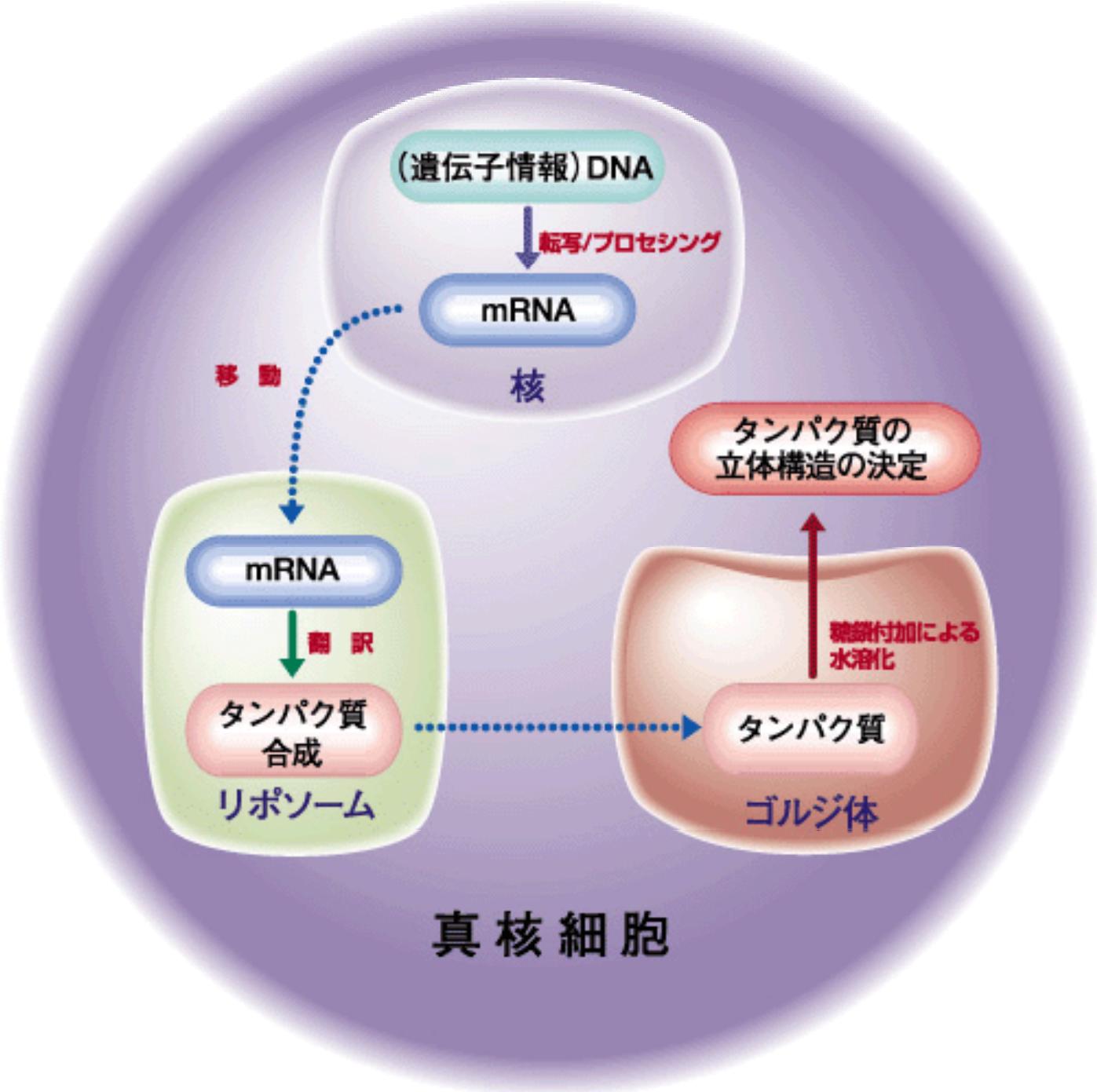


- 前にも述べたように、DNA鎖のうちタンパク質の生成に関わる遺伝子情報を含んでいるのは、全体の10%弱程度（この部分をエキソンと呼びます）です。エキソン部分だけが選択的に集約・圧縮され、情報が含まれない部分（イントロンと呼びます）が除かれる。この過程が「スプライシング」です。

真核生物

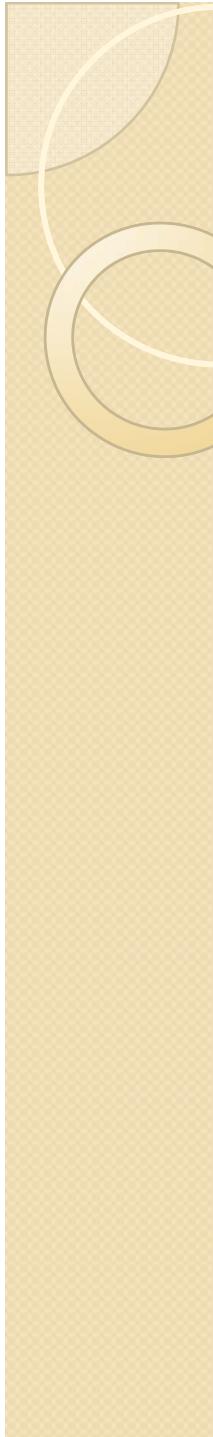


- このように遺伝情報が転写され、圧縮されたmRNAは、核から飛び出して同じ細胞内にあるリボソームと呼ばれる部分へ「移動」します。
- ここで「翻訳」が行われます。
- 最後に、生成されたタンパク質は、細胞内でゴルジ体と呼ばれる部分に移動して糖鎖が付加されることで水溶性となります。



