

# 系統分類学 第二回

藤 博幸

(1) 博物学・本草学

(2) リンネ

(3) ダーウィン

(4) 分類学から体系学へ  
進化分類学 vs 数量分類学 vs 分岐分類学

(5) 分子分類学

- タンパク質のアミノ酸配列
- 核酸の塩基配列

などの分子データを用い、各サイトを形質、各サイトを占めるアミノ酸や塩基を形質状態とする。



**化学物質 = 伝統的分類学者の直感の働かない形質**

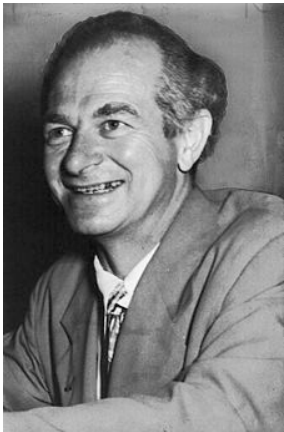
Zuckerkandle & Pauling (1962)

Fitch & Margoliash (1967)

によるアミノ酸配列の解析をきっかけとして始まる

- ※ 進化分類、数量分類、分岐分類は方法の戦い  
分子分類と伝統分類は、分子データと形態データとの戦い

# 分子時計の発見 (1)



ライナス・ポーリング  
Linus Carl Pauling  
(1901-1994)  
量子化学者、生化学者

## Pauling and Zuckerkandle (1962)

二つの生物のヘモグロビンのアミノ酸配列を比較し、その置換数を、化石から推定される、それら生物の分岐時期に対してプロット

近似的な直性関係が得られた。

化石がない生物でも、配列の比較から分岐年代を推定できる。

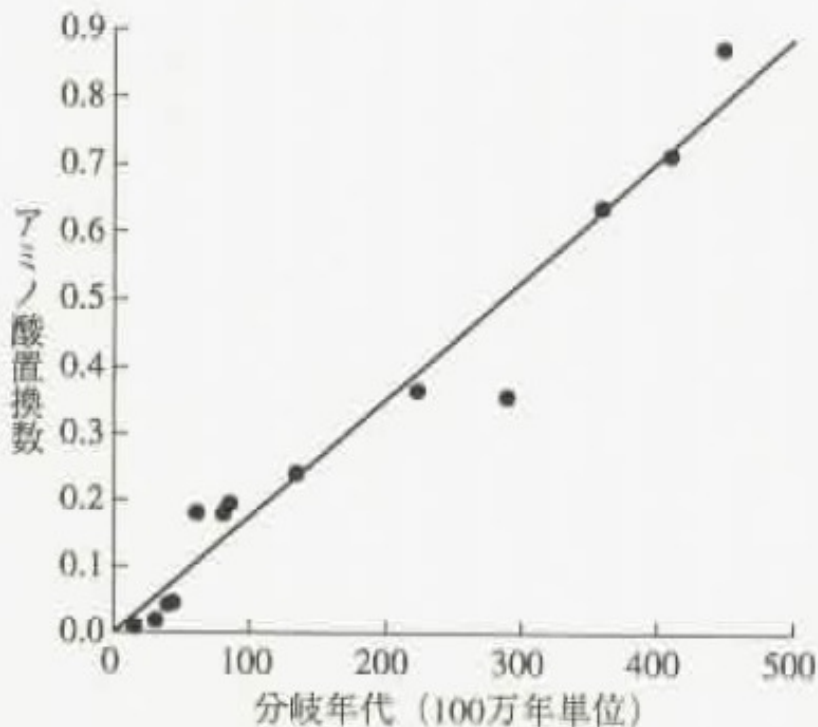


図4-4 ヘモグロビン $\alpha$ 鎖の分子時計

ヒトのヘモグロビン $\alpha$  (上段) と  
マウスのヘモグロビン $\alpha$  (下段)  
アミノ酸は1文字表記で表現

>gi|57013850|sp|P69905.2|HBA\_HUMAN Full=Hemoglobin alpha chain  
MVLSPADKTNVKAAWGKVGGAHAGEYGAELERMFLSFPTTKTYFPHFDLSHG  
SAQVKGHGKVKVADALTNVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSH  
CLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTISKYR

>gi|145301578|ref|NP\_032244.2| hemoglobin subunit alpha  
MVLSGEDKSNIKAAWGKIGGGHGAEYGAELERMFASFPTTKTYFPHFDVSHG  
SAQVKGHGKVKVADALANAAGHLDDLPGALSALSDLHAHKLRVDPVNFKLLSH  
CLLVTLASHHPADFTPAVHASLDKFLASVSTVLTISKYR

# アラインメント (alignment): 相同な配列の対応するアミノ酸 あるいは塩基を対応する位置に並べる操作、あるいは並 べたもの

CLUSTAL format alignment by MAFFT L-INS-i (v7.221)

```
gi|57013850|sp| MVLSPADKTNVKAAWGKVGHAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHG
gi|145301578|re MVLSGEDKSNIKAAWGKIGGHGAEYGAEALERMFASFPTTKTYFPHFDVSHGSAQVKGHG
      ****  **:*:*****:*.*.***** *****:*****
```

```
gi|57013850|sp| KKVADALTNAVAHVDDMPNALSALSSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTP
gi|145301578|re KKVADALANAAGHLDDLPGALSALSSDLHAHKLRVDPVNFKLLSHCLLVTLASHHPADFTP
      *****:*.*.:*:*:*.******:*. **:*
```

```
gi|57013850|sp| AVHASLDKFLASVSTVLTSKYR
gi|145301578|re AVHASLDKFLASVSTVLTSKYR
      *****
```

CLUSTAL format alignment by MAFFT L-INS-i (v7.221)

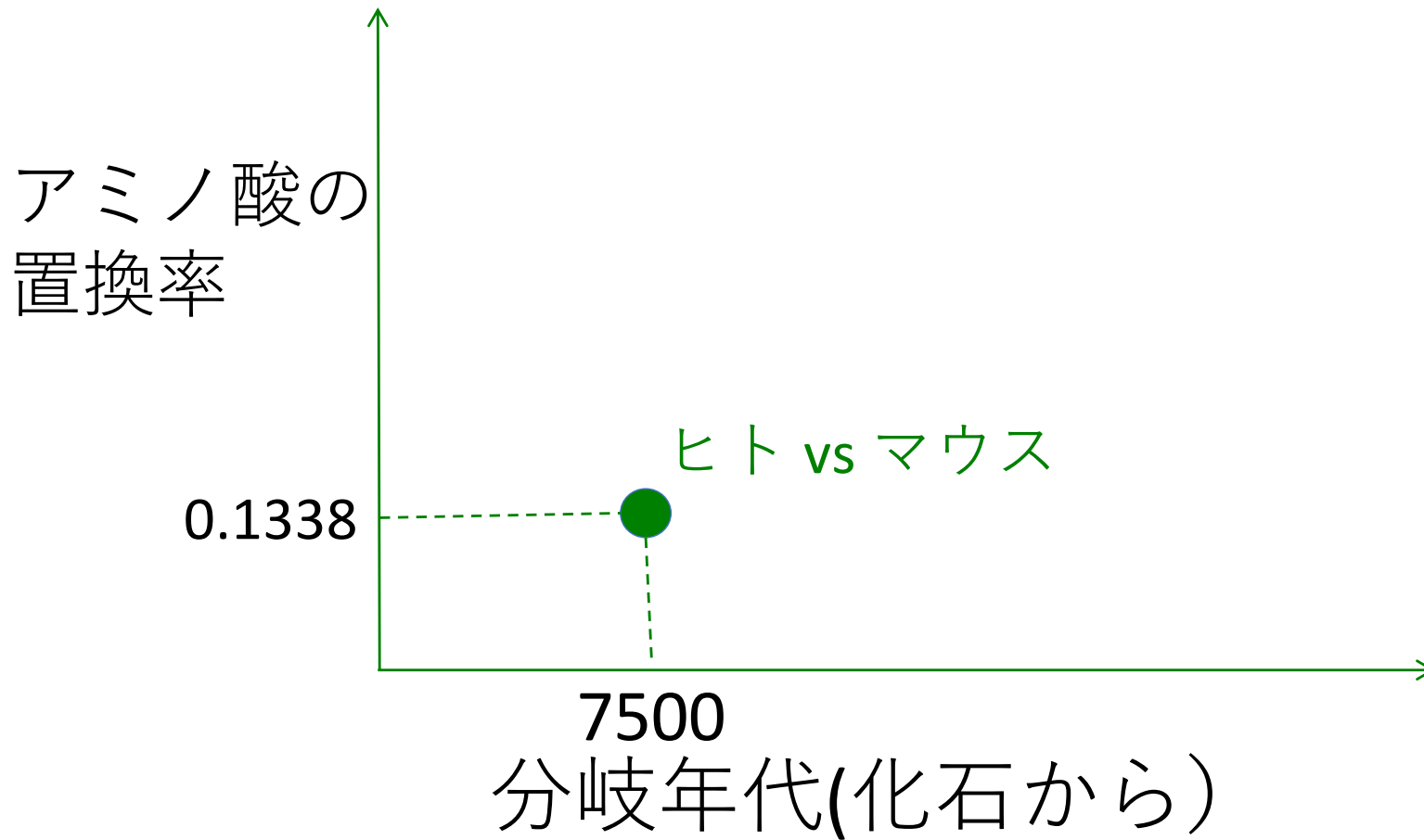
```
gi|57013850|sp| MVLSPADKTNVKAAWGKVGAGHAGEYGAELERMFLSFPTTKTYFPHFDLSHGSAQVKGHG
gi|145301578|re MVLSGEDKSNIKAAWGKIGGHGAEYGAELERMFASFPTTKTYFPHFDVSHGSAQVKGHG
      ****  **:*:*****:*.*..*****  *****:*****

gi|57013850|sp| KKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTP
gi|145301578|re KKVADALANAAGHLDDLPGALSALSDLHAHKLRVDPVNFKLLSHCLLVTLASHHPADFTP
      *****:*.*..*:*:*.*.*****:*.  **:*

gi|57013850|sp| AVHASLDKFLASVSTVLTISKYR
gi|145301578|re AVHASLDKFLASVSTVLTISKYR
      *****
```

