

平成21年度 市民講座 第7回

「ゲノムと情報学」

あなたの体質や生物の進化をゲノムから知る方法とは？

隈 啓一
(国立情報学研究所)



2009年は「ダーウィン・イヤー」でした。
(生誕200年、「種の起原」第1版発表150年)



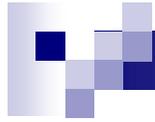
© Bioportal - Jabion

現在はゲノムの時代です。
ゲノムを調べることで、いろいろなことがわかります。
たとえば、「生物進化」・「人間の体質」などです。
特に後者は、社会に影響を与える可能性があります。



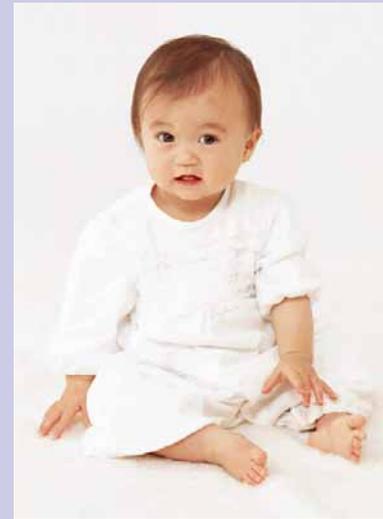
本日の内容

- (1) 遺伝子とゲノムとは？
- (2) 遺伝子やゲノムを生物の間で比較してわかること
- (3) 多型とは？
- (4) あなたの体質を決める1塩基多型



(1) 遺伝子とゲノムとは？

私たち人間は、両親と良く似た姿形や
性質をもって生まれます。



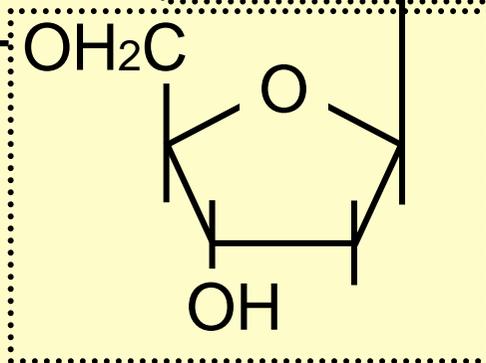
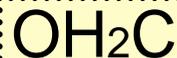
それは、親の遺伝子を受け継ぐからです。

遺伝子とは何でしょうか？

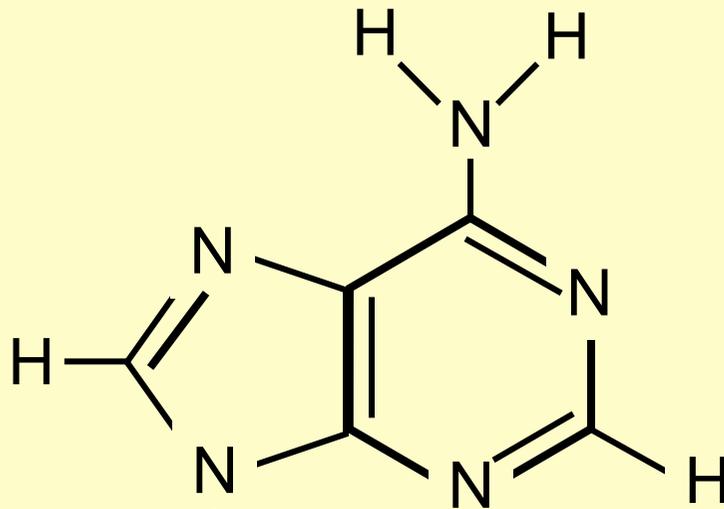
その本体は「DNA*」という物質です。

DNAの1つ
「アデニン」

リン



糖



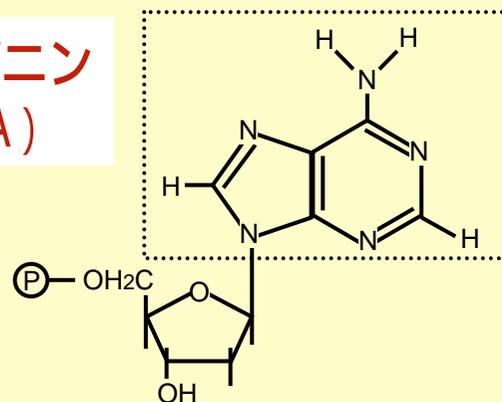
塩基

* DeoxyriboNucleic Acid
「デオキシリボ核酸」の略

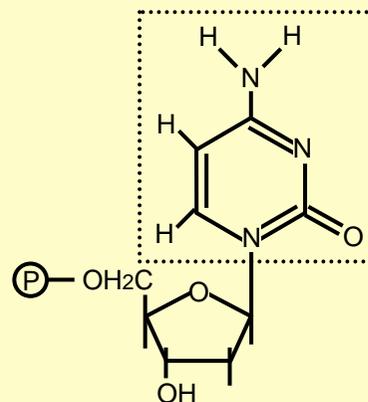
DNAには、互いに少し違った4種類があります。

これは「塩基」と言う部分の違いによります。
それぞれ、アデニン(A)、シトシン(C)、
グアニン(G)、チミン(T)になっています。

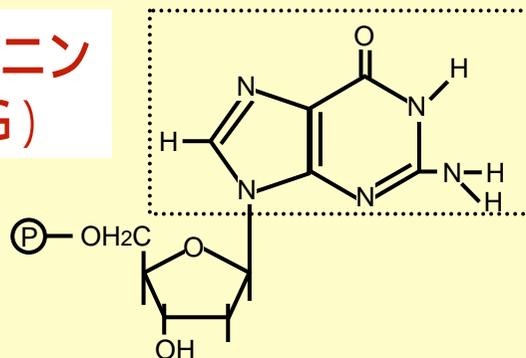
アデニン
(A)



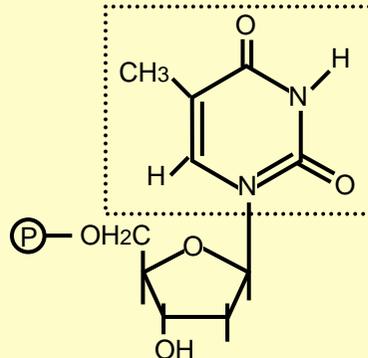
シトシン
(C)



グアニン
(G)



チミン
(T)