《遺伝情報の仕組み》 コドンと それに対応するアミノ酸一覧

		2番目の塩基					
		T	C	A	G		
1番目の塩基	T	Phe	Ser	Tyr	Cys	T	3番目の塩基
	T	Phe	Ser	Tyr	Cys	C	
	C	Leu	Ser	Stop	Stop	A	
	C	Leu	Ser	Stop	Trp	G	
	C	Leu	Pro	His	Arg	T	
	C	Leu	Pro	His	Arg	C	
	A	Leu	Pro	Gln	Arg	A	
	A	Leu	Pro	Gln	Arg	G	
	A	Ile	Thr	Asn	Ser	T	
	A	Ile	Thr	Asn	Ser	C	
	A	Ile	Thr	Lys	Arg	A	
	Met	Met	Thr	Lys	Arg	G	
	G	Val	Ala	Asp	Gly	T	
	G	Val	Ala	Asp	Gly	C	
	G	Val	Ala	Glu	Gly	A	
	G	Val	Ala	Glu	Gly	G	



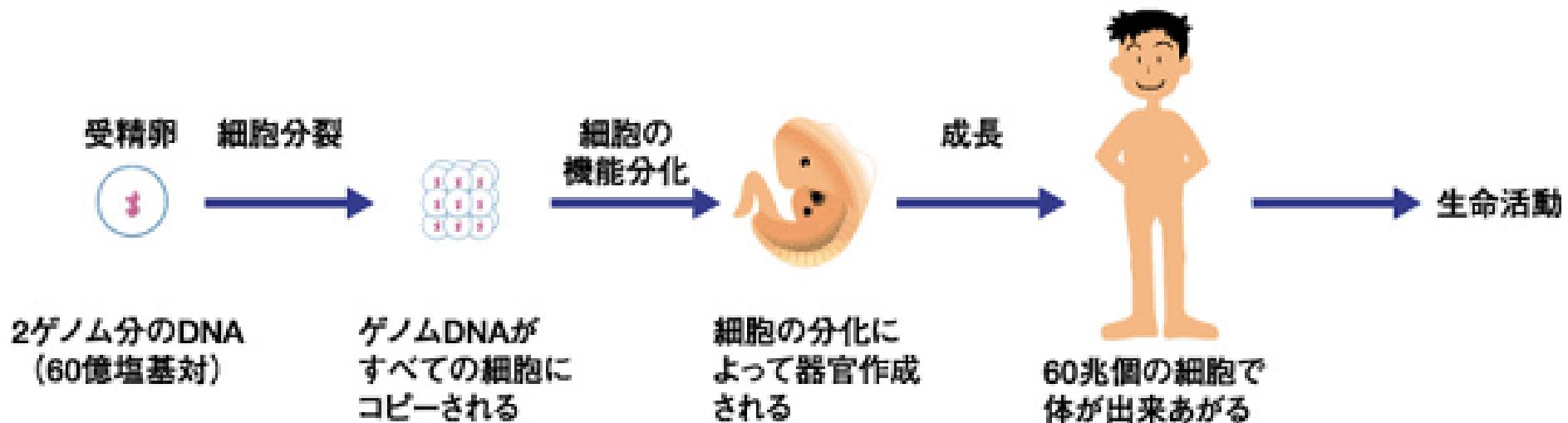
遺伝子と生命活動について

- 以上に述べたような精巧な仕組みによって、生命体の活動に関与する全てのタンパク質が、特定の場所（臓器や器官、細胞など）で、必要なタイミングで、必要なだけ生産されている。遺伝子が「生命の設計図」、「生命の司令塔」とも呼ばれる所以はある。

遺伝子と生命活動について

- ヒトは、約60兆個の細胞で構成されている。そして個々の細胞にあるヒト遺伝子は、その細胞にどんなタンパク質を生産すべきかを命令することで、どんな役割をする細胞になるかを決めている。その結果として、骨細胞は骨を形成・維持し、筋肉細胞は筋肉を形成し、内臓にある腺細胞は必要に応じたホルモンや酵素などを生産・分泌するように働いている。

図4 遺伝子と生命活動について





遺伝子と生命活動について

- ヒトが誕生して、成人するまで継続的に成長し、さまざまな活動を行いながら生命活動を営んでいく基本的な仕組みは、すべて**ゲノム**に刻み込まれた遺伝情報として細胞一つ一つの染色体のなかにしまい込まれている。そして、この情報は親から子へ、子から孫へと永遠と引き継がれ人類の歴史をなしている。



遺伝子と生命活動について

- ・一方、その生物がどのような生物となるかを決定しているのも遺伝子であり、ヒトがヒトであること、サルがサルであること、バクテリアがバクテリアであることを決定しているのも、その細胞にある遺伝子です。



全塩基配列が決定されたゲノム

http://www.genome.ad.jp/kegg/java/org_list.html

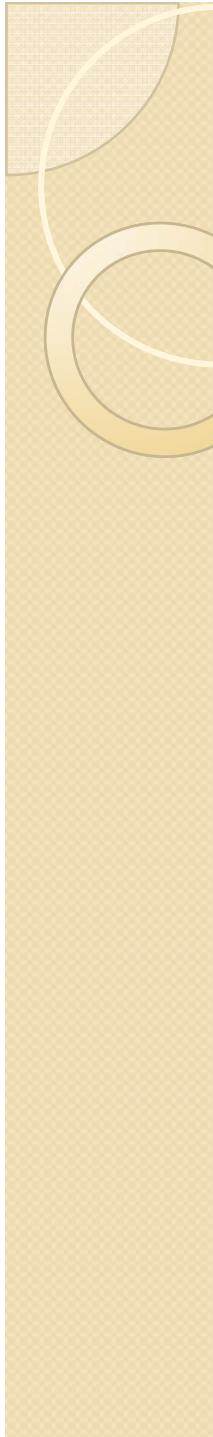
ヒトゲノム情報の解明

- (1) 遺伝子のマッピング：個々の遺伝子が、ヒトで23対ある染色体のうちで、どの染色体のどの位置にあるのかを調べること。
- (2) 遺伝子の機能解明：解読した塩基配列のどこからどこまでが遺伝子であり、それが人体においてどのような働きを司っているのかを調べること。