$$19/142 = 0.1338028$$

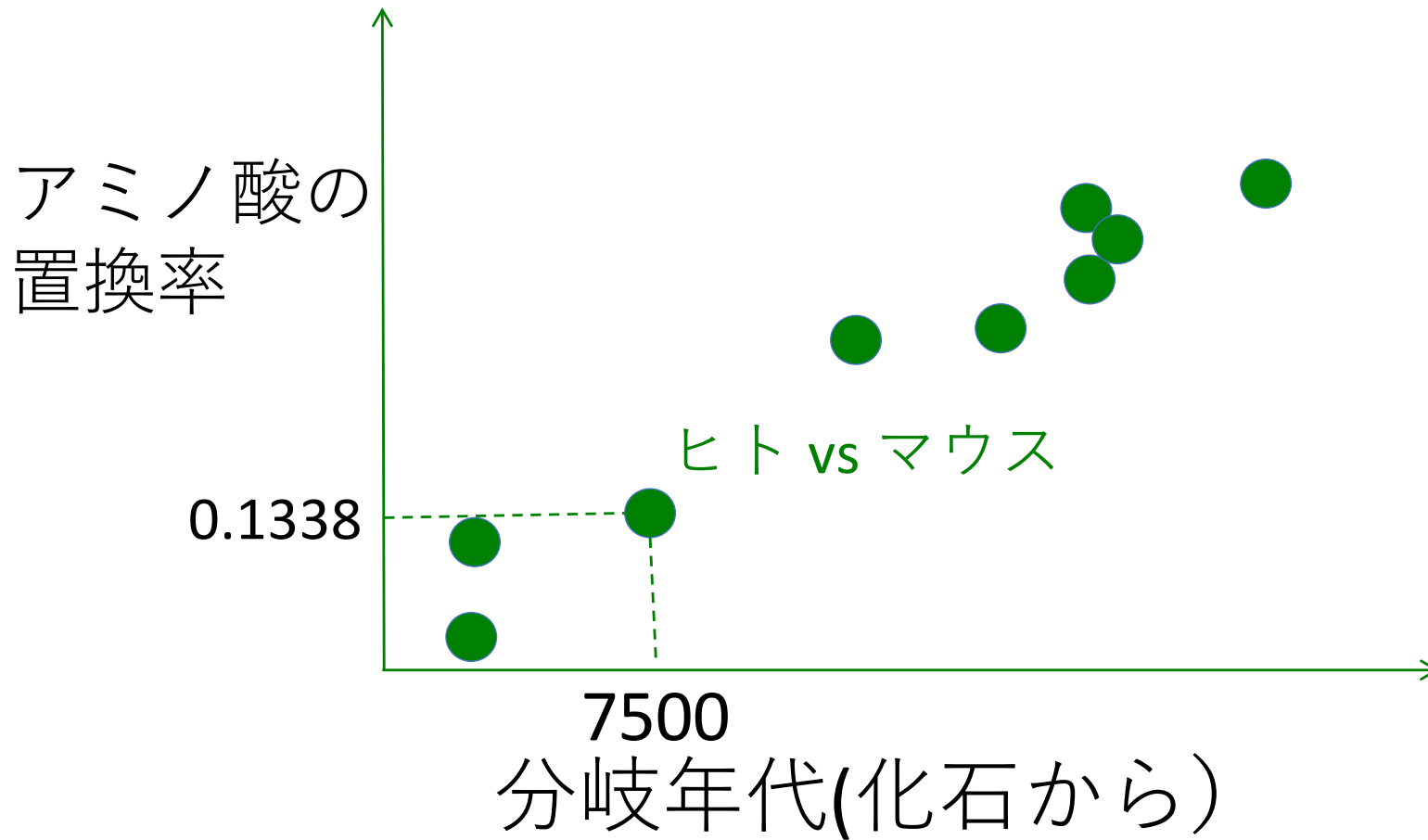
ヒトとマウス      化石から約7500万年前に分岐



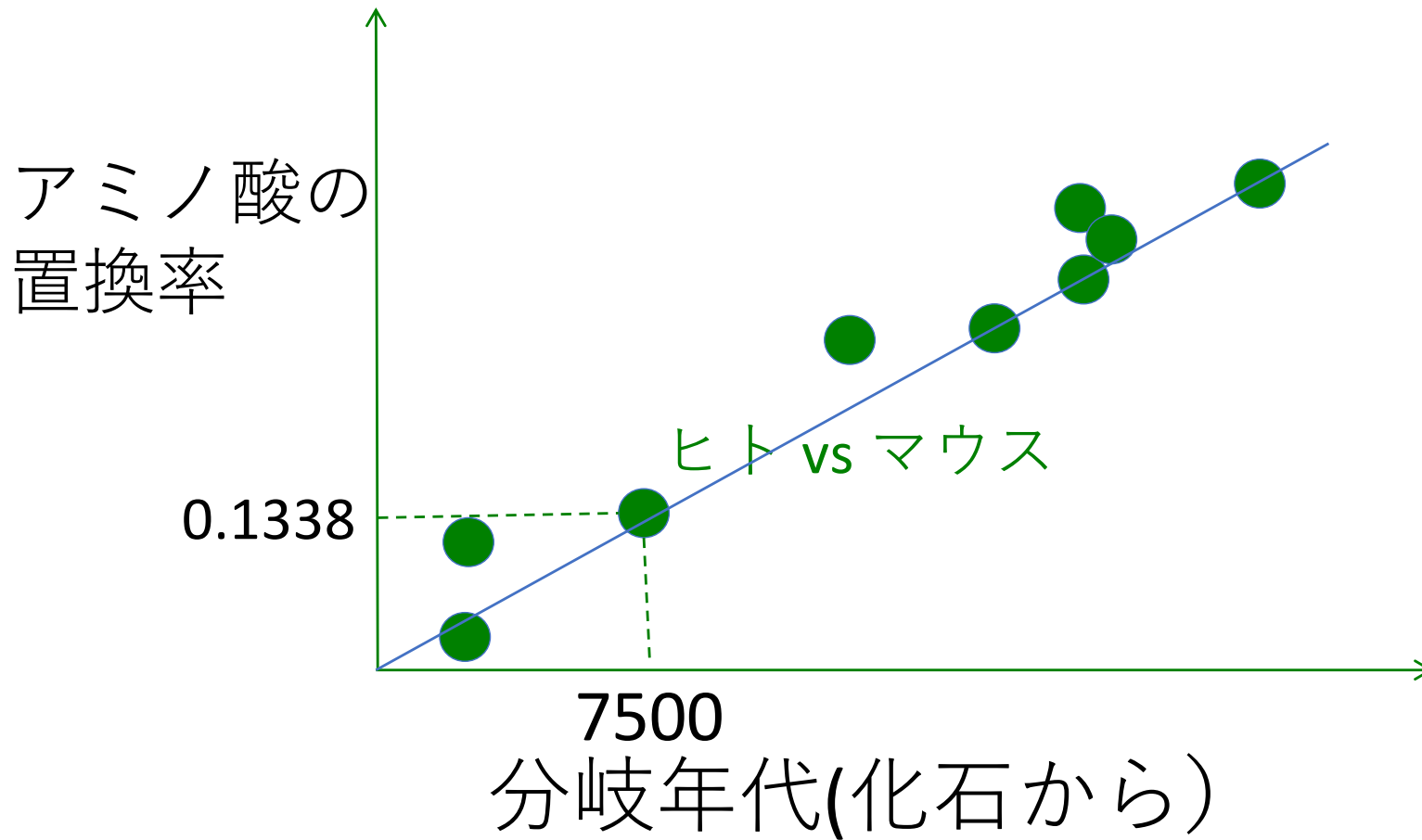
$$19/142 = 0.1338028$$

ヒトとマウス      化石から約7500万年前に分岐



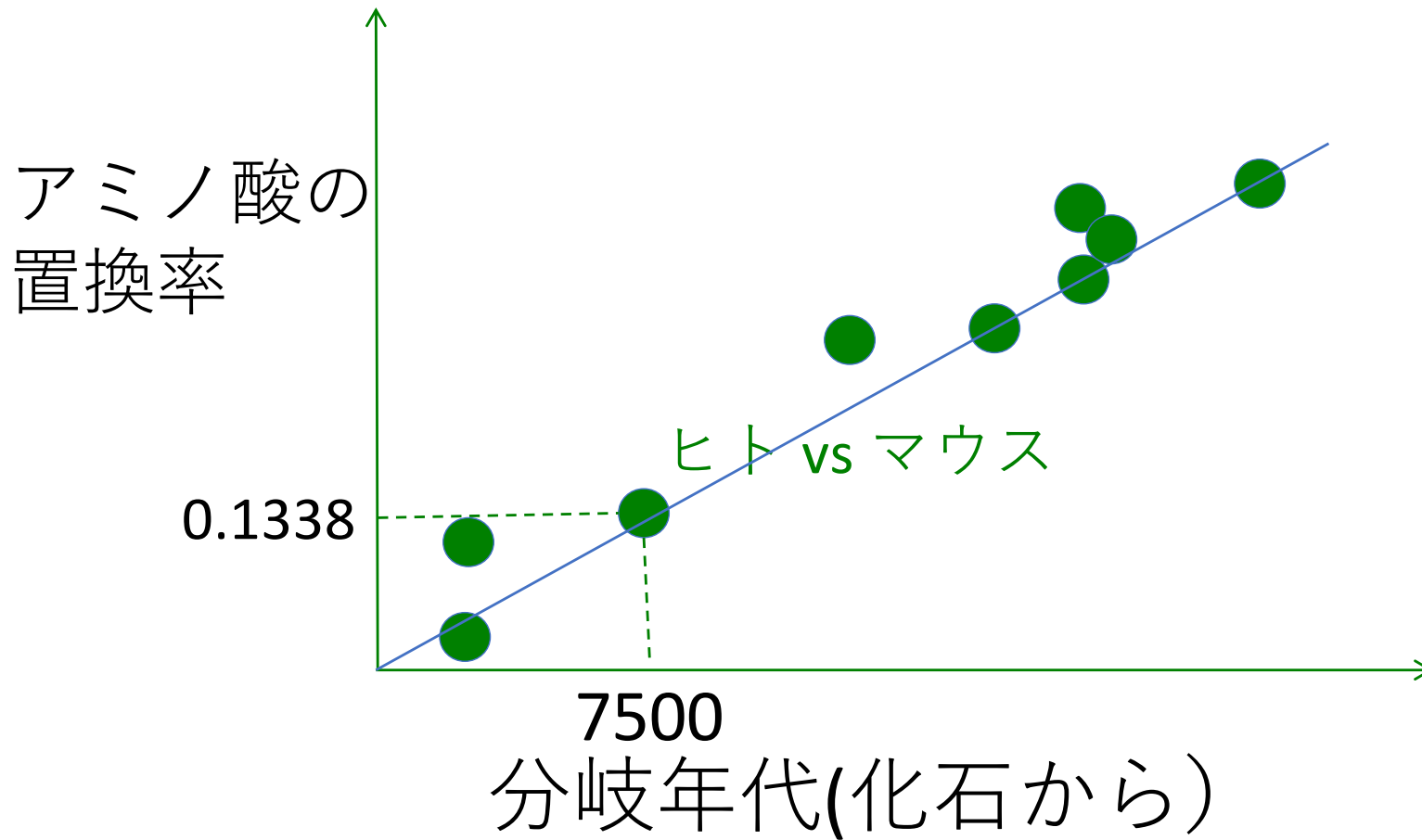


様々な生物のペアについて同様のプロットを作成

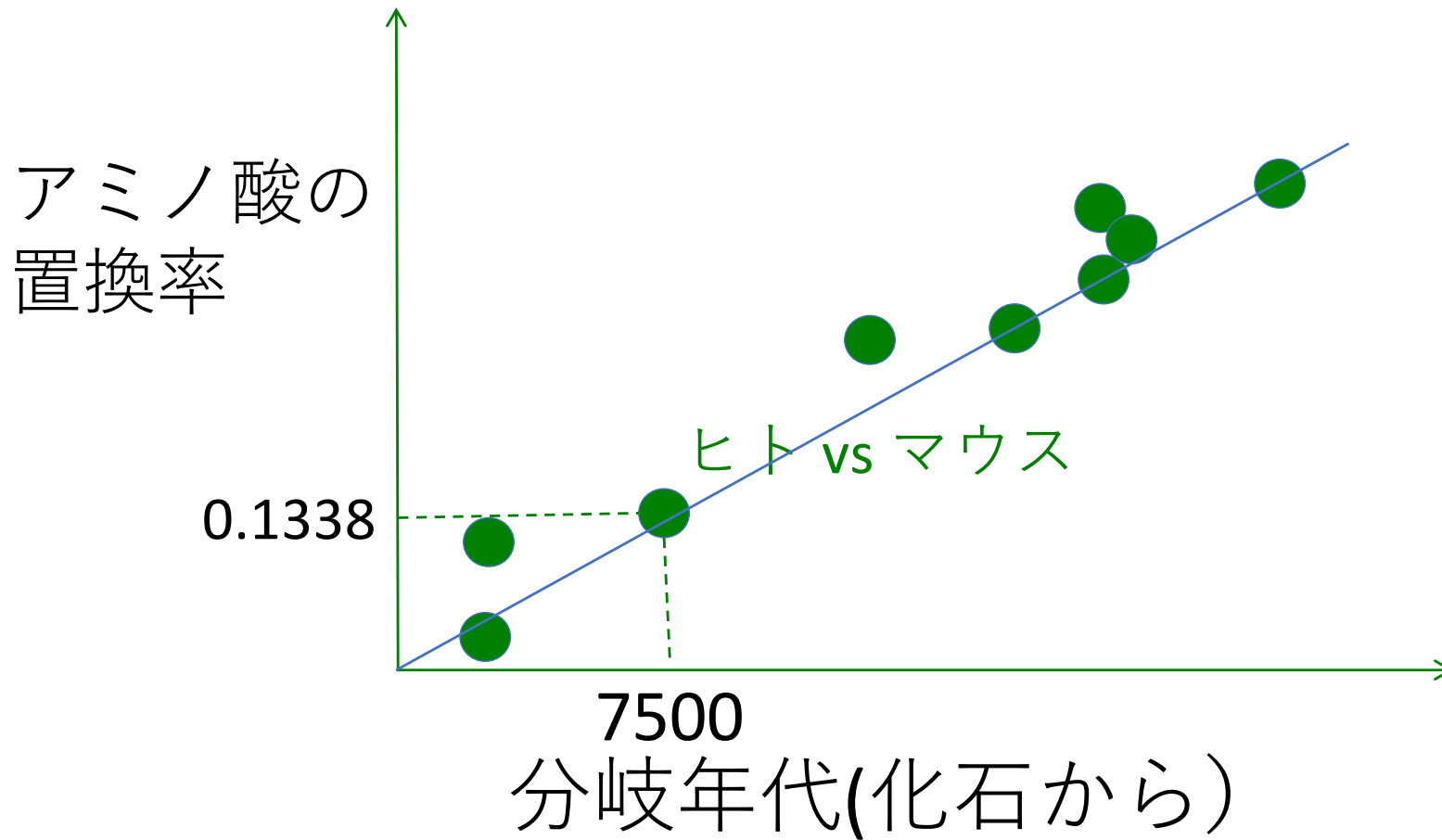


近似的な直線関係

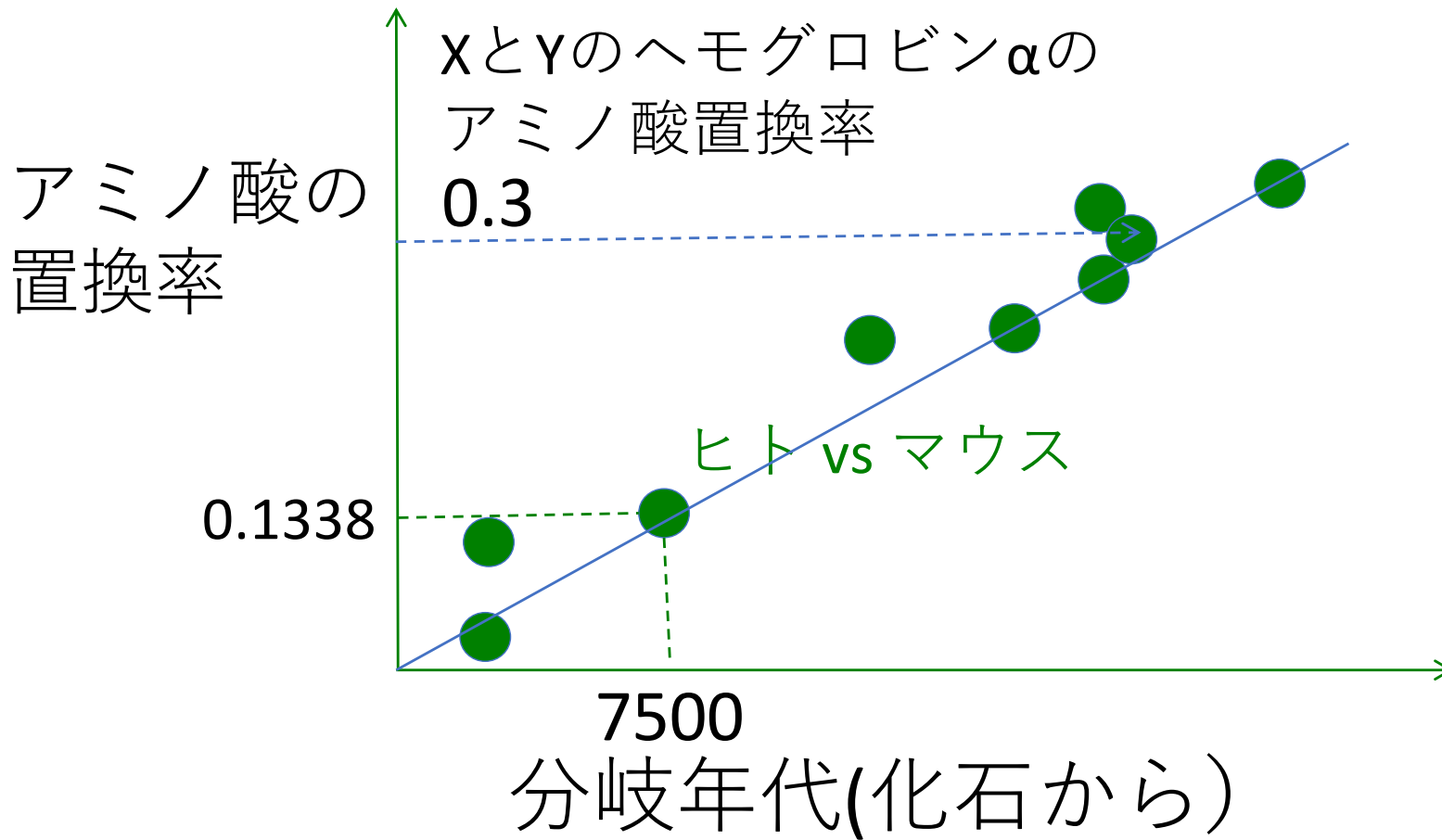
アミノ酸置換数は分岐年代に比例し  
一定のペースで置換している