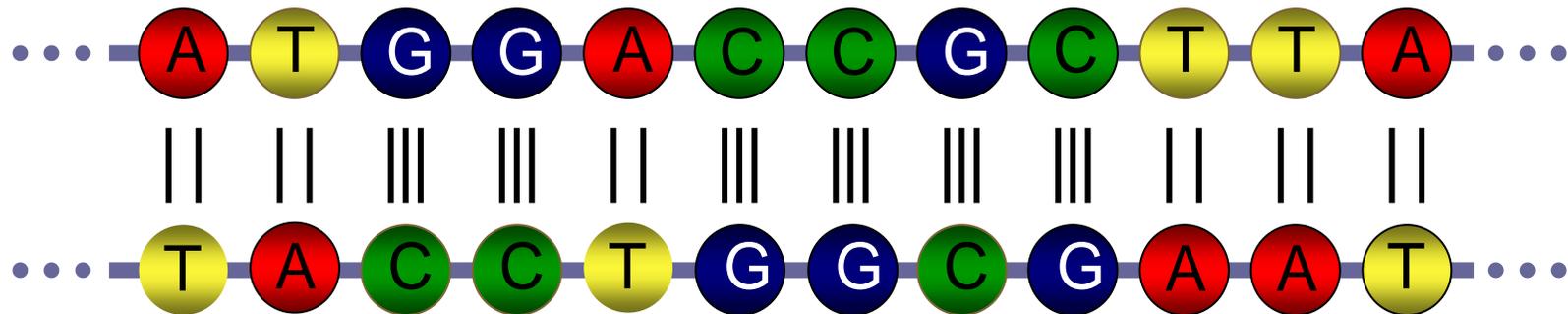




私たちの細胞の1つ1つの中には、この4種類のDNAが鎖のように長くつながったものがあり、「染色体」とよばれます。

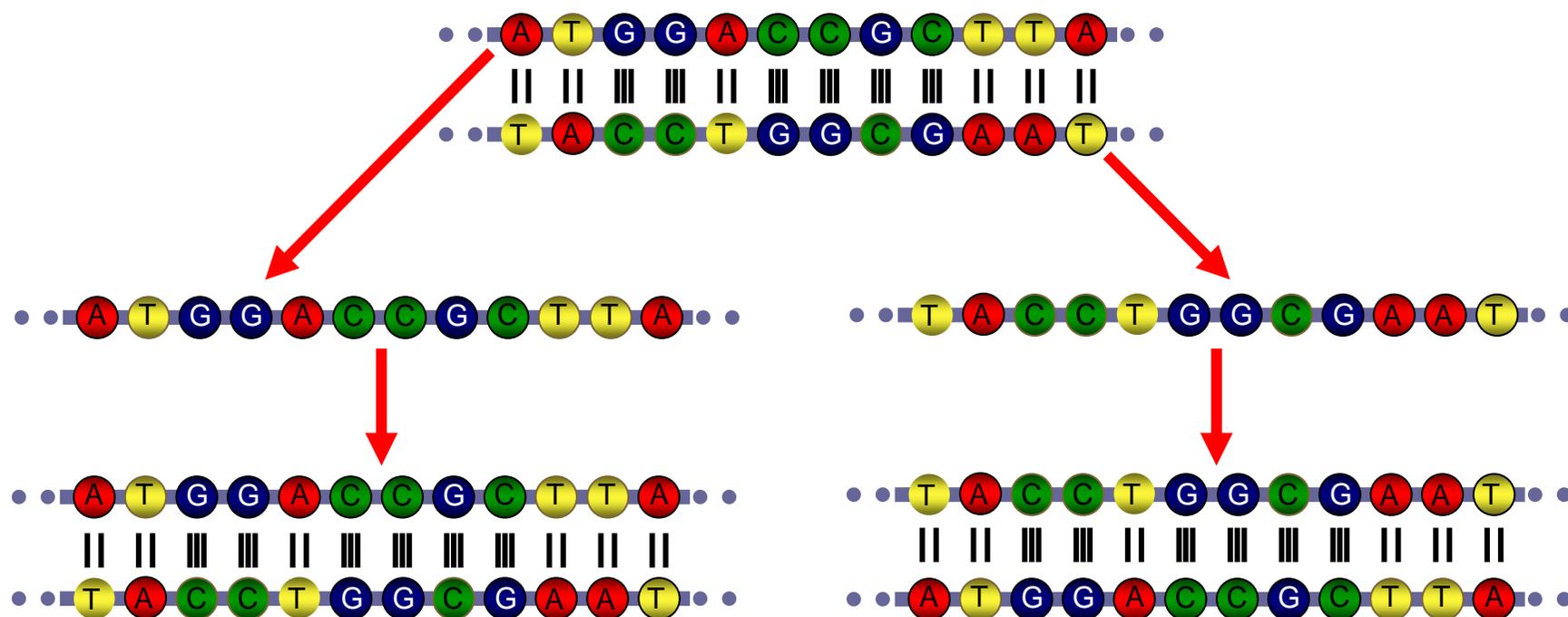


実際には、この「鎖」は「対(ペア)」を組む
2本鎖であり、お互いの鎖は、写真のネガ
(陰)とポジ(陽)のような対応関係にあります。



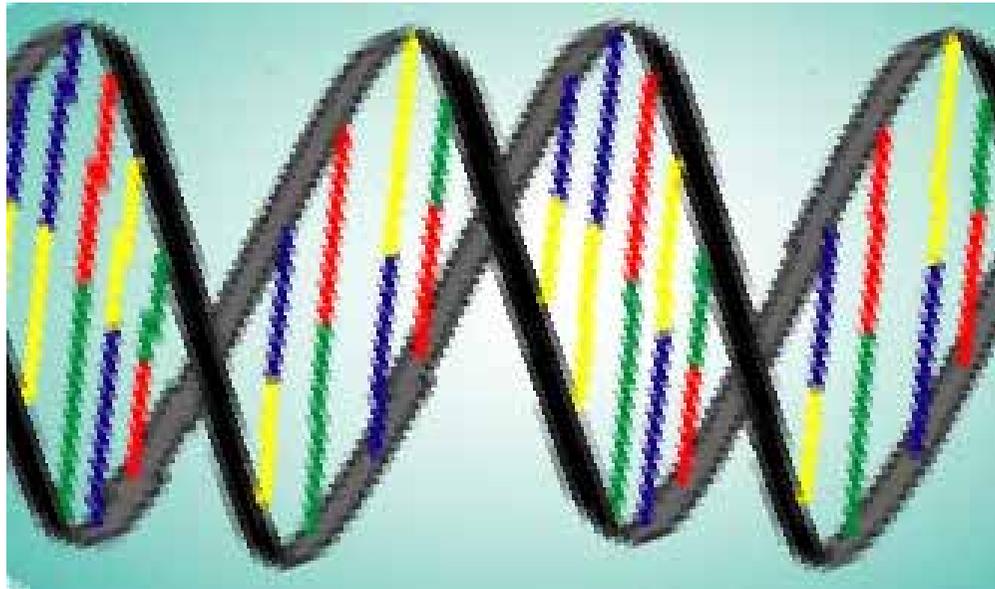
AにはT、CにはGが対応する。||と|||は水素結合を表す。

その理由は、2本鎖が2つに分かれると、自動的にそれぞれが2本鎖を作れるからです。



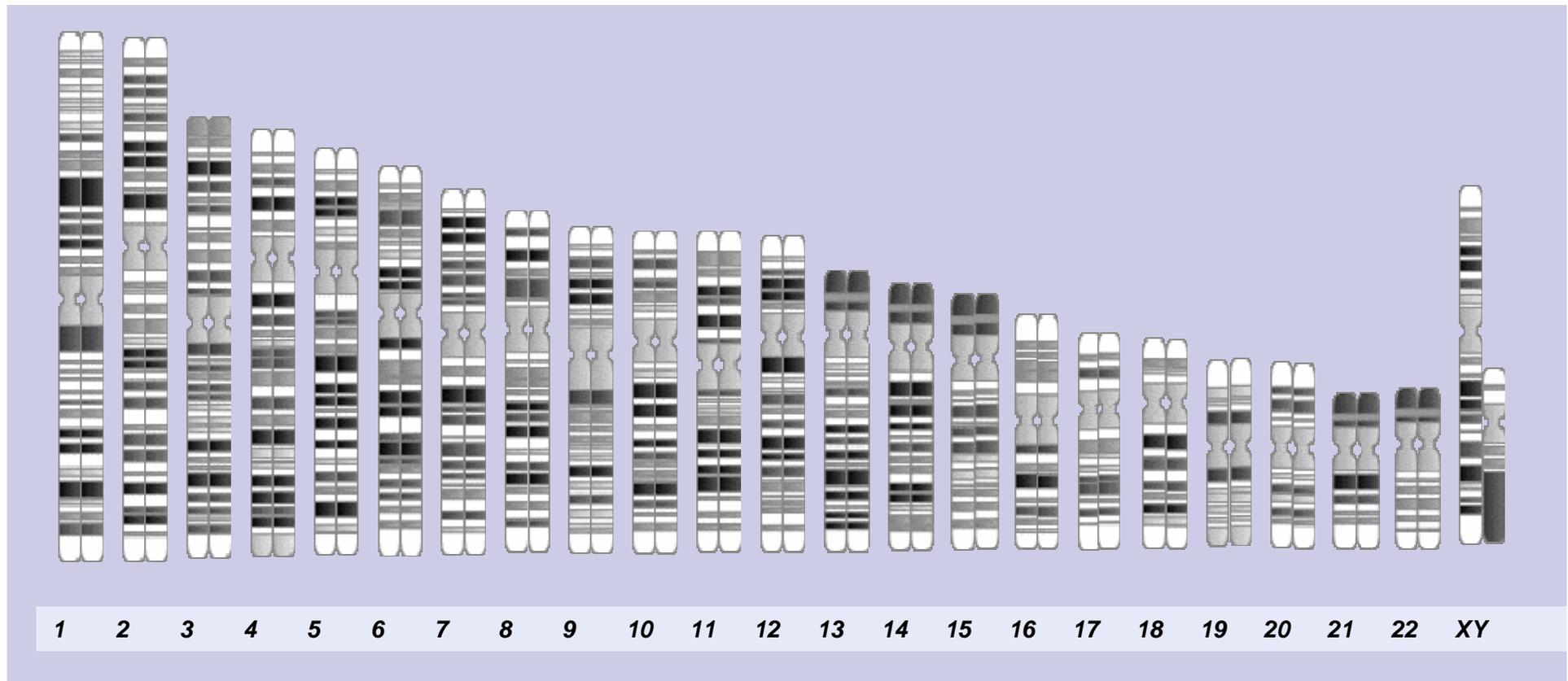
つまり、簡単に複製(コピー)を作れます。

この2本鎖がバネの形のような2重らせん構造をとることにより、染色体はつくられています。



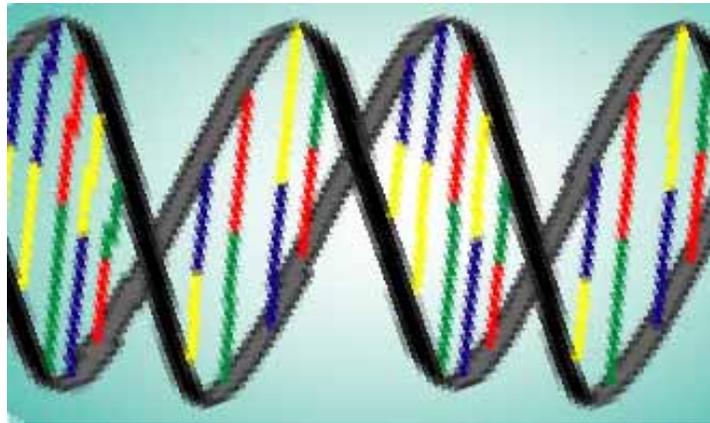
© Bioportal - Jabion

人間では染色体は23本(種類)ありますが、
父親由来と母親由来のものがあるので
 $23 \times 2 = 46$ 本存在します。



染色体DNAの総数は、約30億にも達します。

(ただし、片方の親由来のものだけで、かつ2本鎖の片方の鎖のみを数える)