ポスト・ゲノム時代の研究の方 向

1. cDNA（相補型DNA）解析
2. バイオインフォマティクス
3. プロテオーム解析

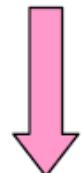


I) cDNA（相補型DNA）解析：

- cDNAはmRNAのコピーであり、遺伝子だけを効率よく研究するためには必須の技術である。

◆メッセンジャーRNA (mRNA)

タンパク質の配列情報をリボソームへ運ぶ伝令役のRNAである。

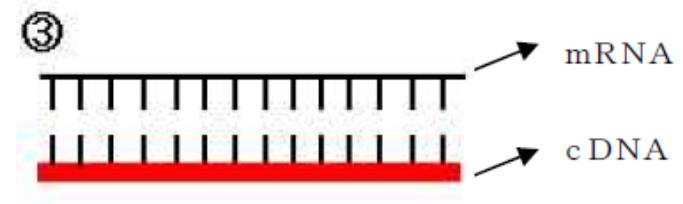
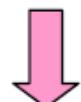


◆タンパク合成

リボソーム内でmRNAの運んできた配列情報からタンパク質が合成される。

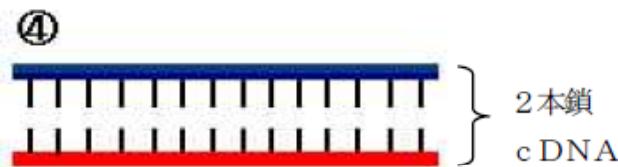
◆逆転写

逆転写酵素によりmRNAからDNAへ配列情報を逆転写し相補的な1本鎖DNA (c DNA) を合成する。



◆2本鎖c DNA

1本鎖を2本鎖にする酵素(DNAポリメラーゼ)により2本鎖c DNAとする。



◆c DNAクローン

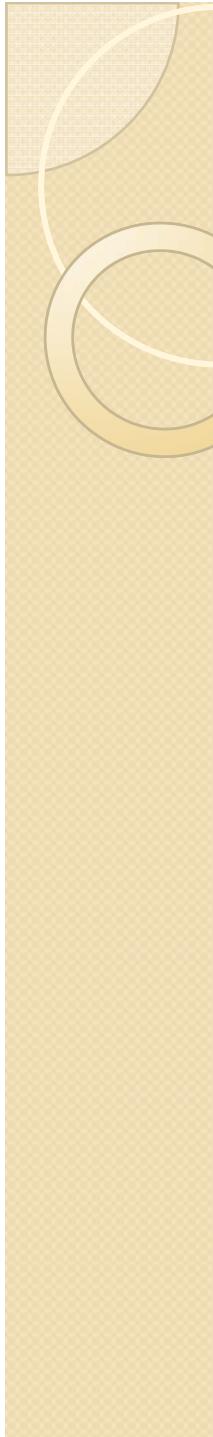
大腸菌に入れてクローンとする。

2)バイオインフォマティクス：

- 生命情報科学などと訳されているが、生命化学と情報科学の境界に生まれた新しい研究分野である。
- 現在、生命関連のデータベースは、DNA配列、アミノ酸配列、タンパク質立体構造などですが、これらのデータを効率良くコンピュータ処理して、新薬に結びつく情報を引き出す研究、技術をいう。
- <http://www.ncbi.nlm.nih.gov/>

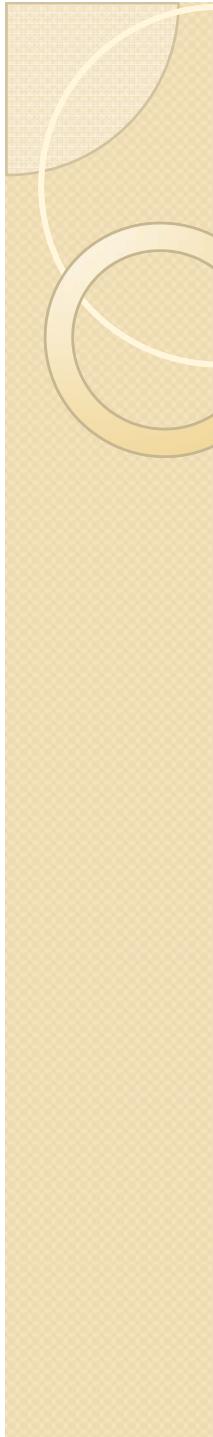
3) プロテオーム解析：

- プロテオーム (proteome) とはprotein (タンパク質) とgenome (ゲノム) を組み合わせた造語であり、ゲノムが一個の生物の持つ全ての遺伝情報を指すのに対し、プロテオームは、細胞内で発現している（発現する可能性をもつ）全タンパク質のことを指す。



3) プロテオーム解析：

- 病態と正常の細胞中のプロテオームを比較することで、疾患に関係しているタンパク質を見出すことが出来る。

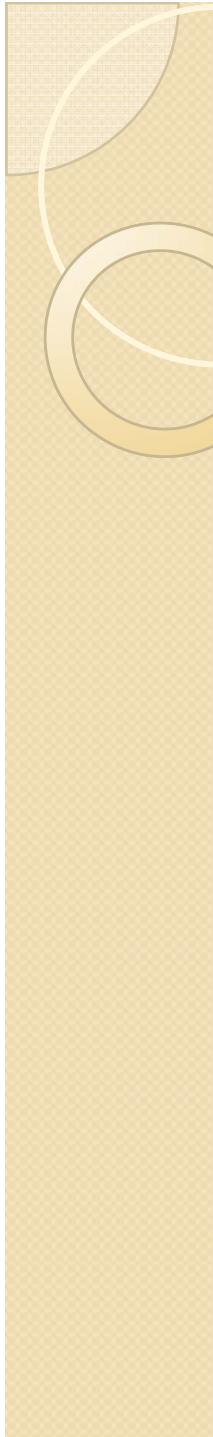


3) プロテオーム解析：

- 生体内では、多数の遺伝子から作られる多様な種類のタンパク質が働いて生命活動を支えているが、それらのタンパク質の構造や機能は、これまで膨大な時間と労力を費やして個別に単離して解析するしかなかった。