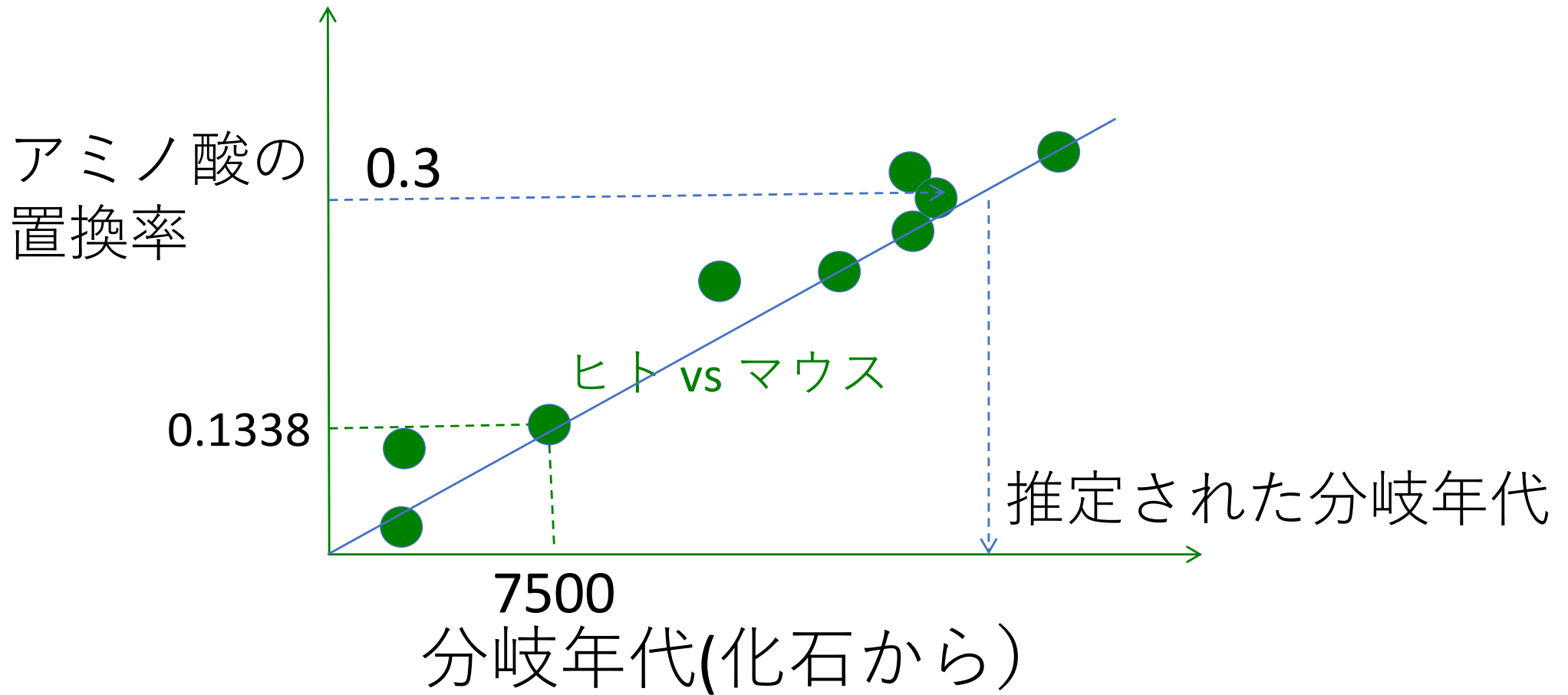
生物の進化の過程での分子の変化：分子進化(molecular Evolution)  
 分子の変化が一定のペースを刻むこと：分子時計(molecular clock)  
 変化の速度:分子進化速度(molecular evolutionary rate)  
 = 直線の傾き = 単位時間あたりのアミノ酸の置換数



分子時計が成立していれば、化石がなくてもアミノ酸配列から分岐年代を推定できる。

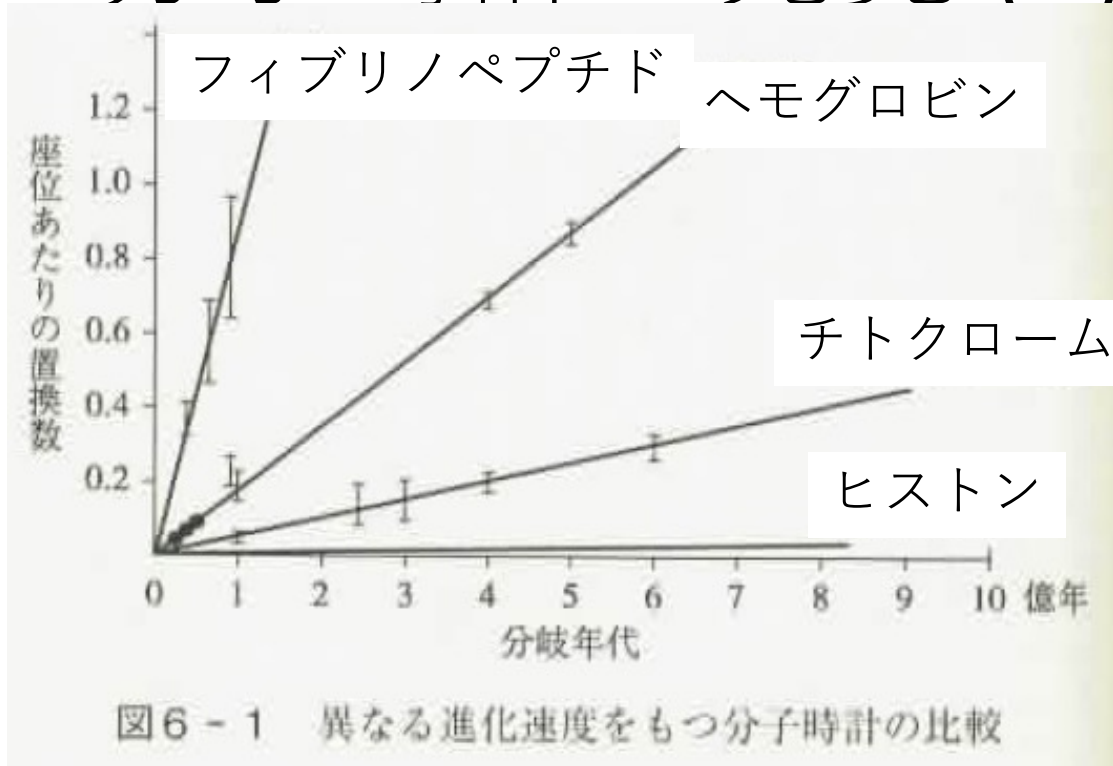


今、現存の生物Xと生物Yの分岐を示す化石はないが、ヘモグロビン $\alpha$ の置換率が0.3であったとすると、分岐年代を直線関係から推定できる。



今、現存の生物Xと生物Yの分岐を示す化石はないが、ヘモグロビン $\alpha$ の置換率が0.3であったとすると、分岐年代を直線関係から推定できる。

# 分子時計の発見 (2)



## Dickerson (1971)

様々はタンパク質で分子時計が成立していること、タンパク質によって分子進化速度が違うことを発見

進化速度の違いは機能的制約の強さを反映している。

||

生物にとって機能的な重要性の高い分子は進化速度が遅く、それほどでもない分子は速く変化する。

# 分子時計の発見 (3)

今日の分子進化学者は、すべての分子に対して、分子進化速度の一定性が成立するとは考えていない。

- 分子によっては変動が激しく、分岐時間にも依存
- 網レベルの比較では近似的な一定性が認められるが、目、科、属、種などのレベルでは一定性が成立しないものも多い。