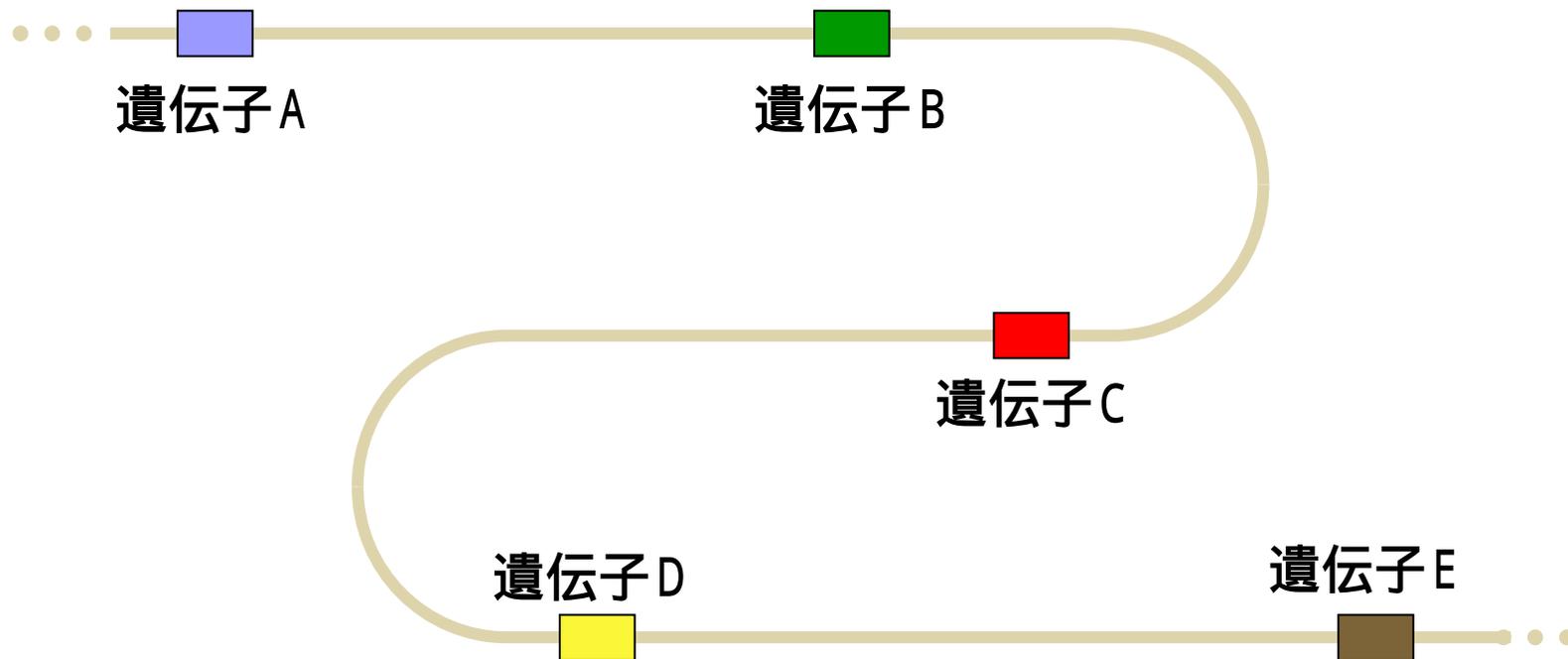


© Bioportal - Jabion

この約30億のDNAは、その人の遺伝情報のすべてであり、これを「ゲノム」といいます。

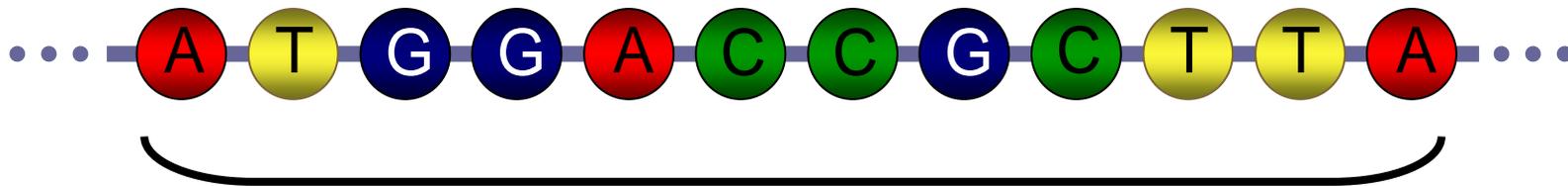
(正確には片親からの遺伝情報のすべて)

約30億のDNAには、約26,000個の
「遺伝子」がまばらに載っています。



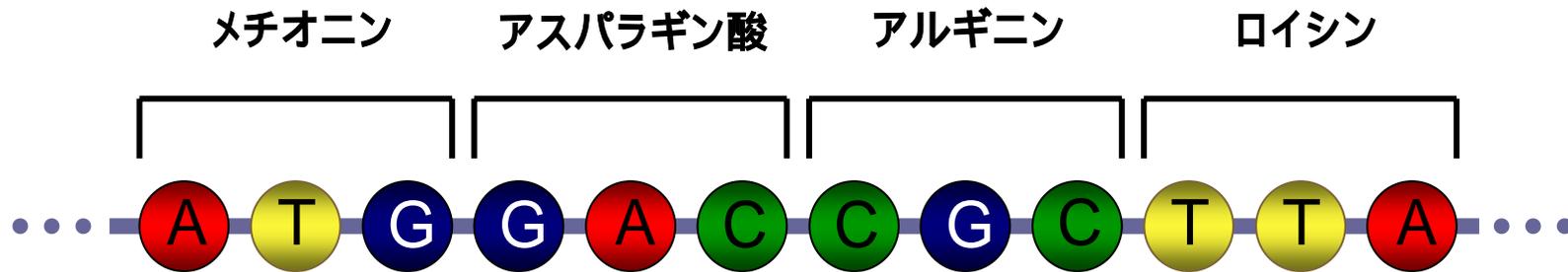
実際には、遺伝子の内部にも隙間(イントロン)があることが多い。

それぞれの「遺伝子」は、単なるDNAの並びですが、「タンパク質」を作るための設計図が「暗号」で記されています。

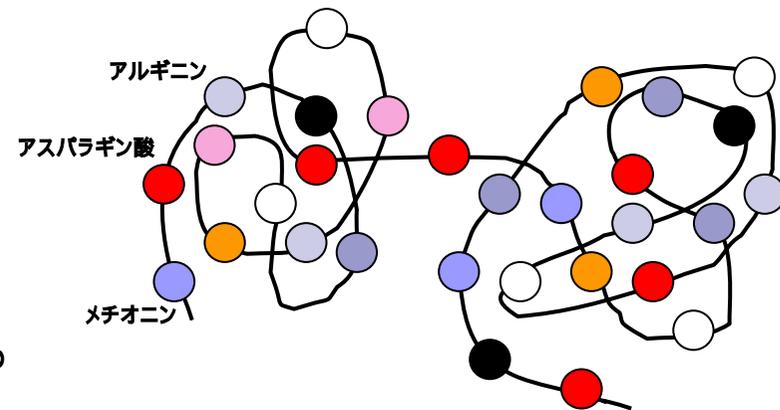


遺伝子の一部

3つのDNAの並び(コドン)が、
1つの「アミノ酸」を指定しています。
アミノ酸がつながったものがタンパク質です。



タンパク質は針金細工のように
折れ曲がり、それぞれの用途に
合った立体的な構造物を作ります。



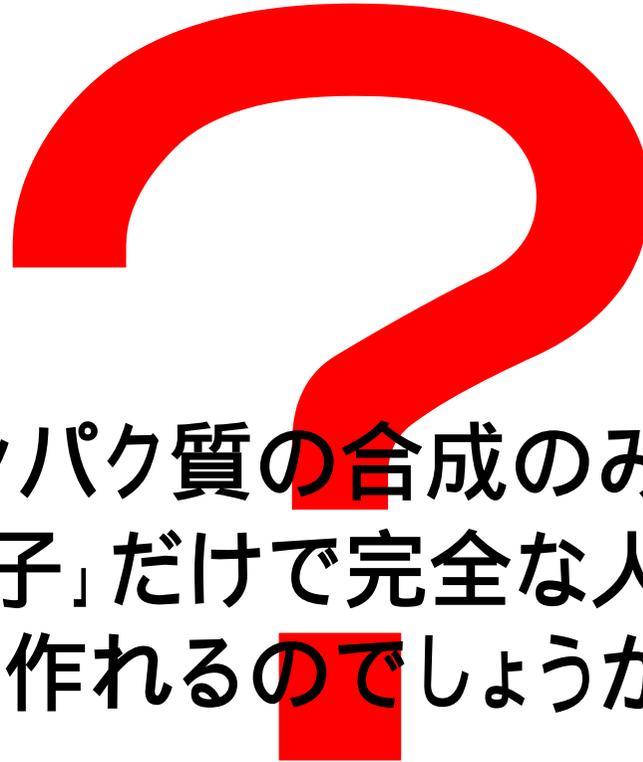
3つのDNAの並びは全部で
64(4×4×4)通りですが、これらが
20種類のアミノ酸と「終わり」を指定しています。

TTT = フェニルアラニン	TCT = セリン	TAT = チロシン	TGT = システイン
TTC = フェニルアラニン	TCC = セリン	TAC = チロシン	TGC = システイン
TTA = ロイシン	TCA = セリン	TAA = 終わり	TGA = 終わり
TTG = ロイシン	TCG = セリン	TAG = 終わり	TGG = トリプトファン
CTT = ロイシン	CCT = プロリン	CAT = ヒスチジン	CGT = アルギニン
CTC = ロイシン	CCC = プロリン	CAC = ヒスチジン	CGC = アルギニン
CTA = ロイシン	CCA = プロリン	CAA = グルタミン	CGA = アルギニン
CTG = ロイシン	CCG = プロリン	CAG = グルタミン	CGG = アルギニン
ATT = イソロイシン	ACT = スレオニン	AAT = アスパラギン	AGT = セリン
ATC = イソロイシン	ACC = スレオニン	AAC = アスパラギン	AGC = セリン
ATA = イソロイシン	ACA = スレオニン	AAA = リジン	AGA = アルギニン
ATG = メチオニン	ACG = スレオニン	AAG = リジン	AGG = アルギニン
GTT = バリン	GCT = アラニン	GAT = アスパラギン酸	GGT = グリシン
GTC = バリン	GCC = アラニン	GAC = アスパラギン酸	GGC = グリシン
GTA = バリン	GCA = アラニン	GAA = グルタミン酸	GGA = グリシン
GTG = バリン	GCG = アラニン	GAG = グルタミン酸	GGG = グリシン

「遺伝暗号」表

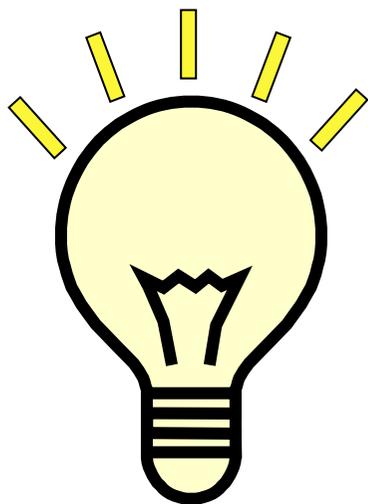


しかし、人間の体はタンパク質だけから
できているわけではありません。



なぜ、タンパク質の合成のみ*を指示する
「遺伝子」だけで完全な人間の体を
作れるのでしょうか？

*正確には、「RNA」の合成も指示する。



それはタンパク質が「万能選手」
だからです。

もちろんタンパク質は、筋肉・臓器(アクチン・ミオシン)、毛髪や爪(ケラチン)、眼の水晶体(クリスタリン)、腱・軟骨(コラーゲン)などの成分です。

しかし、他にも実にさまざまなことに使われます。

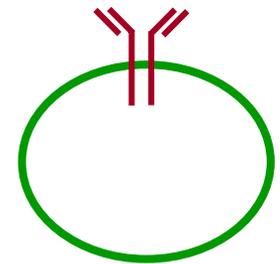
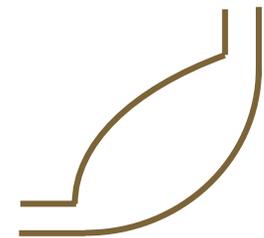
食べ物を消化するための「消化酵素」になります。
トリプシン・リパーゼなど