3) プロテオーム解析：

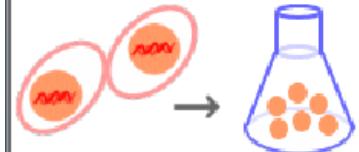
- しかし近年になって、微量のタンパク質断片の質量を正確に測定することができるようにになりましたので、種々の電気泳動やクロマトグラフィーを組み合わせて分離された少量のタンパク質を断片化してその質量を測定し、得られたデータをゲノム解析から推定されるタンパク質のアミノ酸配列データと比較して同定することが可能になりました。



プロテオーム解析の流れ

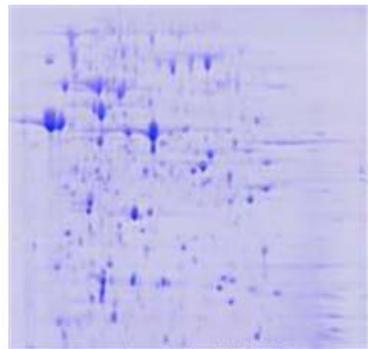
- (1) タンパク質の発現
- (2) タンパク質の分離
- (3) タンパク質の同定
- (4) ゲノムデータベースへの反映、
という手順を経て行われます。

(1)タンパク質の発現



目的とする微生物を特定の条件下で培養し、タンパク質を発現させます。その後、培養した微生物から発現したタンパク質を抽出します。

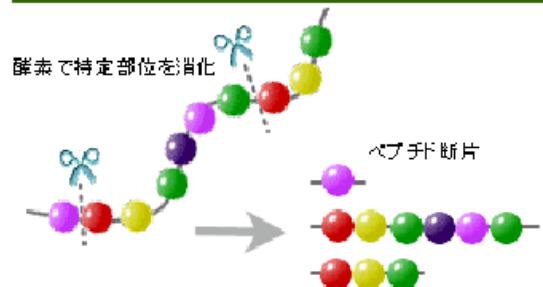
(2)タンパク質の分離



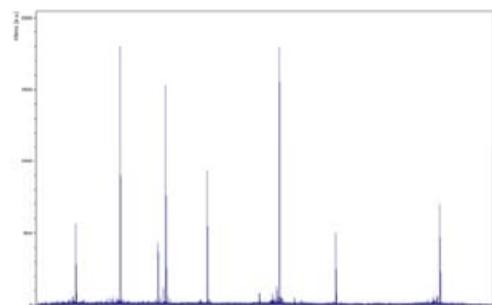
抽出したタンパク質(混合物)を二次元電気泳動を用いて個々のタンパク質に分離します(左図)。

二次元電気泳動像

(3)タンパク質の同定



分離した各タンパク質のスポットを切り出し、消化酵素を用いてタンパク質を切断します(左図)。タンパク質を切断すると、数～数十アミノ酸からなる断片になります。この断片をペプチド断片といいます。



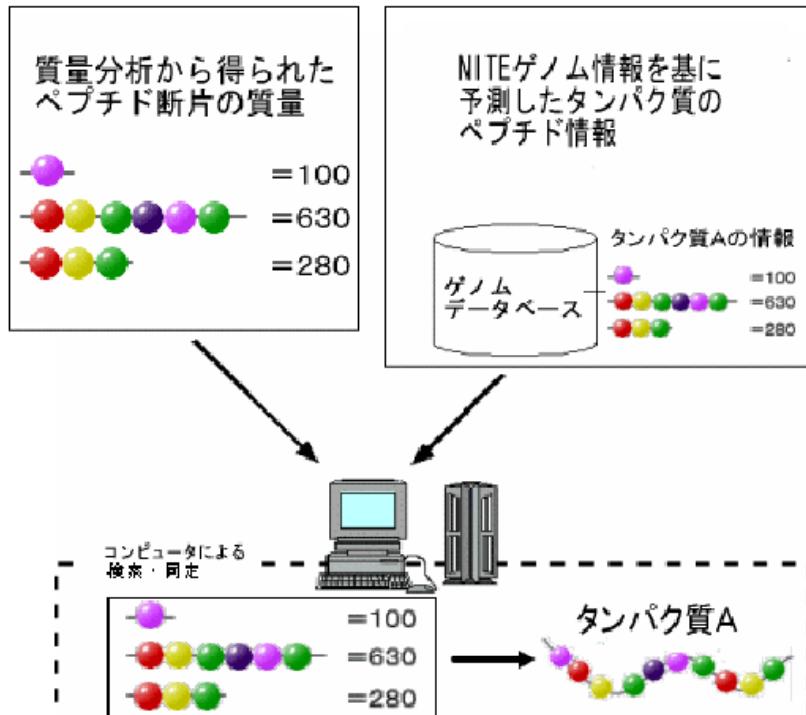
MSスペクトルデータ

質量分析計を用いて、ペプチド断片の質量(分子量)を測定します。質量分析の結果は、左図のようなMSスペクトルデータとして得られます。



質量分析計の例(MALDI-TOFMS:マトリックス支援レーザー脱離イオン化飛行時間型質量分析計)

質量分析から得られたペプチド断片の質量とゲノムデータから予測されるタンパク質のペプチド情報(質量等)を比較し、目的とするタンパク質を同定します。



(4)ゲノムデータベースへの反映

NITEでは、解析データの信頼性を高めるために、ペプチドマスフィンガープリント法の他、ショットガンプロテオーム法、多次元クロマトグラフシーケンス法、N-末端アミノ酸シーケンス法を用いて解析を実施しています。