**primates slow down, rodents speed up**

霊長類では進化速度は遅くなり、齧歯類では速くなる傾向がある。



# 分子系統樹の構築 (1)

## Construction of Phylogenetic Trees

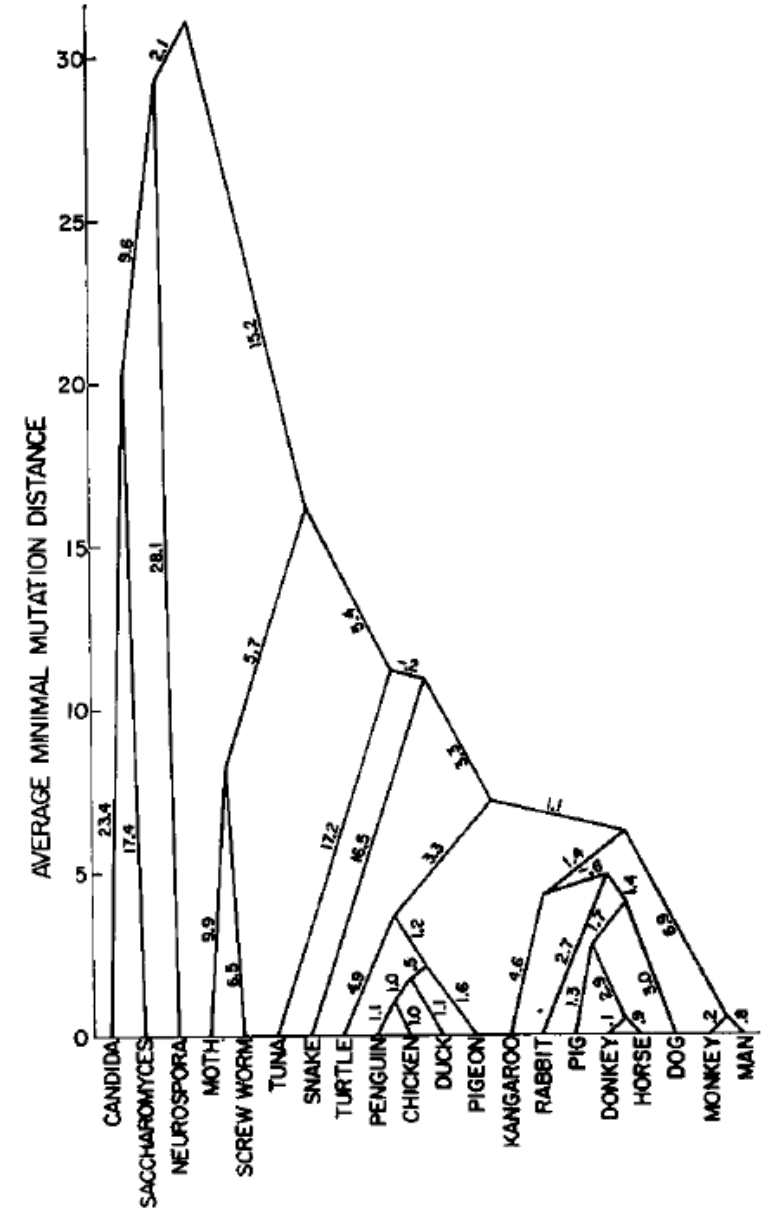
A method based on mutation distances as estimated from cytochrome *c* sequences is of general applicability.

Walter M. Fitch and Emanuel Margoliash

### Fitch and Margoliash (1967)

チトクローム *c* というタンパク質の配列を比べ、その置換数からほ乳類、鳥類、は虫類、昆虫、菌類を含む系統樹を構築した。

これは、それ以前の形態の比較に基づく分類では不可能なものであった。



## 分子系統樹の構築 (2)

分子系統樹は、これまでの分類を反映しており、これまでの分類と食い違う場合、分子系統樹が正しいことがしばしばあった。

特に”**隠蔽種 (cryptic species)**”の発見に力を発揮  
(見た目には区別がつかないが、DNAレベルでは全く異なるもの)

例：ウーズは**16S r RNA**を利用して、界レベルでの隠蔽されていた**古細菌**を発見。

現在、分子系統学的手法は、系統分類の手法として広く受け入れられ、力を発揮している。

# 系統樹とは？

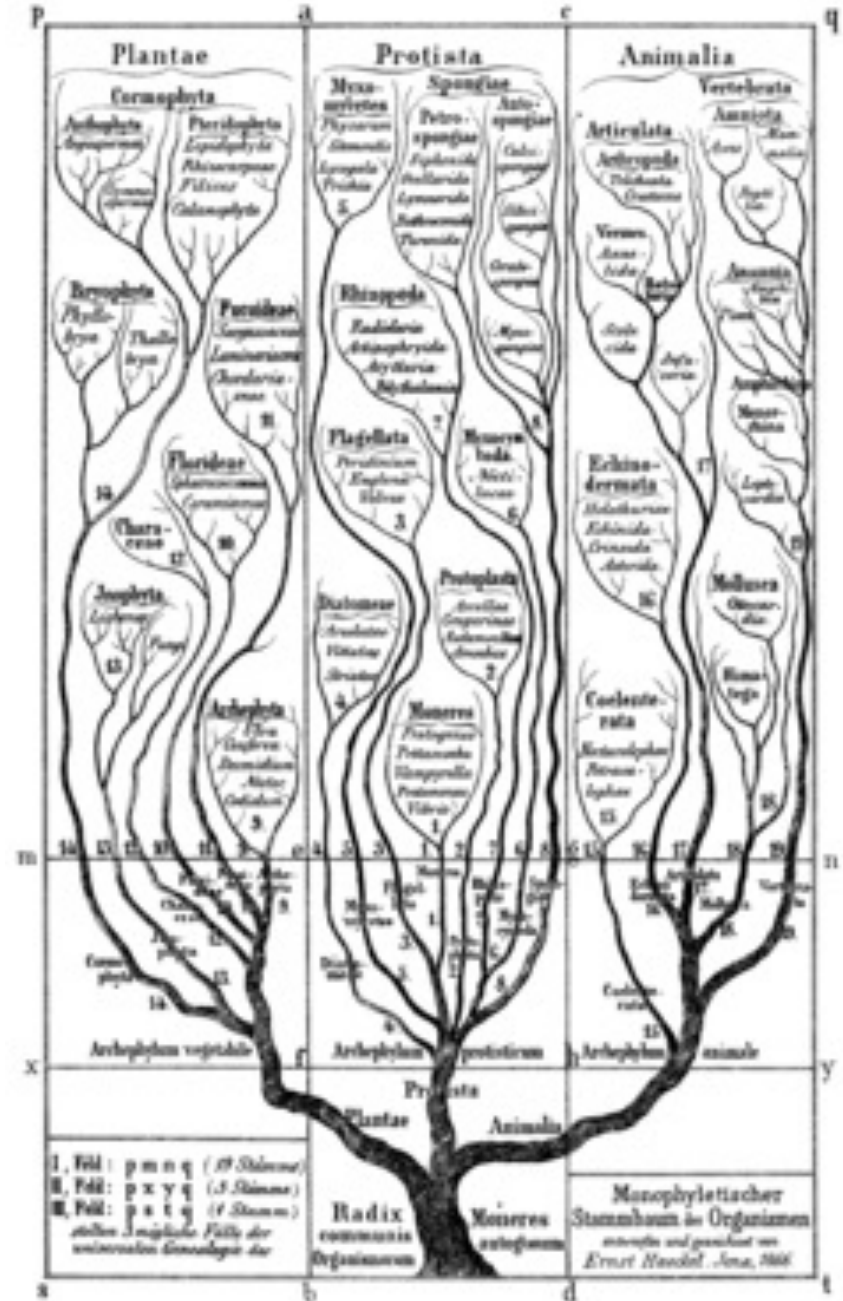
最初に描かれた系統樹  
 エルンスト・ヘッケル (1866)

生物の進化を樹になぞらえて書いた。

- 枝の先に各生物群
- 高い所ほど複雑な体制の生物
- 近縁な系統は互いに近い枝に配置

このような傾向はあるが、確固たるルールに従っているものではない。

ダーウィンの「種の起源」(1859)に、進化の模式的な図が記載されている。



分子系統解析では  
相同な配列（塩基配列 or アミノ酸配列）  
を比較して、分子系統樹が構築される。

## 相同タンパク質の形成

- ・種分化                      ortholog
- ・遺伝子重複                paralog

機能の多様化に特に重要

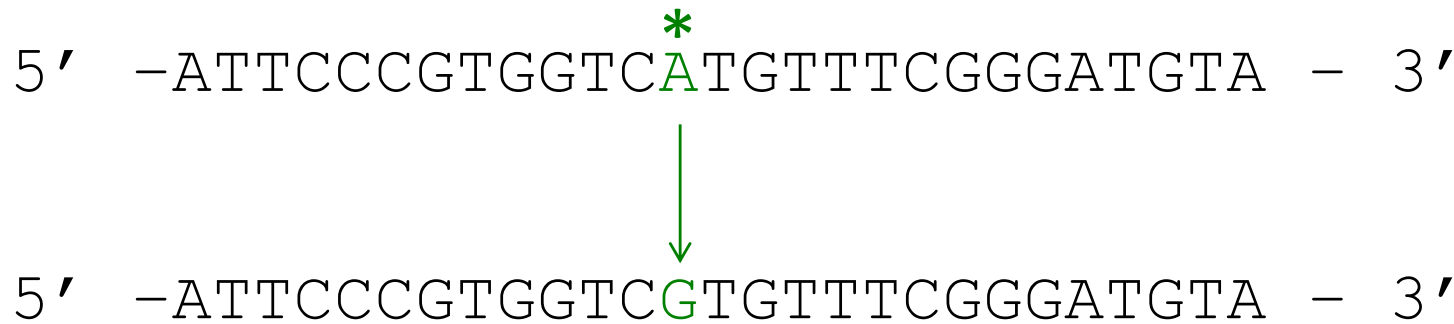
## 分子進化 (Molecular Evolution)