いろいろな「ホルモン」とその「受容体」をつくります。
成長ホルモン・エストロゲン受容体など

外来の異物に対する「抗体」をつくります。
免疫グロブリン

タンパク質以外の物質を「輸送」します。
ヘモグロビン・メタロチオネインなど

その他、いろいろな化学反応を触媒したり、
細胞内装置の部品になったりします。

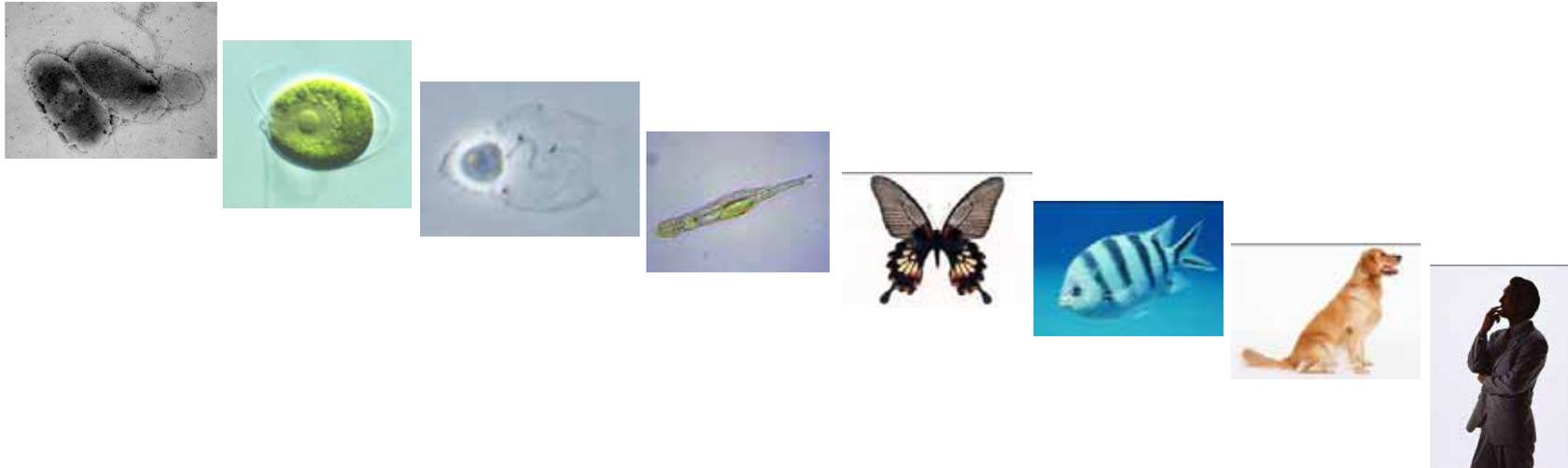




**(2) 遺伝子やゲノムを生物の
間で比較してわかること**

ゲノムをもつものは人間だけではありません。
生物は、すべてゲノムをもっています*。

それは、大昔に存在した、すべての生物の
祖先生物から受け継いだものです。



* ウイルスは生物ではないという考えが一般的だが、ゲノムをもつ。

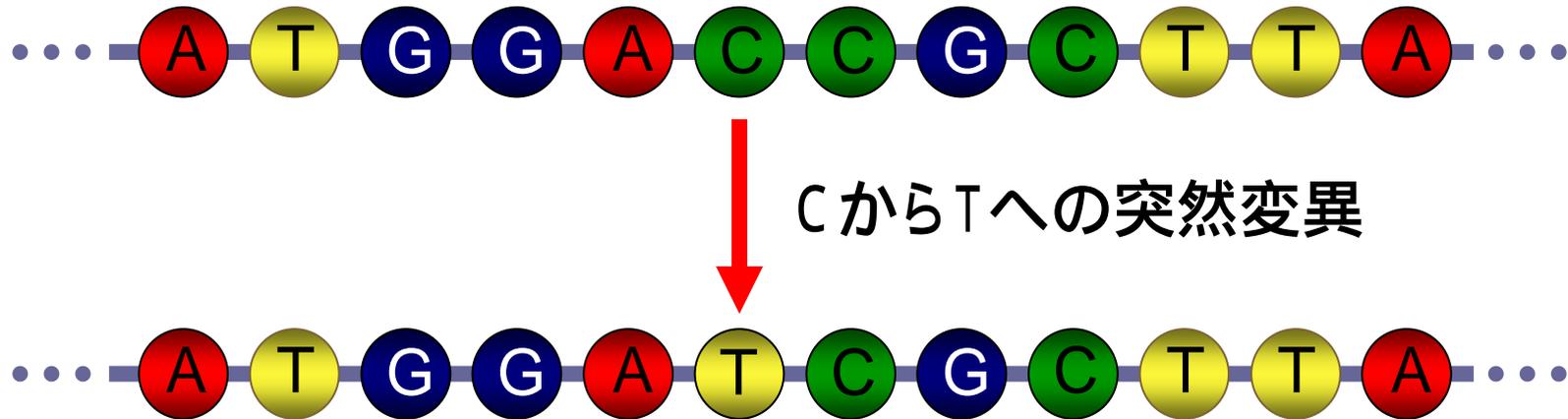
しかし、ゲノムや、その一部である遺伝子をくらべると、生物ごとに少しずつ違ってきます。

ヒト		...GTCTAAGTACG...
イヌ		...GTCTAAGTACG...
サカナ		...GTCTAAGTACA...
チョウ		...GTCTCAGTGCA...