得られたタンパク質情報は、NITEのゲノムデータベースに反映させ、ゲノム情報の高付加価値化を図ります。

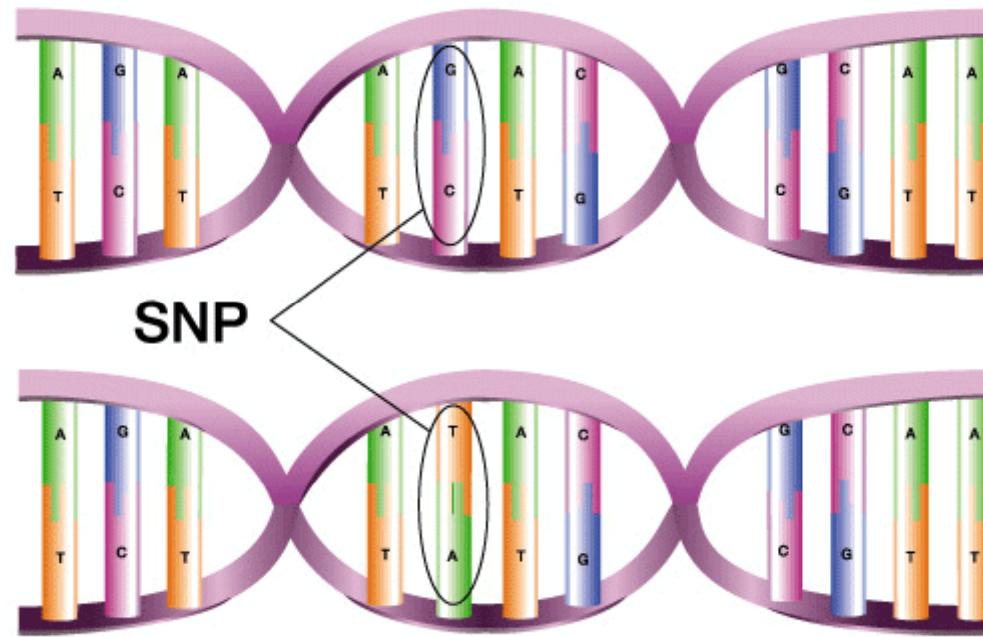
3) プロテオーム解析：

- ・ このようにして、生物のもつタンパク質の構造や機能を網羅的に解析することをプロテオーム解析という。
- ・ ゲノム解析で得られる情報は、特定のタンパク質を作る可能性をもった遺伝子があるという情報であるから、プロテオーム解析によって実際に細胞内で働いているタンパク質についての情報を付加することにより、生物のゲノム情報は産業に利用しやすくなる。

さらに応用面で注目されるものとして

- SNP（スニップ）：遺伝的に継承される塩基の置き換わり（单塩基多型）のことを言い、個人の体質に関係しています。
- DNAチップ（ジーンチップまたはDNAマイクロアレイともいいます）：シリコンチップまたはガラス版上に塩基配列がすでに分かっている一本鎖DNAを整列化/固定したものです。SNP解析、遺伝子鑑定、固体識別、疾患関連遺伝子研究などに用いられています。

Single Nucleotide Polymorphism (SNP)



- 2人以上のヒトのゲノムを比較すると
- 3 0 0 0 M B のうち、99.9%は塩基配列が同じ
 - 残り 0.1%（約 3 0 0 万ベース）は個人間で差