上の機構で分岐した遺伝子に突然変異が生じることで、配列が変化していくこと

基本ステップは

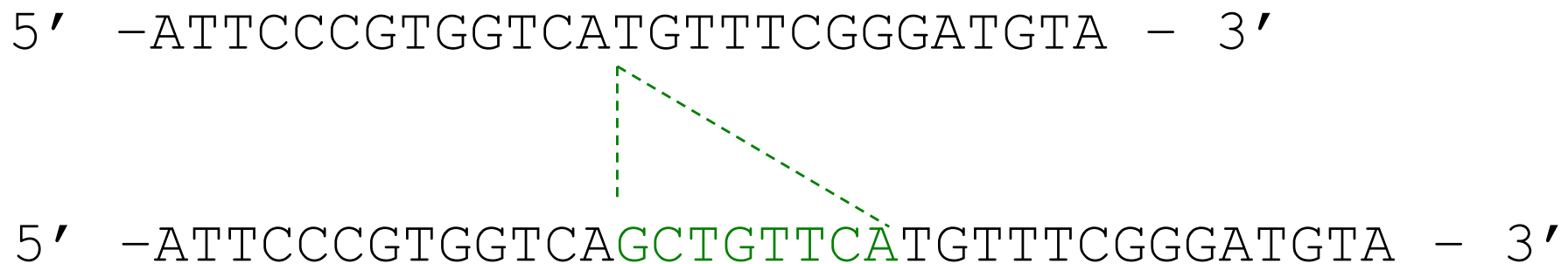
- 塩基(アミノ酸)置換
- 挿入/欠失

## 置換 (substitution)

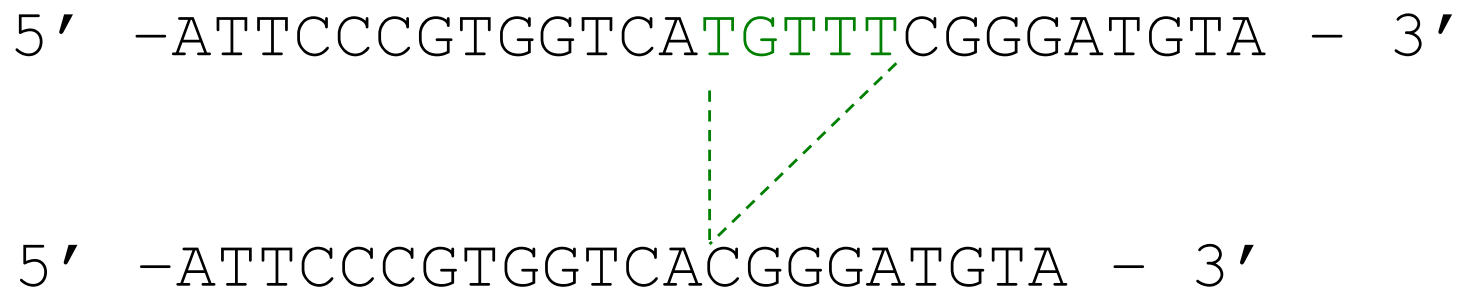


挿入と欠失  
はまとめて  
**INDEL**  
とよぶことも  
ある。

## 挿入 (insertion)



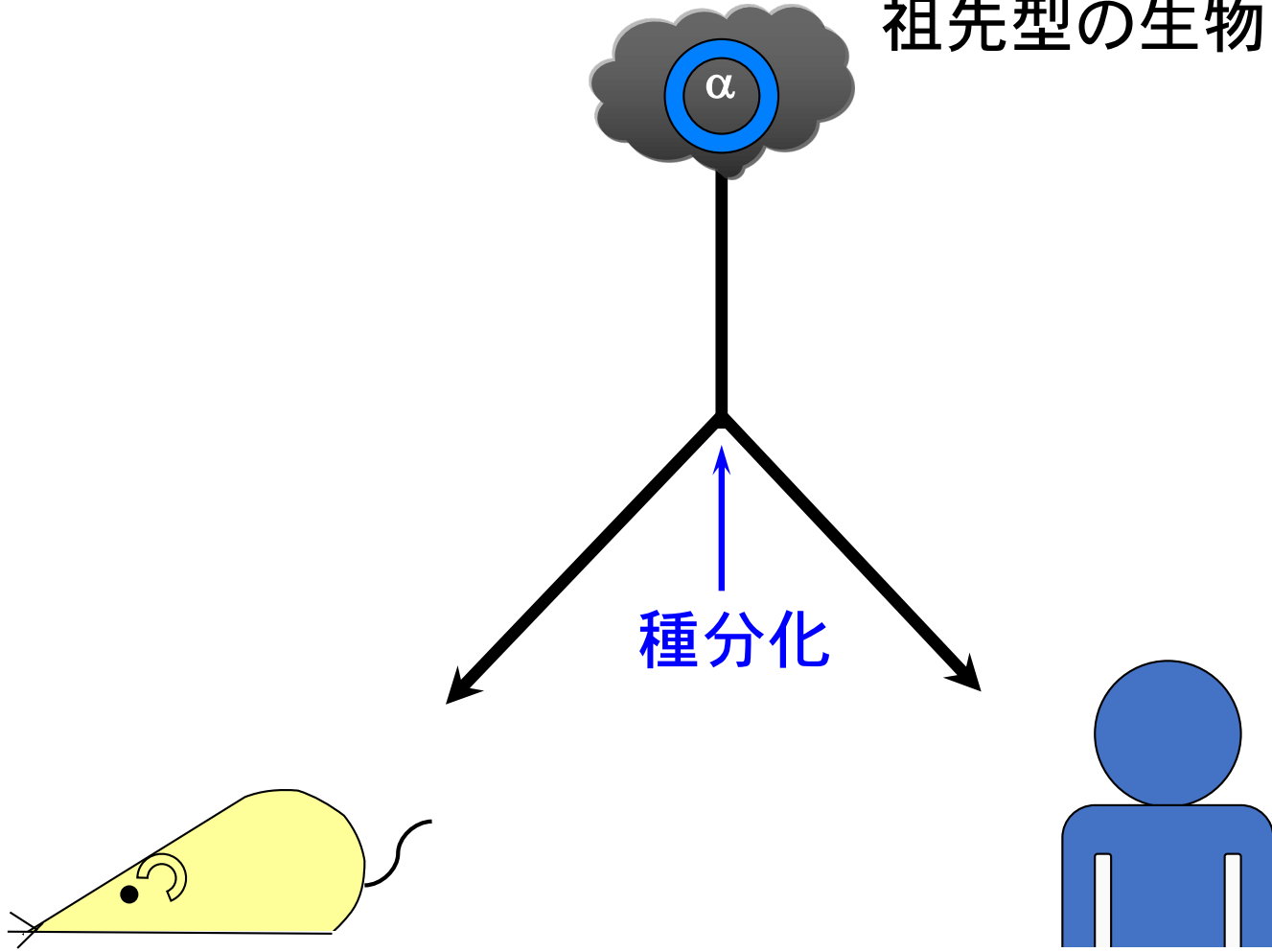
## 欠失 (deletion)