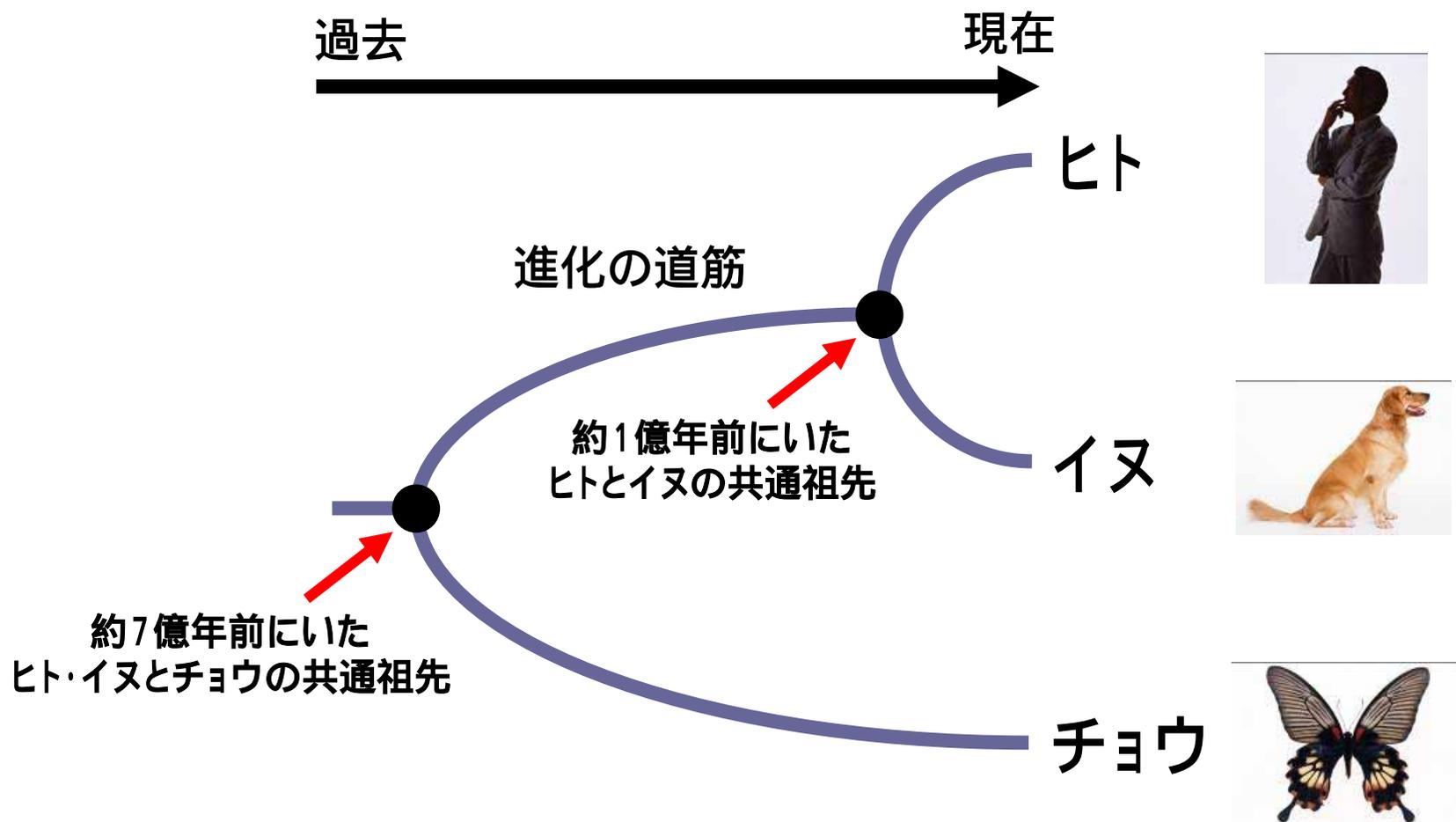
18SリボソームRNA遺伝子の配列の一部

ヒトに「近い」生物ほどヒトに遺伝子が似ています。これはなぜでしょうか？

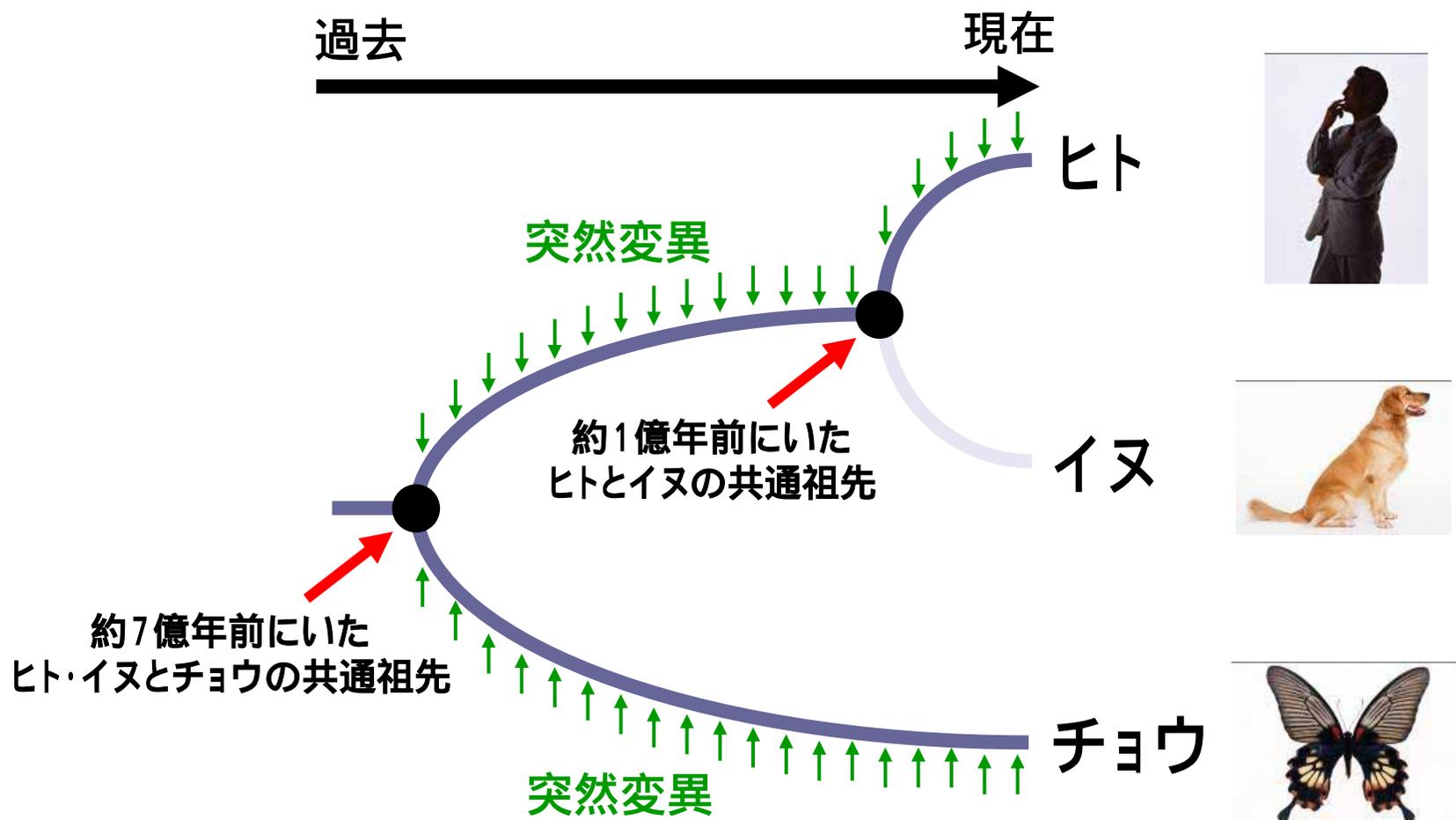
それは、DNAが複製される際に、ある確率で間違い(突然変異)がおこり、それが子孫に伝えられることがあるからです。



ヒトとイヌの共通祖先は約1億年前の生物で、
ヒト・イヌとチョウの共通祖先は約7億年前の生物です。



ヒトとチョウは約7億年の間にそのゲノムにそれぞれ、このような突然変異を受けてきました。





以上の理由で、チョウよりもイヌの方が
ヒトのゲノムに似ているのです。

一般に、2つの生物が別種として分かれた
時点が古ければ古いほど、両者のゲノムは
異なっていきます。これは、一定の確率で
突然変異がおこるため、時間におおよそ
比例して、突然変異が溜まっていくからです。

この原理(分子時計)を利用すれば、
次のような2つの研究が可能です。

(1) ある生物が他のどの生物に近いがわかる

近縁の生物同士であるほどゲノム(遺伝子)が似る



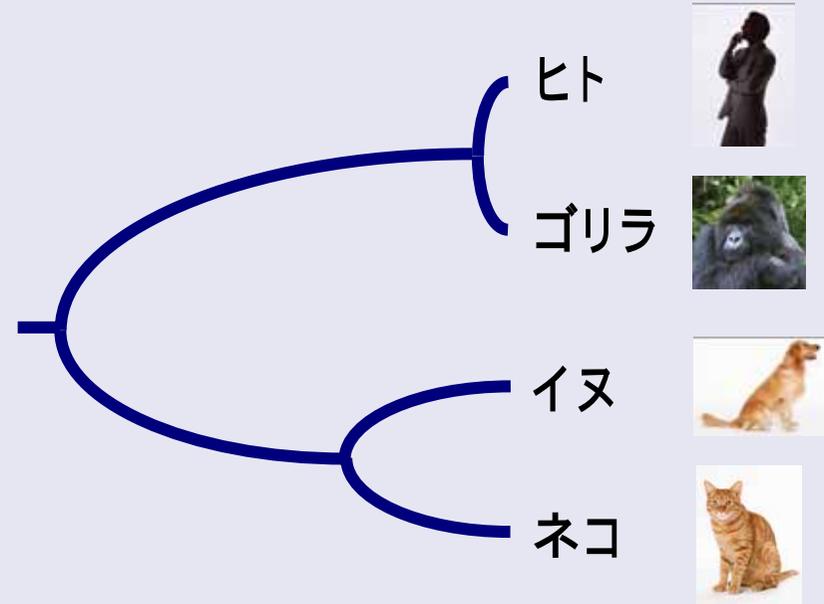
遺伝子が似ている生物同士であるほど、より近縁



遺伝子の似ている度合いを数値化して、いろいろな生物で比較すれば、それらの生物の類縁(親戚)関係がわかる(図解できる = 分子系統樹)!

たとえば、ある遺伝子をヒト、ゴリラ、イヌ、ネコで比較して、違っているDNAの数を数えたとします。

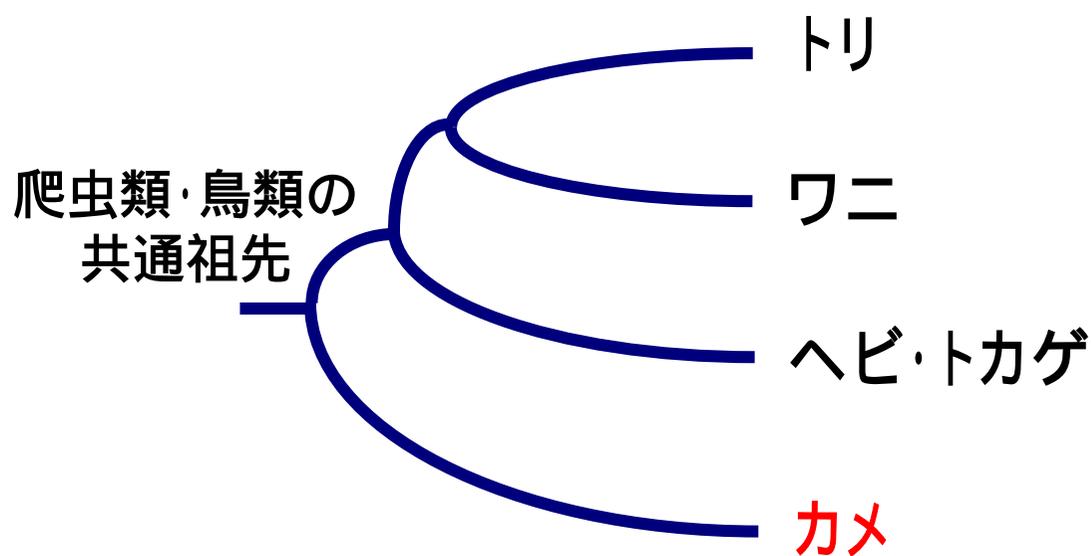
	ヒト	ゴリラ	イヌ	ネコ
ヒト	0	1	10	10
ゴリラ		0	10	10
イヌ			0	4
ネコ				0



その結果が、左の表の通りだとすると、右のような類縁(親戚)関係が推定されます(ヒトとゴリラが近縁で、イヌとネコも近縁)。

研究例：「カメは古いタイプの爬虫類ではない！」

カメは甲羅など、他の爬虫類とは大きく異なる形をしており、また頭蓋骨の穴(側頭窓)がありません。よって、カメは最も古いタイプの爬虫類だと思われてきました(下図)。

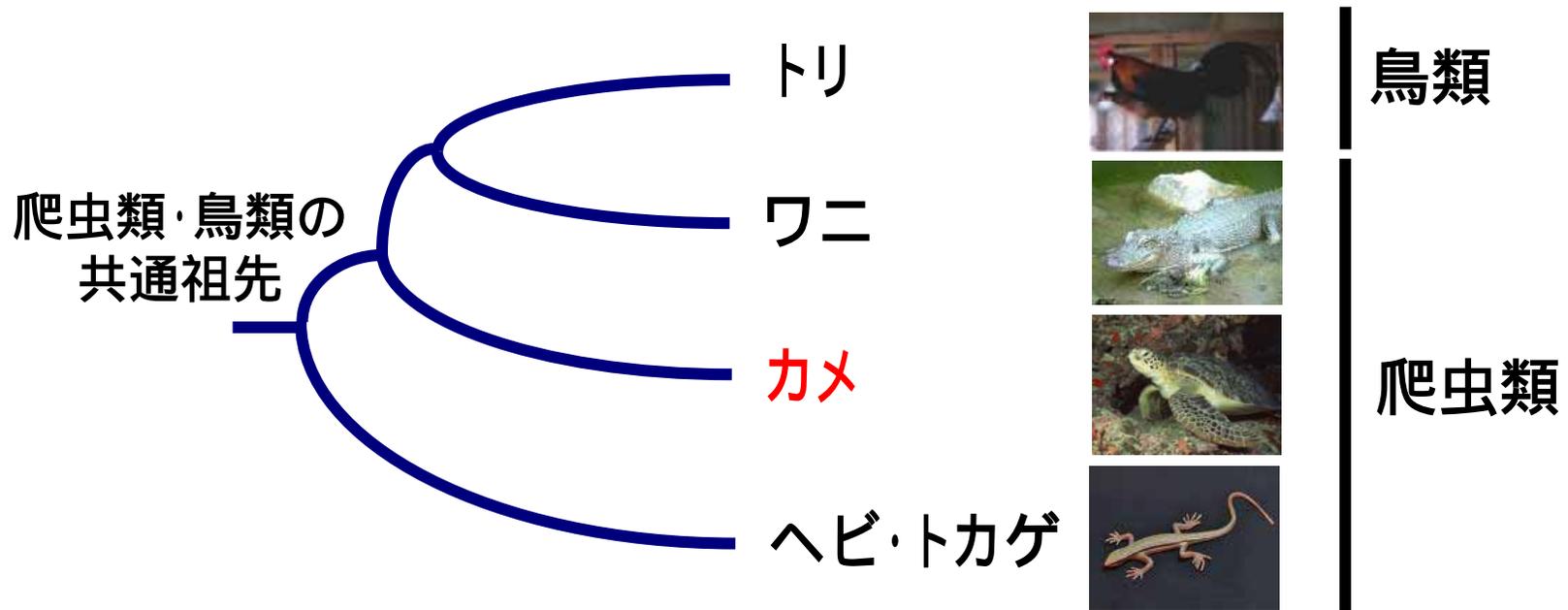


鳥類

爬虫類

しかし、Cao Y et al. (2000)、Iwabe et al. (2005)などの、
遺伝子を比較した研究によれば、最も古いタイプの
爬虫類はヘビ・トカゲ類らしいのです。