遺伝子変異解析

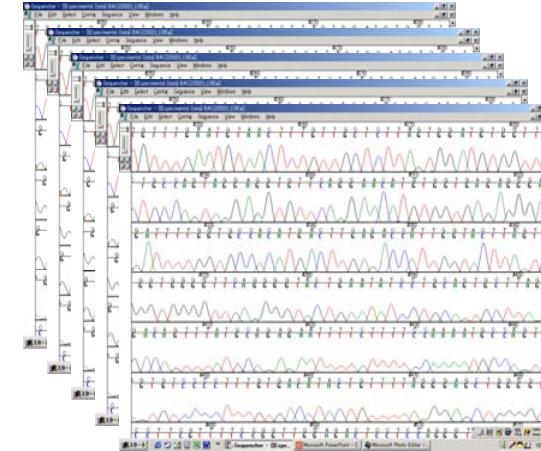
サンプル：ゲノムDNA



特定領域增幅用オリゴデザイン



PCRにて特定領域を増幅



データ解析、比較により変異部位検出



DNAシークエンサー配列決定

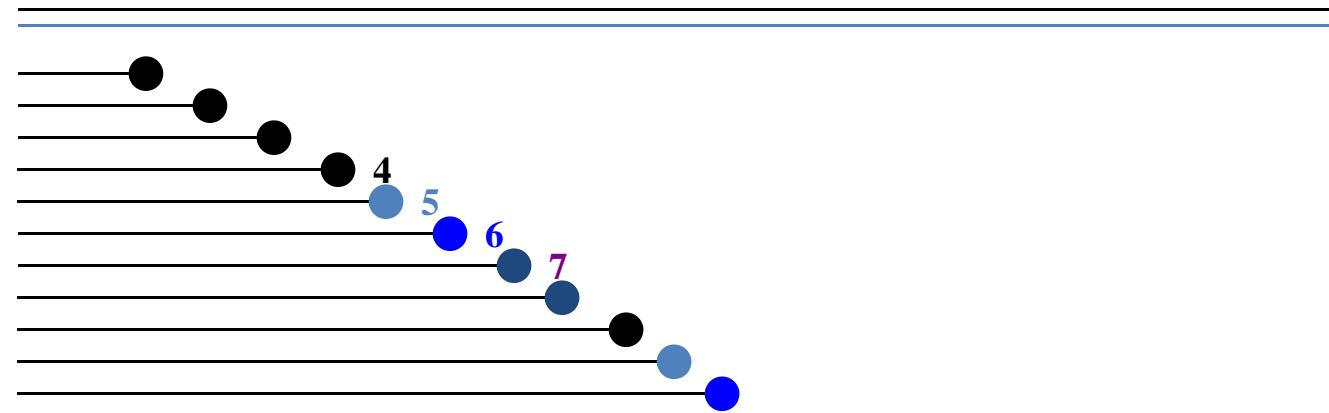
DNA Sequencer



DNA Sequencer MegaBACE 1000

96本キャビ°ラリータイ°の DNA Sequencerにより 1台あたり最大 **460,000b/day**のデータを測定します。

DNA Sequence



GGGATCCGATAATCGTCCCCTAGGCTATGGACA

4 CCCTAGGCTATTAGCAGGG●

5 CCCTAGGCTATTAGCAGGGG●

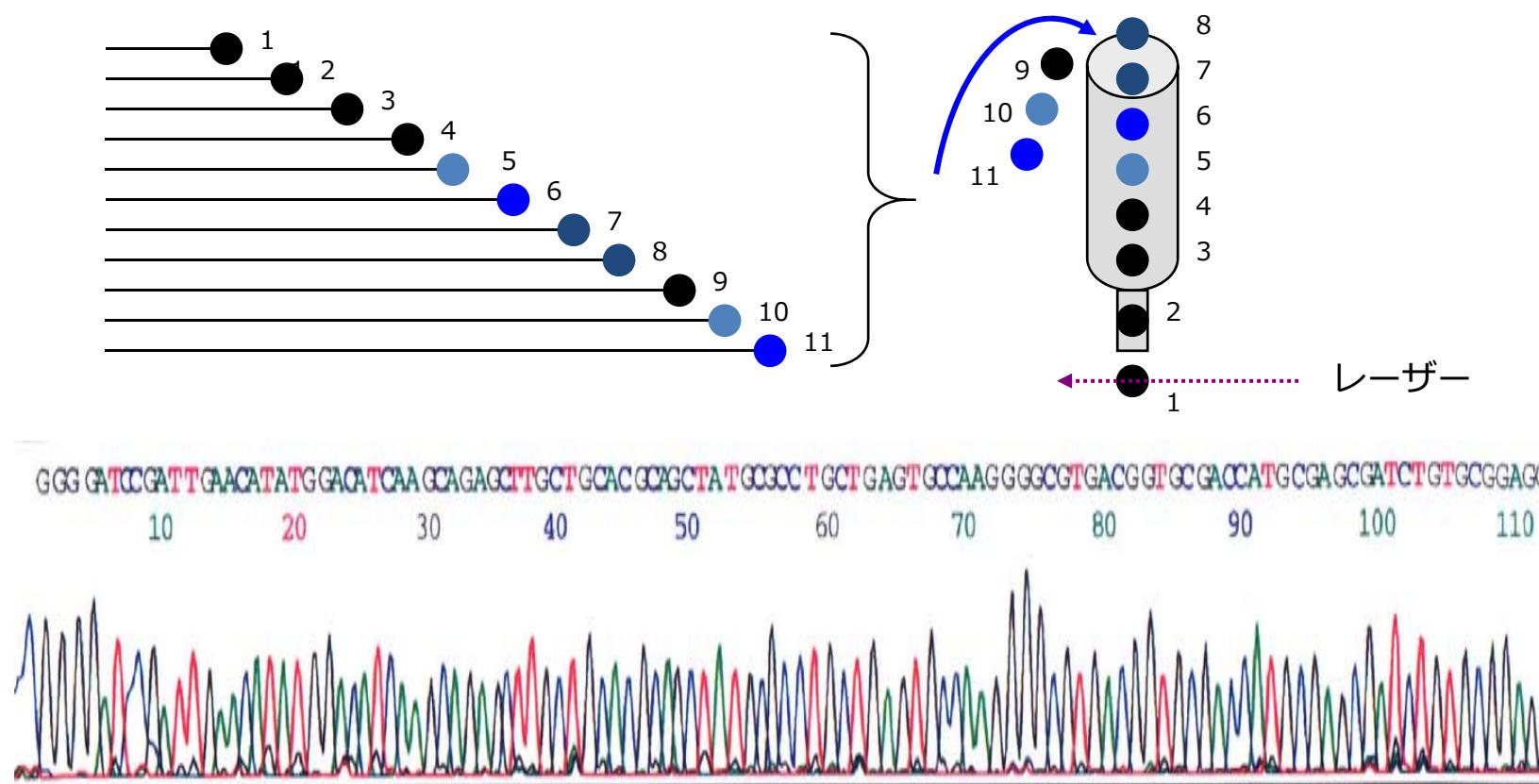
6 CCCTAGGCTATTAGCAGGGGA●

7 CCCTAGGCTATTAGCAGGGGAT●

DNA合成

プライマー

DNA Sequence



薬剤代謝酵素チトクロームP450の遺伝子多型解析 DNAチップ

- 「遺伝子多型解析キット AmpliChip CYP450」の販売
- ¥1,423,800 / 24テスト（税込み価格）
- 2006年5月9日



Diagnostics



遺伝子と病気について

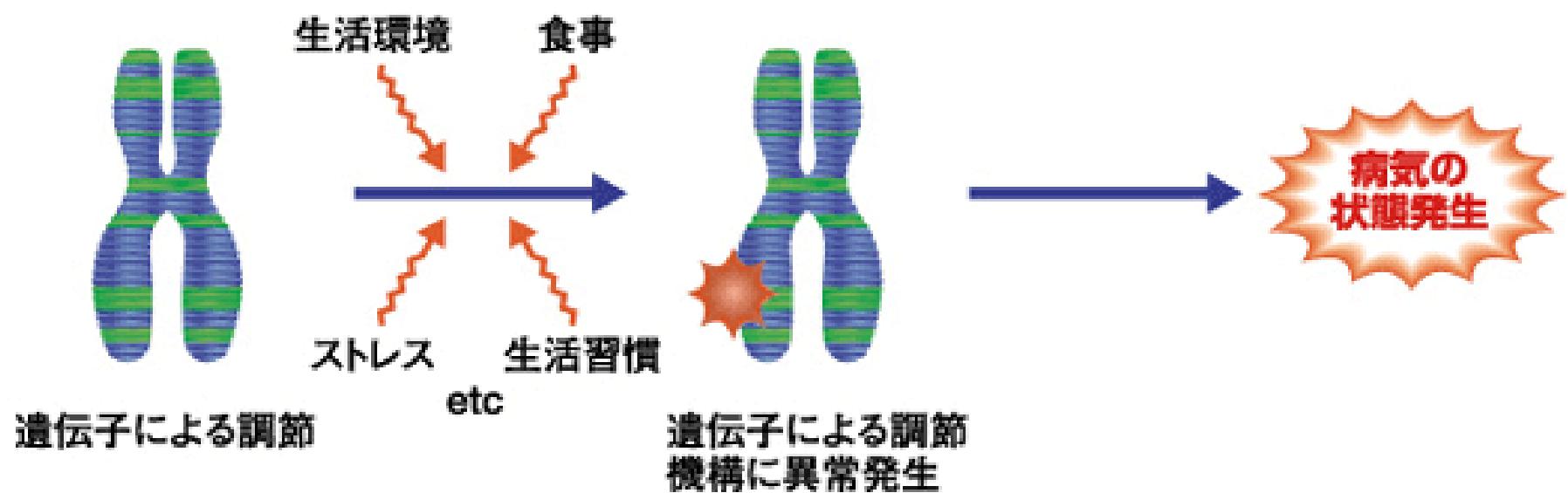
- ・ヒトに限らず生命活動においては、ゲノムに書き込まれた情報にそって、遺伝子がいつ、どこで、どのように働けばよいかが、非常に厳密にコントロールされている。このプログラムの運用によって、必要なタンパク質が必要な細胞でタイムリーに必要なだけ生成され、健康な生命活動が維持されている。