種分化による相同遺伝子の形成

# 種分化

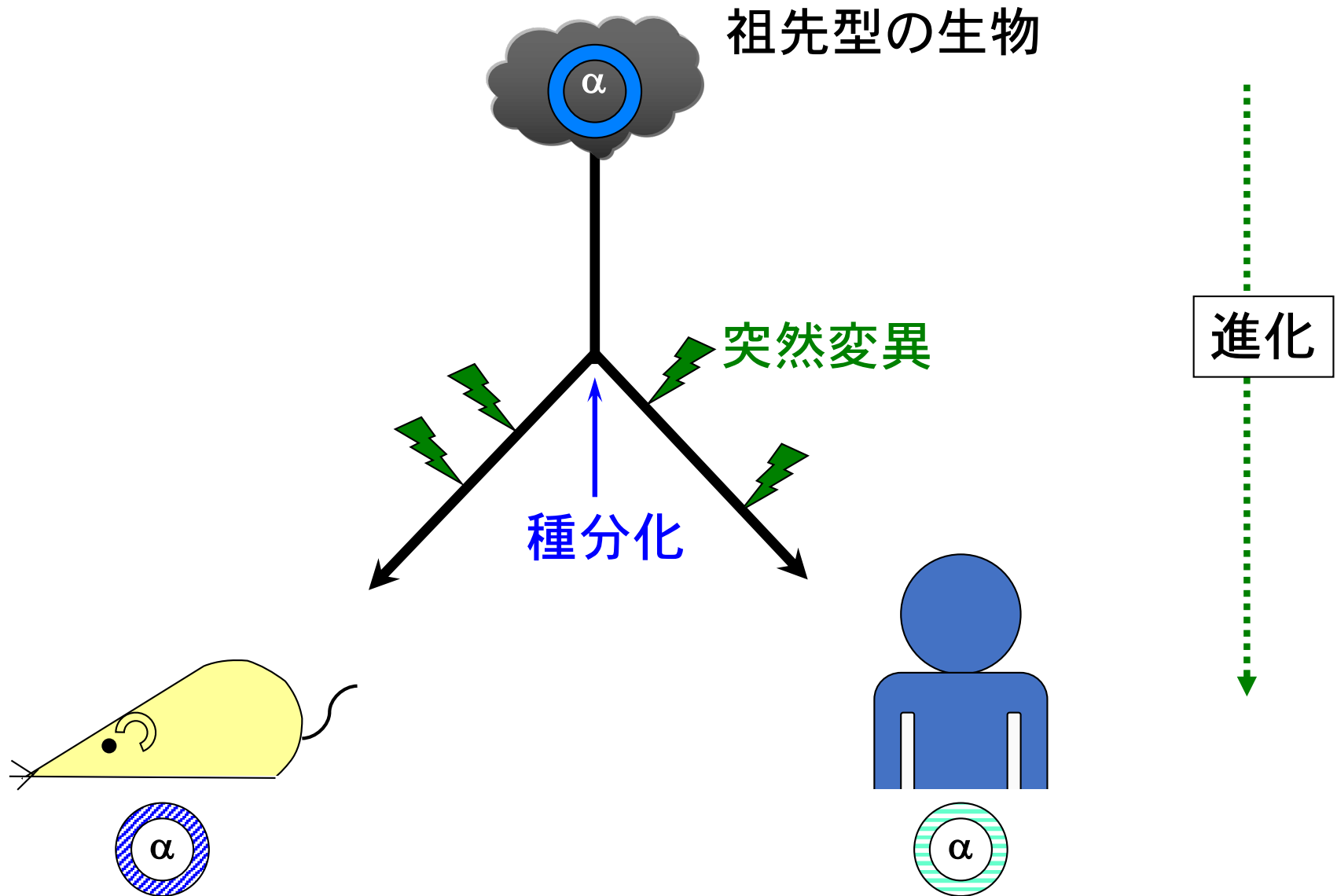
祖先型の生物



進化



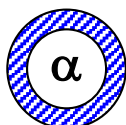
# 種分化



生物の種分化を反映したタンパク質の増加

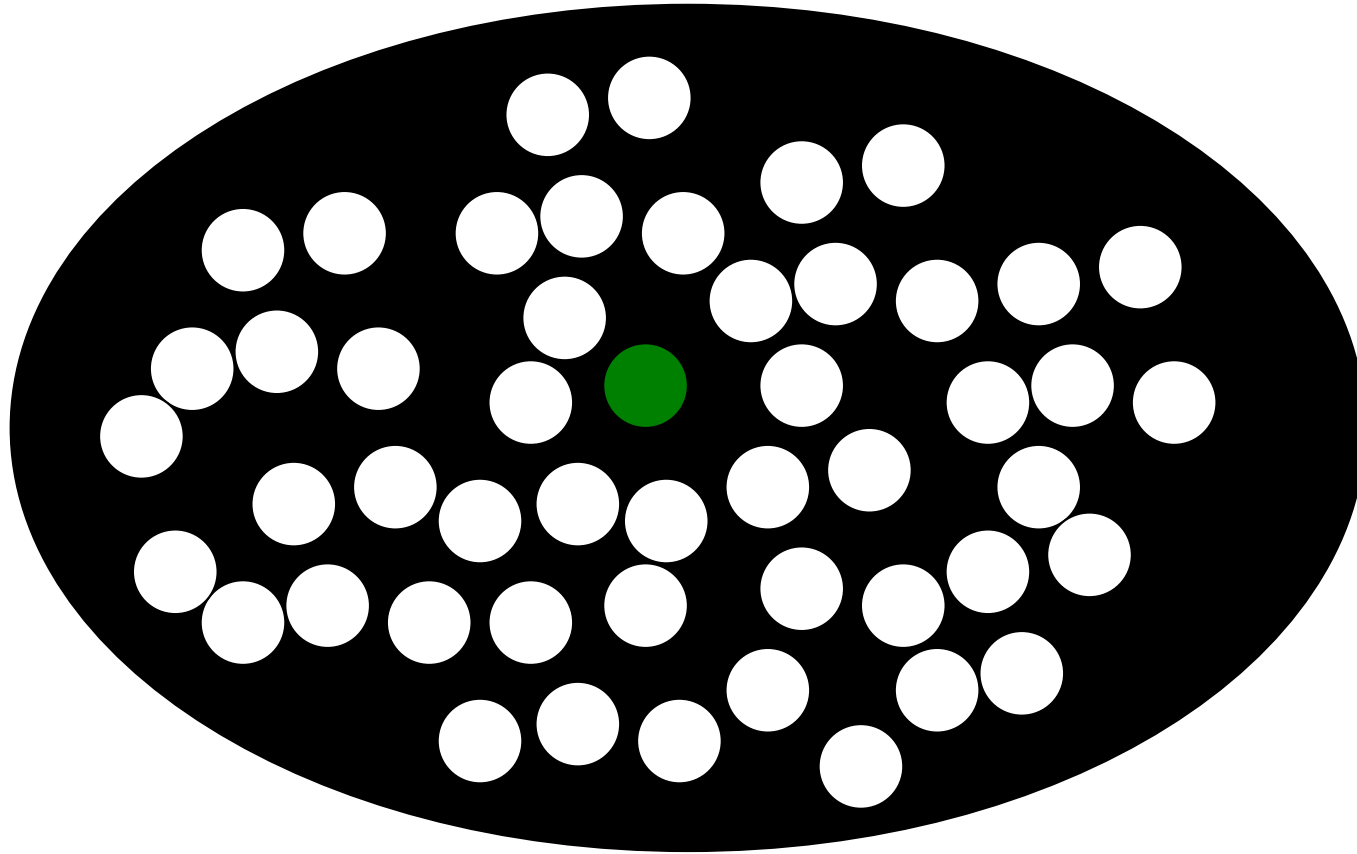
# 種分化

	10	20	30	40	50	
human	VLS	PA	DK	TN	VK	AA
	WG	KV	GA	HAGE	YG	AE
	AL	ER	MF	LS	FP	TT
	KT	YF	PH	FD	LS	H
	SH					
	...	...	...	...	...	...
mouse	VLS	GE	DK	SN	IK	AA
	WG	KI	GG	HA	EY	GA
	EA	LE	RM	FA	SF	PT
	TK	TY	FP	PH	FD	VS
	SH					
	10	20	30	40	50	
	60	70	80	90	100	
human	GSA	QV	KG	HG	KK	VAD
	AL	TNA	VAV	HV	DD	MP
	NA	LS	AL	SA	LS	DL
	HA	HK	LR	VP	VN	FK
	KL					
	...	...	...	...	...	...
mouse	GSA	QV	KG	HG	KK	VAD
	AL	AS	AG	HL	DD	LP
	GA	LS	AL	SA	LS	DL
	HA	HK	LR	VP	VN	FK
	KL					
	60	70	80	90	100	
	110	120	130	140		
human	LSH	CL	LV	TL	AA	HL
	PA	EF	TP	AV	HA	SL
	DK	FL	AS	VS	TV	LT
	SK	YR				
	...	...	...	...	...	...
mouse	LSH	CL	LV	TL	AS	HP
	AD	FT	PA	VH	AS	LD
	DK	FL	AS	VS	TV	LT
	SK	YR				
	110	120	130	140		



生物の種分化を反映したタンパク質の増加

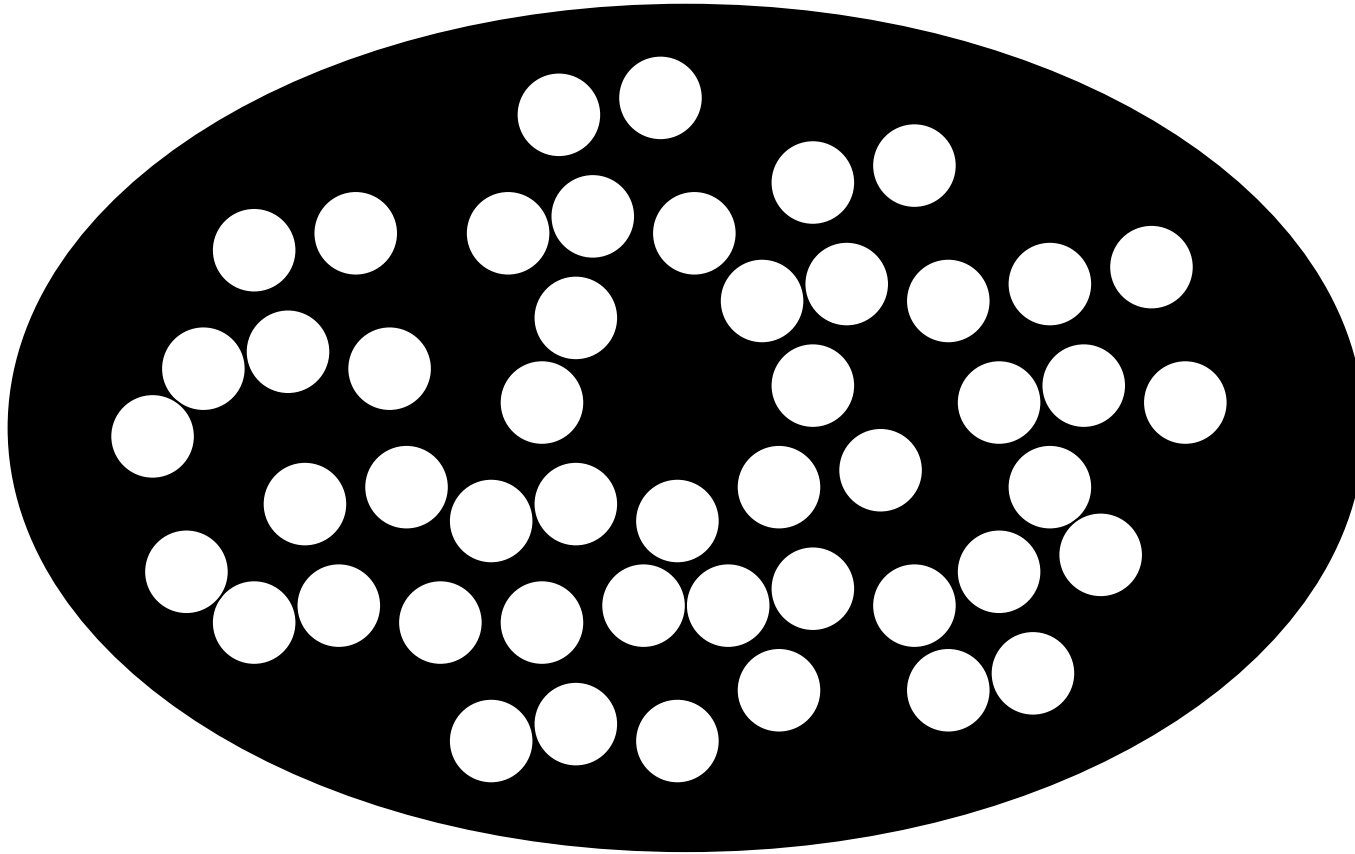
# 突然変異(mutation)と置換 (substitution)