現在では、このことは広く認められつつあります。



このような研究は、化石があまり残っていない生物や、外見にあまり特徴がなく、分類しにくい生物(細菌などの微生物)の間の類縁関係を探るときに特に有効です。



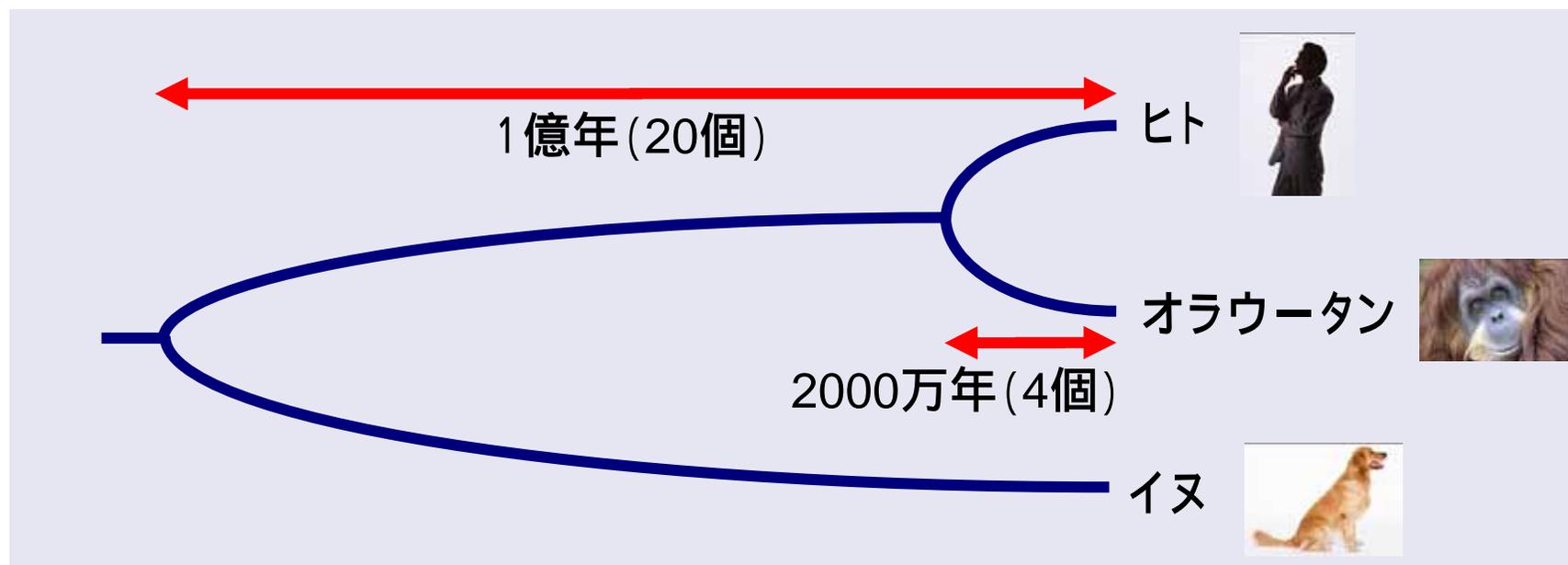
(2) ある生物と他の生物がいつごろ 別種として分かれたかがわかる

突然変異が時間にほぼ比例して蓄積するならば、
DNAが異なる程度は分かれてからの時間に比例する



ある生物と他の生物が、いつ別種として分かれたかが
わかっていれば、それをものさしにして、いろいろな
生物がいつ頃別種として分かれたかがわかる！

たとえば、ヒトとイヌが約1億年前に別種として分かれたとします。また、ある遺伝子をヒトとイヌで比較すると、違っているDNAは20個だったとします。



もし、ヒトとオラウータンで同じ遺伝子を比較したとき、違っているDNAは4個だったならば、ヒトとオラウータンは約2000万年前に別種として分かれたと考えることができます。

研究例：「現生人類と、現生人類に一番近いネアンデルタール人は、約37万年前に種として完全に分かれた」



Science (2006) 314, 1071

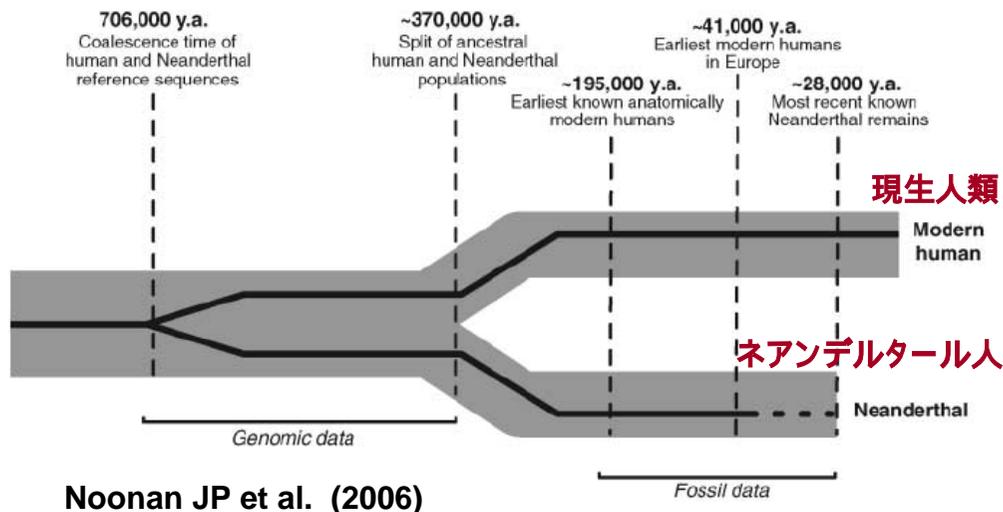


Cell (1997) 91, 1

ネアンデルタール人は、今から約2～3万年前まで存在していました。彼らと私たちがどのくらい近縁なのか、また遺伝的交流(結婚)があったのかは非常に関心のあることでした。

38,000年前のネアンデルタール人の化石*からDNAを取り出してその一部を(65,250bp)を決定、現生人類およびチンパンジーと比較しました。

ヒトとチンパンジーが分かれた時期(約650万年前)をものさしにして考えると、現生人類とネアンデルタール人は、約(71~)37万年前に分かれたと考えられました。



彼らは私たちとは、それほど近縁ではなく、遺伝的交流もなかったようです。

* 化石の状態にもよるが、一般に化石からとれたDNAの配列を決められるのは、10万年前くらいが限界。

また、ゲノム全体を比較することでゲノム(遺伝子)の
どんな違いが、生物間の形や性質の違いを
決めているかを探る研究も盛んです。

多細胞生物と単細胞生物の違い
陸生生物と水棲生物の違い
放射線に対する耐性の違い

生物のゲノムは、現在1,000以上も決定しており、
ますますその数は増加する見込みです。

脊椎動物10,000種ゲノム
計画なども提案されています。



ゲノムが現在決定している生物

真核生物 (219)	動物 (57)	脊椎動物	21
		頭索動物	1
		尾索動物	3
		キョクヒ動物	2
		節足動物	21
		線形動物	3
		扁形動物	3
		刺胞動物	2
		板状動物	1
	植物		90
	菌類		43
	原生生物		29
原核生物 (1020)	真正細菌		951
	古細菌		69
合計			1239

ゲノムサイズ(遺伝子数)は生物によって大きく異なります

生物	ゲノムサイズ(遺伝子数)
カルソネラ(真正細菌)	約16万(182個)
大腸菌(真正細菌)	約460万(約4,100個)
コウボ菌(菌類)	約1200万(約5,900個)
ショウジョウバエ	約1億8000万(約14,000個)
トラフグ	約3億5000万(約22,000個)
ニワトリ	約12億(約18,000個)
トウモロコシ	約23億(約32,000個)
人間	約30億(約26,000個)
サンショウウオ	約400億(?)
肺魚(エチオピクス)	約1300億(?)
アメーバ(<i>Amoeba dubia</i>)	約6700億(?)