遺伝子と病気について

- しかし、食事や生活環境、ストレスなどの外的要因や内的要因によって、この調節機構に異常が生じると、生命活動の維持に必要な物質にアンバランスが生じ、その結果としてさまざまな病気を引き起こすことになる。

図5 遺伝子と病気について





遺伝子と病気について

- ゲノム情報の活用・研究によって、いつ、どこで、どのようなタンパク質がつくられるのか、また、その調節の仕組みが明らかになれば、それと対比させる形で、病気の際にはどこに、どのようなアンバランスが原因なのかを解明することができると考えられる。

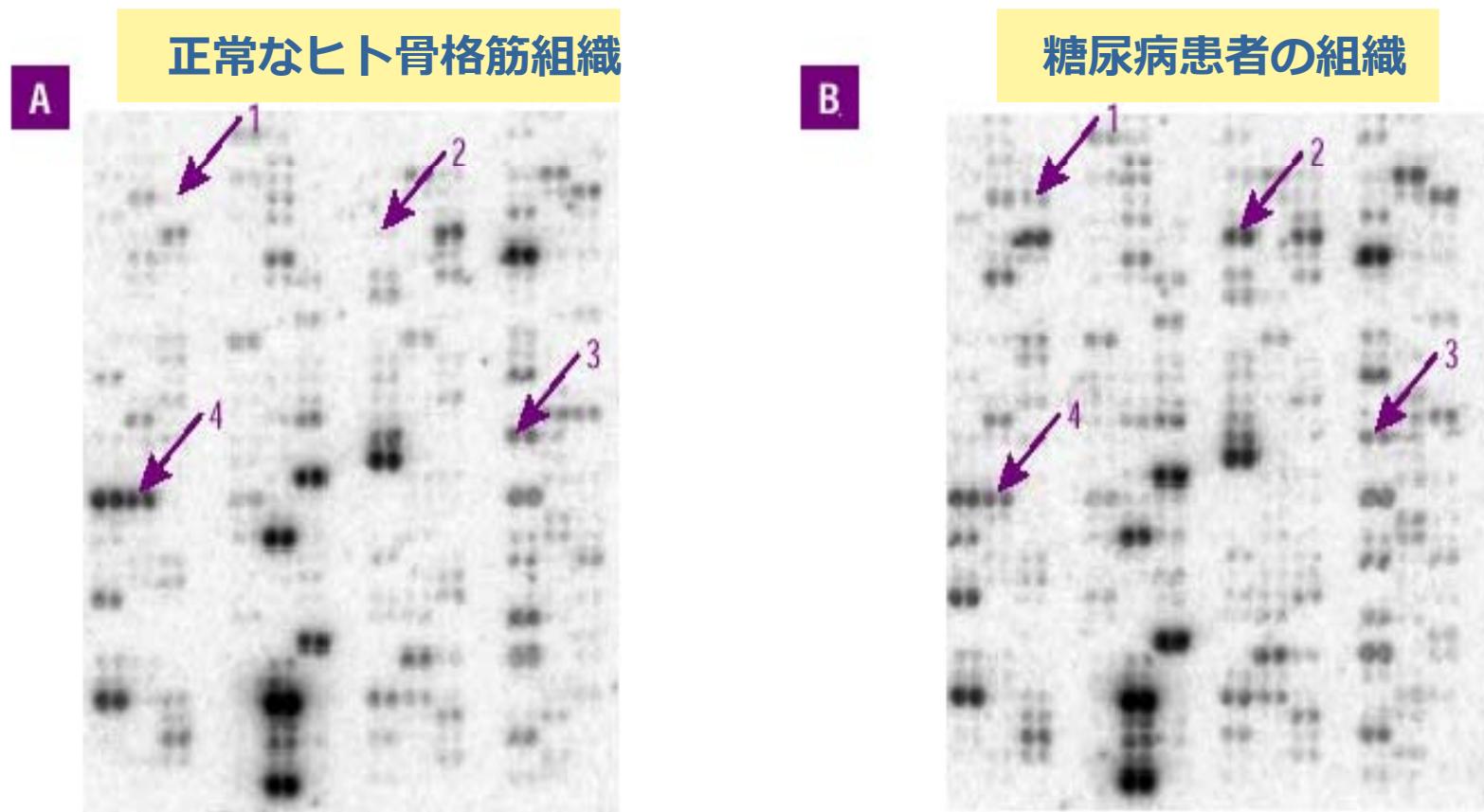
病気に対する危険因子と決定因子

- 病気に罹るケースにおいて関係する遺伝的な要因のウェイトは、病気の種類によって大きく異なる。
- 遺伝性疾患の場合は、遺伝子異常が病気になるかどうかの運命をほぼ決定しており、その遺伝子を持つ人は数十%から100%の確率で特定の病気になるものと予測される。このような要因を「**決定因子**」と呼んでいる。
- 一方、遺伝的変化（個人の体質を特徴づける遺伝子のバリエーション）を有する人が、ある病気にかかりやすい（普通の人の数倍高い確率で病気になる）場合には、その遺伝的な要因を「**危険因子**」と呼んでいる。