真核生物ゲノムは「ヘテロクロマチン領域」を除いた大きさの場合もある

© Bioportal - Jabion

参考:最大のゲノムをもつウイルス:ミミウイルス 約120万(911個)

最近では、ある環境にいる生物(主に微生物)全てのゲノムを一度に決める、「メタゲノム計画」も盛んです。

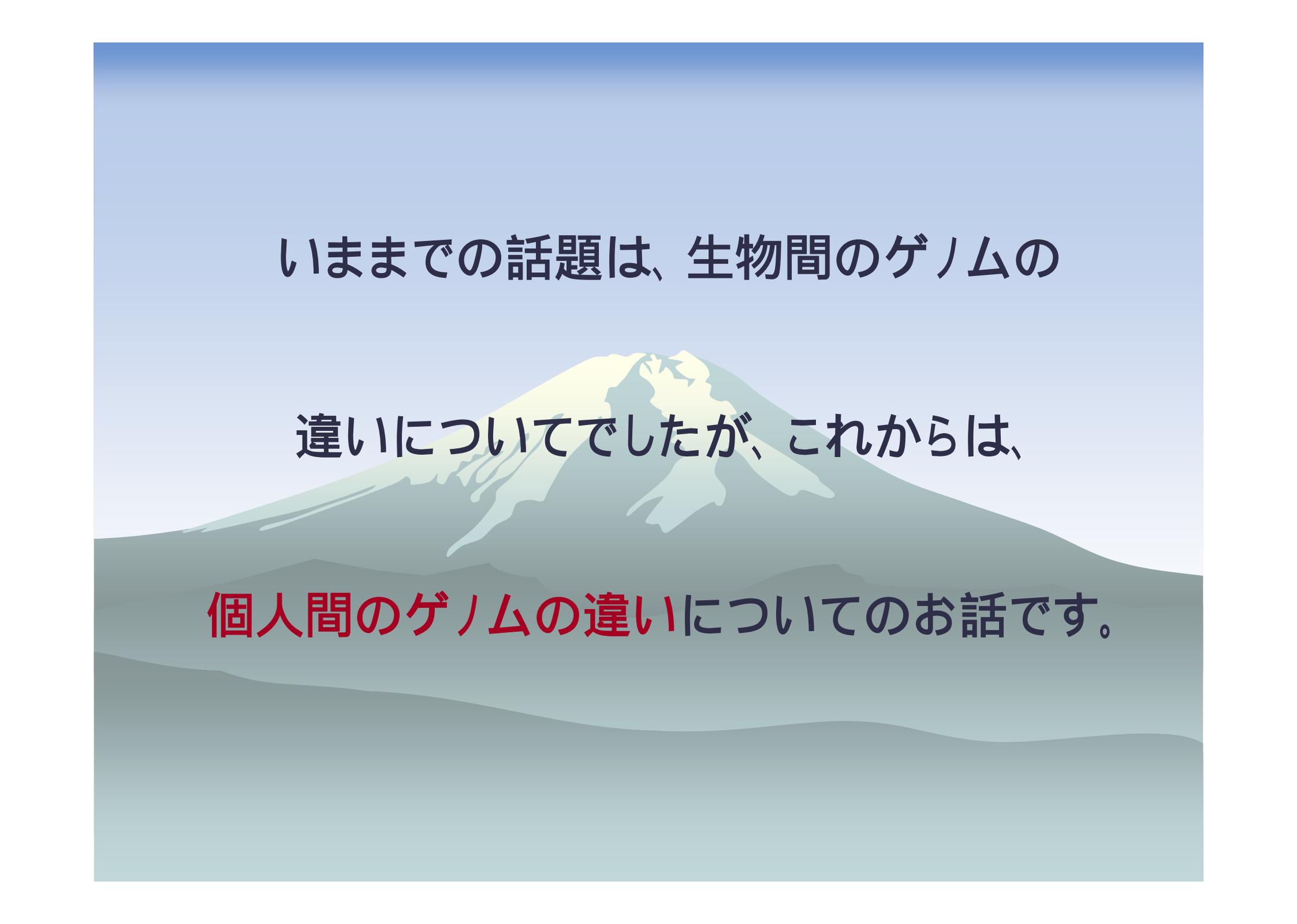
メタゲノム解析の例(予定も含む)

- 鉦山廃水に存在する細菌群
- サルガッソーの海洋細菌群
- マンハッタンの大気中に存在する細菌群
- 海底の鯨骨細菌群と農場土壌細菌群
- ヒト小腸常在細菌群
- シロアリ共生細菌群
- ヒト腸内細菌群
- 超好塩性細菌集団
- 海岸の砂浜にいる生物死体からのDNA
- 腸管のない線虫からの共生細菌集団
- メタンを酸化する深海古細菌集団
- 人間の身体各部の細菌群

環境と生物の関係が研究できる
生物と生物の相互作用の研究ができる
培養できないため未知だった微生物
(99%にも達する)のゲノムの決定が可能
有用な働きをする微生物の発見



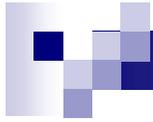
© Bioportal - Jabion



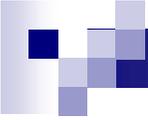
いままでの話題は、生物間のゲノムの

違いについてでしたが、これからは、

個人間のゲノムの違いについてのお話です。



(3) 多型 (たけい) とは？

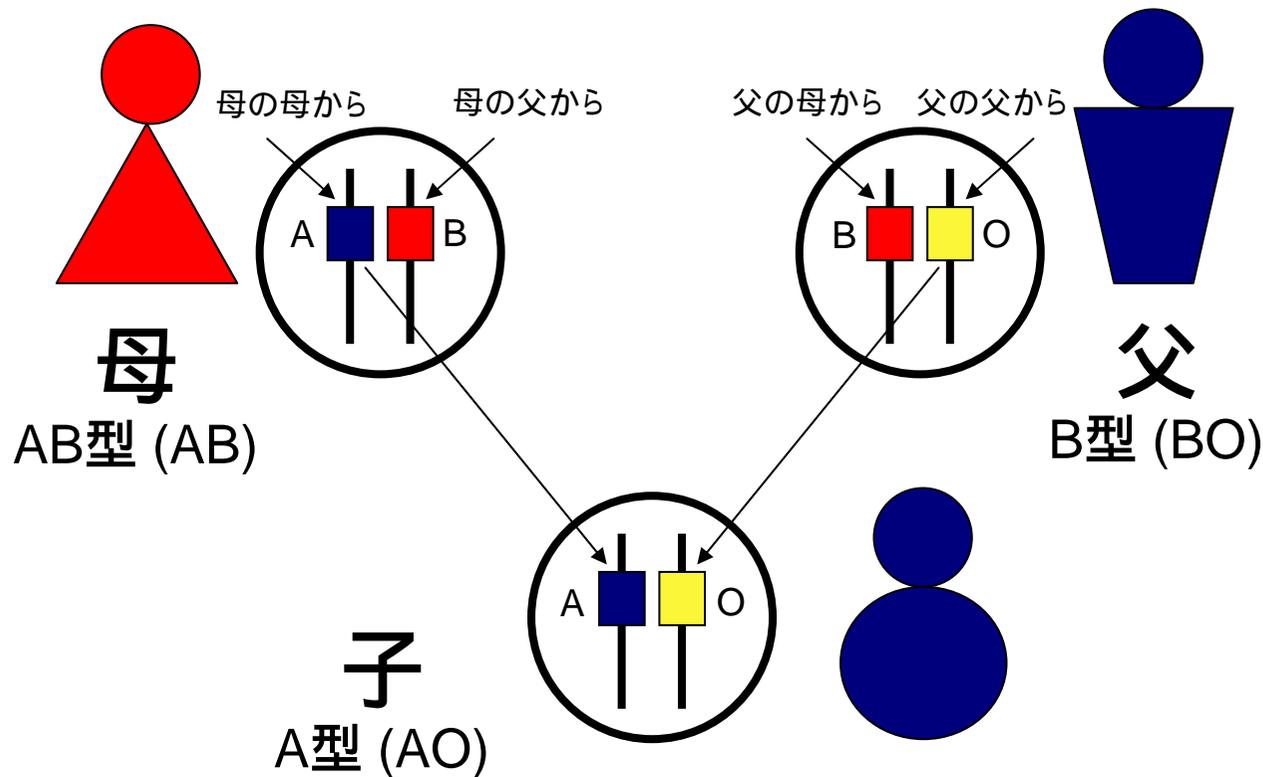


「多型」とは、個人のゲノム間で違いが見られることをいいます。なかでも、比較的知られているものが、「遺伝子多型」です。

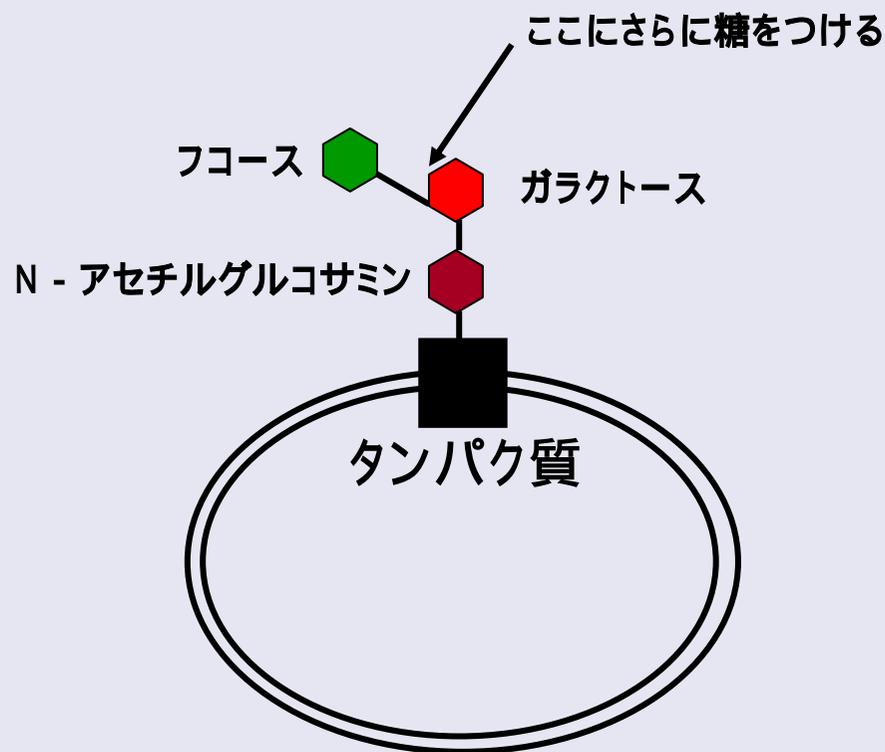
その最も代表的なものが「血液型」です。

A **B** **O** **AB**

人間は、母と父由来の2つのゲノムを持つため、ほとんどの遺伝子を2個もちますが、それらは互いにわずかに異なることもあります。血液型の遺伝子の場合、3種類あります(A, B, O)。



血液型 (A B O 式) を決める遺伝子は、
「糖転移酵素」で、細胞の表面にある
タンパク質の「糖鎖」にさらに糖を付け加えます。



A遺伝子:



N - アセチルガラクトサミンをつける

B遺伝子:



ガラクトースをつける

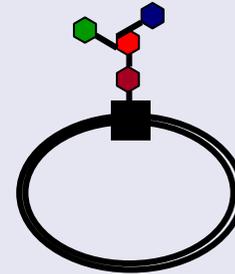
O遺伝子:

何もつけない

血液型と遺伝子型の関係

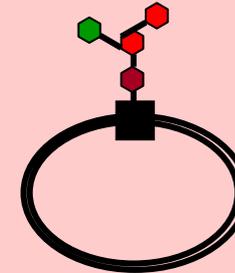
A型:

両親からA遺伝子とO遺伝子(AO)、
あるいはA遺伝子を2つ(AA)受け継ぎ、
A型の糖鎖をもつ



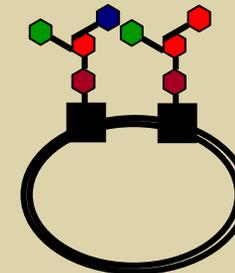
B型:

両親からB遺伝子とO遺伝子(BO)、
あるいはB遺伝子を2つ(BB)受け継ぎ、
B型の糖鎖をもつ



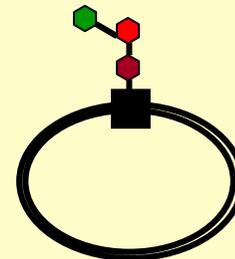
AB型:

両親からA遺伝子とB遺伝子(AB)を受け継ぎ、
A型・B型両方の糖鎖をもつ

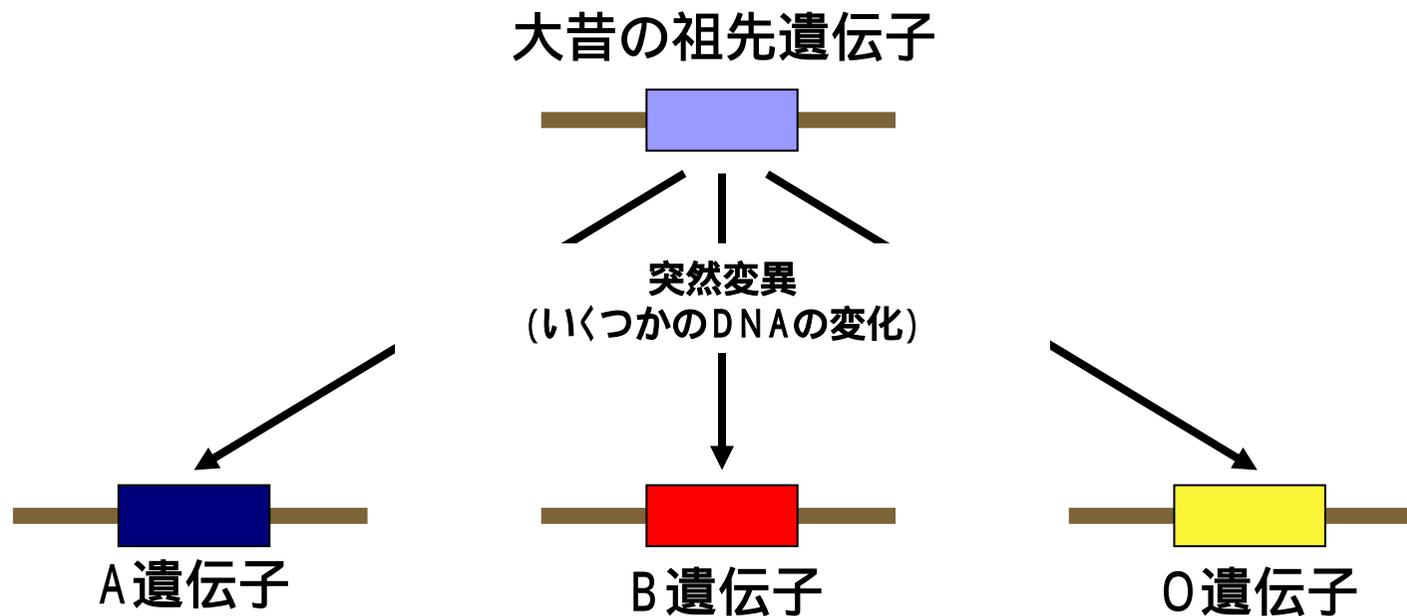


O型:

両親からO遺伝子を2つ(OO)受け継ぎ、
基本型(O型)の糖鎖をもつ



このように、A、B、O遺伝子の働きは異なるのですが、
お互いのDNA配列は非常に似ており、同じ祖先から
突然変異を蓄積することによって生じたのです。



A, B, O 遺伝子全長 (DNAの長さ = 1,059) にわたって
比べても、下の8箇所しか違いがありません。

	261	297	526	657	703	796	803	930
A遺伝子 (A101)	G	A	C	C	G	C	G	G
O遺伝子 (O101)	-	A	C	C	G	C	G	G
B遺伝子 (B101)	G	G	G	T	A	A	C	A

これ以外にも多くのA, B, O 遺伝子があります。この例でのAとOの違いは、Oでは261番目のDNAが失われたことだけです。しかし、これによってOは糖をつける機能を失いました。このように、機能を失えばOになるので、O遺伝子は突然変異でしばしば新たに作られます。

(齋藤・山本 1997を改変)



1塩基多型(SNP, スニップ)とは？

A B O式血液型を決める遺伝子だけでなく、多くの遺伝子では、個人間、あるいは父由来と母由来のものの中でわずかにDNA配列が異なります。

また、「遺伝子」という単位に限定せず、染色体DNA配列の1つの部位にそのような個人間・父・母由来間で違いがあれば、それを「1塩基多型(*SNP)」と呼びます。

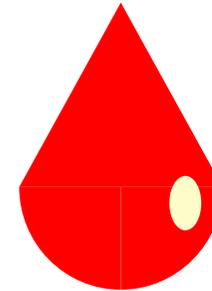


(4) あなたの体質を決める

1 塩基多型 (SNP)

1つの遺伝子の1塩基(または少数の塩基)の違いによって病気を起こす例(一因子疾患)は、
今までにも多く知られていました。

(例) 鎌型赤血球貧血症
血友病
フェニルケトン尿症
伴性ガンマグロブリン血症
色素性乾皮症
ハンチントン病*
など

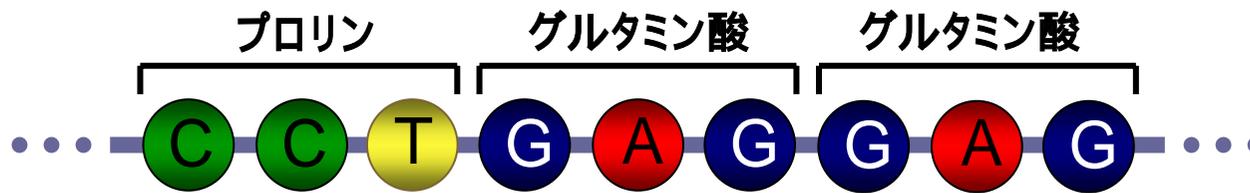


* 第4染色体上のhuntingtin遺伝子内におけるCAGの繰り返し回数の違いによって起こる

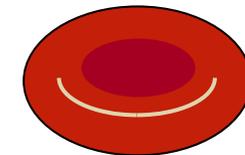
鎌型赤血球貧血症

酸素を運搬するタンパク質である、ヘモグロビン 鎖を指定するDNAが1塩基変わることによって起こる。

正常型



赤血球



貧血症型



赤血球

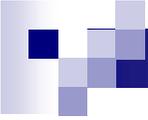


しかし、多くの病気では、ある特定の場所の1塩基の違いが直接病気には結びつかず、病気になりやすい傾向を示すにすぎないということがわかってきました。

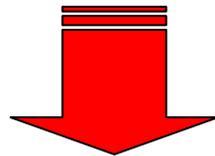


© Bioportal - Jabion

たとえば、第15染色体長腕上(15q24-25.1)の1塩基多型では、通常Gである塩基がTに変わると肺ガンになる危険性が高まりますが、喫煙をしなければその危険性は減ります。

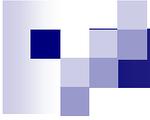


また、ある病気が起きやすくなるような1塩基多型は1箇所とは限らず、複数の遺伝子の中の複数の箇所が関係していることが多いこともわかってきました。



そこで、ゲノム中の全ての1塩基多型と病気の間を徹底的に研究しようとする動きがでてきました。

ゲノムワイド関連解析, GWAS (ジーバス)
(GWAS: Genome-Wide Association Study)



GWASによって1塩基多型との関係が調べられている病気には、次のようなものがあります。

糖尿病
ぜんそく
パーキンソン病
変形性関節症
関節リウマチ
乾癬

統合失調症
痛風(高尿酸血症)
脳動脈瘤
胃ガン
肺ガン
大腸ガン

なお、病気以外のいろいろな性質(身長など)も複数のSNPが関係していると考えられており、これからは、そのような研究も進むと思われます。



自分のゲノムがわかれば、
ある病気へのかかりやすさ、または
太りやすさなどもわかることになります。



大変良いことばかりのようですが、
実は非常に注意が必要でもあります。





ゲノム情報は「個人情報」です。

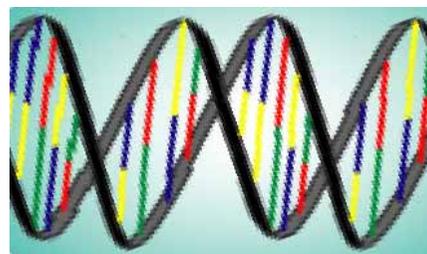
しかも、他の個人情報以上に取り扱いを
厳重にする必要があります。

もし、あなたのゲノム情報を公開することにより、
たとえばある病気になりやすいということで、
何らかの差別を受けることがあってはなりません。

個人ゲノムの公開に関しては、今後多くの有識者による議論とそれに基づく法整備が不可欠でしょう。

また、ゲノム情報はただの「個人情報」ではありません。

それは、あなたの「近親者の情報」でもあります。
あなたのゲノム情報を明らかにすることは、
あなたの親または子供のゲノム情報の
半分について明らかにすることもできます。



© Bioportal - Jabion

要は、他の科学成果と同じで使い方次第です。

厳重な情報管理のもと、重篤な病気に
罹る可能性を前もって知り、それに
対応した生活を行うことで、
健康で長生きできる
可能性があるのですから。

ご静聴有難うございました。