病気に対する危険因子と決定因子

- ・ 高血圧や糖尿病に代表される生活習慣病では、複数の遺伝的要因の積み重ねによって、かかる。要因といえれば、危険因子がいくつあるかによつて、かかりやすさが異なる傾向がある。要因といった病気は、多くの環境要因や遺伝要因（10～20種類程度）が危険因子として関わっていることから多因子疾患とも呼ばれている。
- ・ 遺伝子研究によって、病気の原因となる質行の使い方をうなづくことができる。個人治療する方法が最も適しているか、どういうくすりがベストなのかを理論的に選択することが出来るようになるかもしれません。

マイクロアレイを用いた発現 プロファイリング

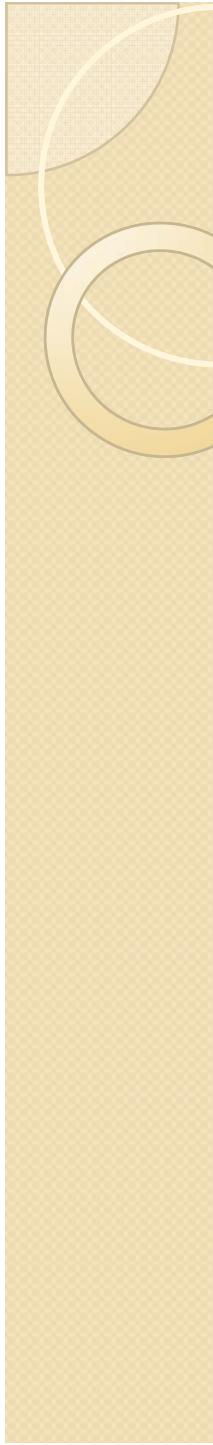


(AtlasTM Plastic Human 8.0 Microarray)



ゲノム研究に期待される健康への貢献

- 病気をおこす原因やメカニズムが、科学的根拠をもって遺伝子レベルおよびタンパク質レベルで解明されることが期待される。
- これによって、今まで原因がわかつていない病気の発症メカニズムが明らかになり、新たな治療法開発への道が切り開かれる可能性がある。



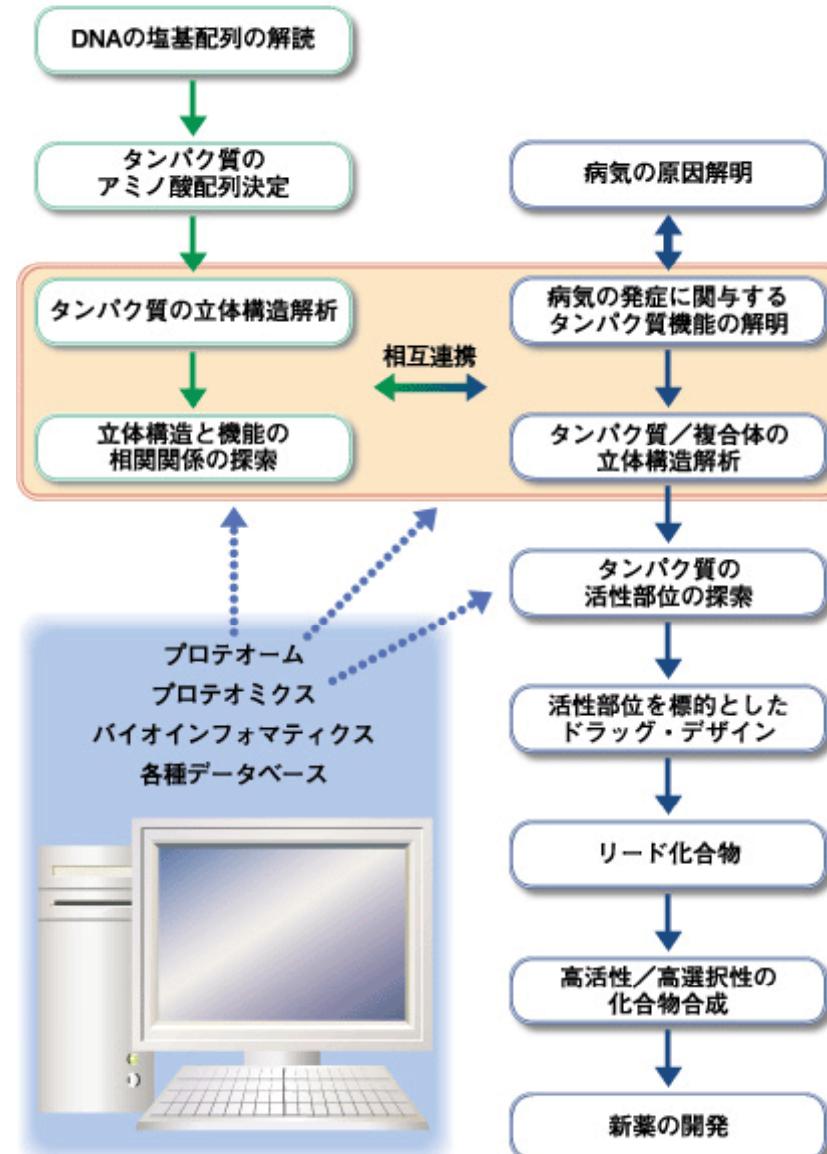
ゲノム研究に期待される健康への貢献

- 病気をおこす原因物質または原因となるメカニズムを標的とした、画期的な新薬がつくり出される可能性がある。従来、根本的な薬物治療が不可能であったり、その効果が不十分であった病気に有効な新薬の誕生が期待される。

ゲノム創薬の戦略



ゲノム情報を活用した創薬のまとめ



ゲノム研究

くすり開発