突然変異は集団中の個体に生じる

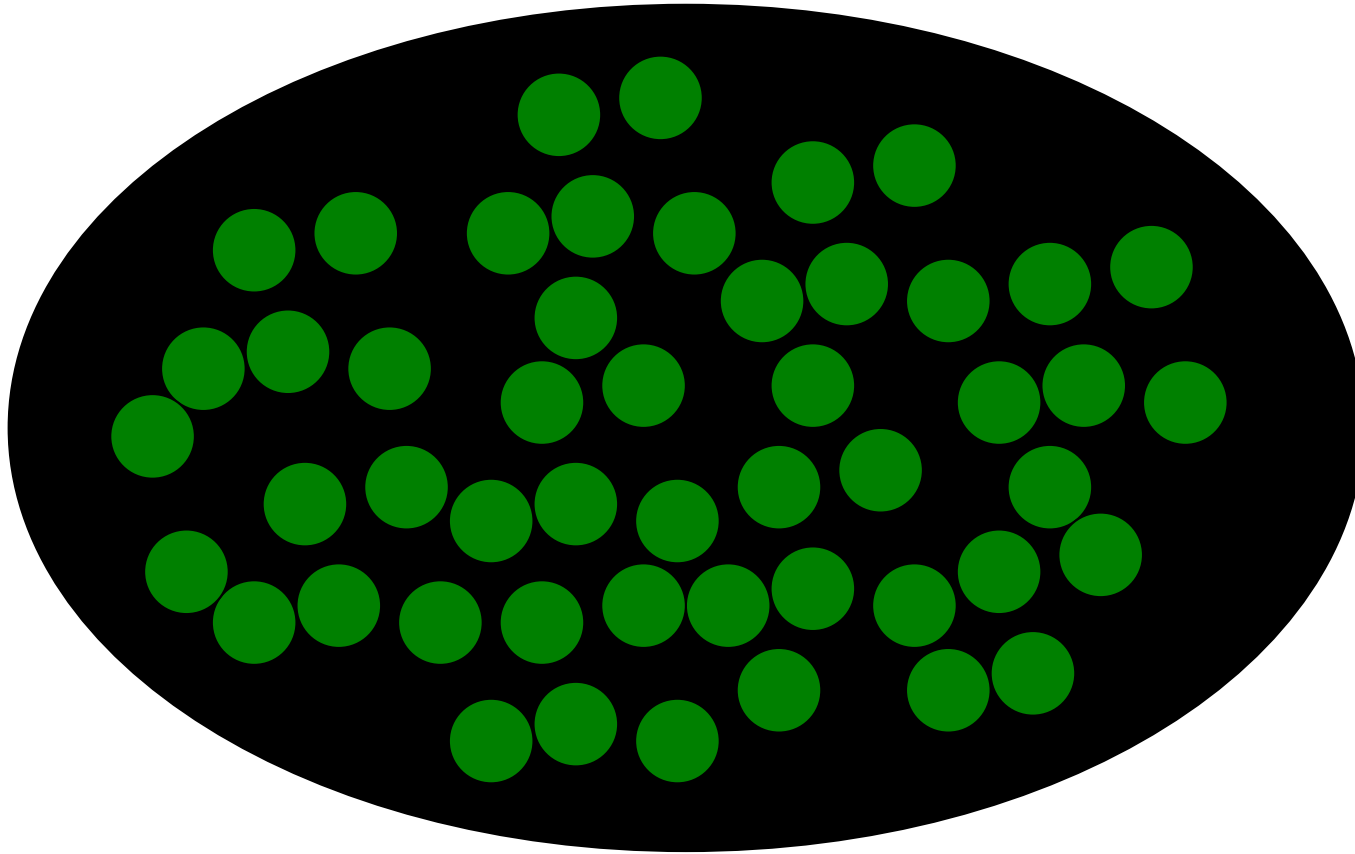
※ 進化に寄与するのは体細胞ではなく生殖系列の細胞に生じる突然変異

# 突然変異(mutation)と置換 (substitution)



有害な突然変異は、集団から除去される  
(負の選択 or 純化淘汰)

## 突然変異(mutation)と置換 (substitution)



有利な突然変異は急速に集団中に広まり  
集団全体がその突然変異遺伝子で置き換  
えられる (正の選択)

# 工業暗化 (industrial melanism)



もともと淡色型と暗色型の体色を持つオオモリエダシャクが存在

19世紀後半から、ヨーロッパの工業都市が発展するにつれて、その付近に生息するガ（蛾）に暗色の個体が増加した

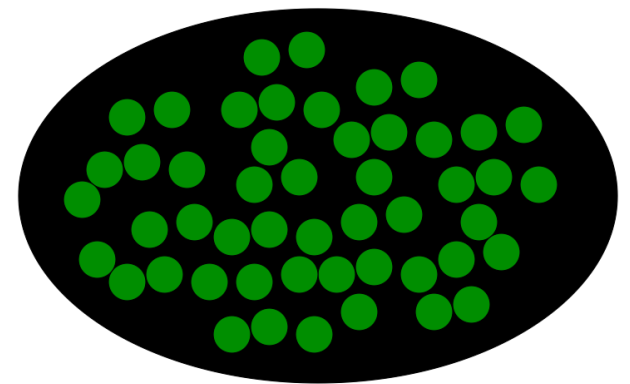
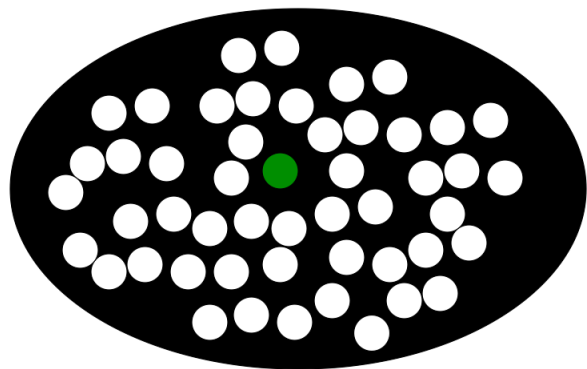
田園都市では白っぽい地衣類が木にはえていて淡色型の方が保護色となり鳥の補色を免れて生き残るのに対して、工業化に伴い地衣類が枯れ、煤煙で木が黒くすすけてしまったため暗色型のほうが保護色となるため



樹皮は白っぽい地衣植物でおおわれており、  
明るい色をしている。



大気汚染のために地衣植物がなくなり、  
樹皮は暗い色をしている。



補足

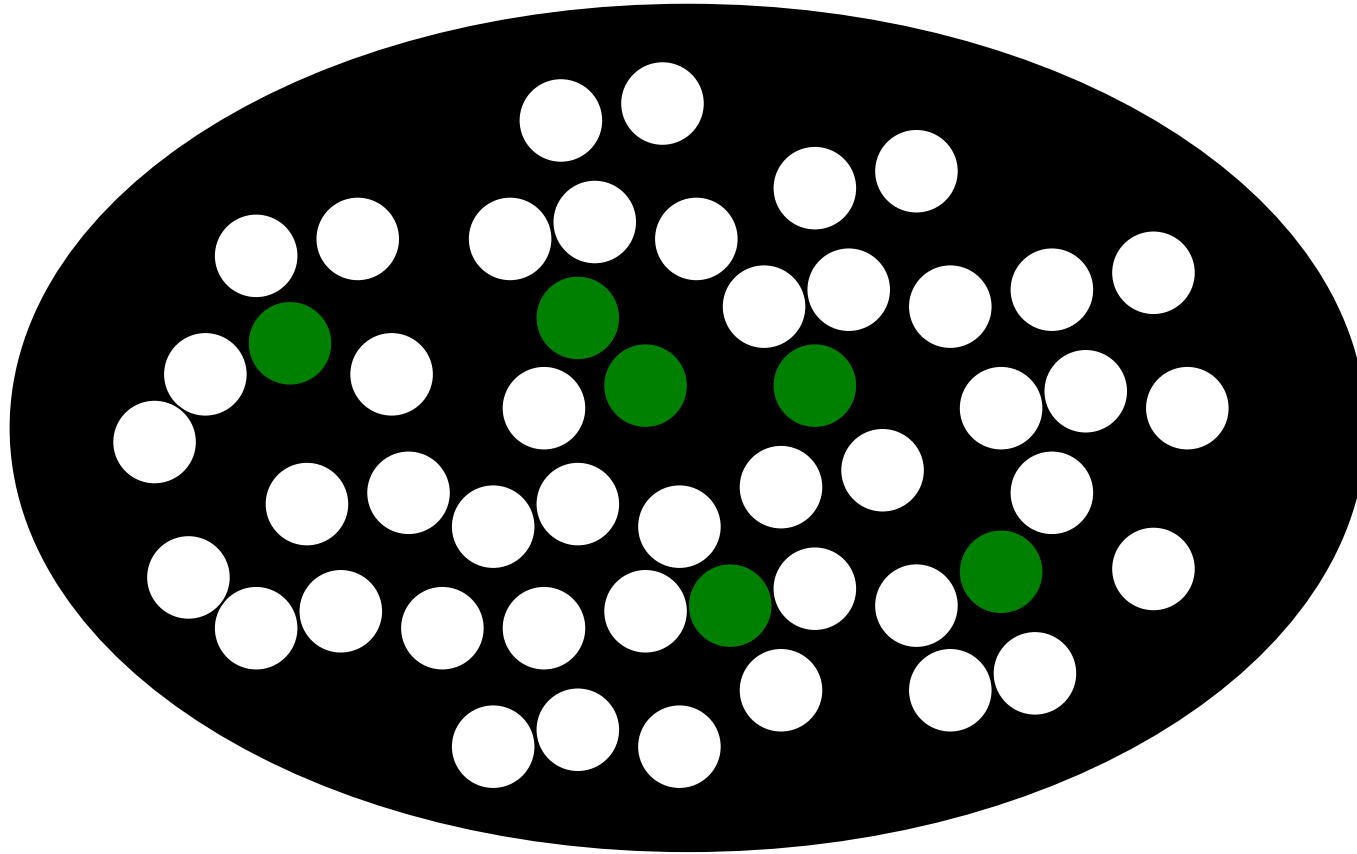
## 大進化と小進化

工業暗化では、種内での形質の変化は起きているが  
“種分化”はおきていない。このようなレベルの進化を小進化とよぶ

これに対し、新しい種あるいは種より高次の分類群の形成、また絶滅  
などのレベルの進化を大進化とよぶ



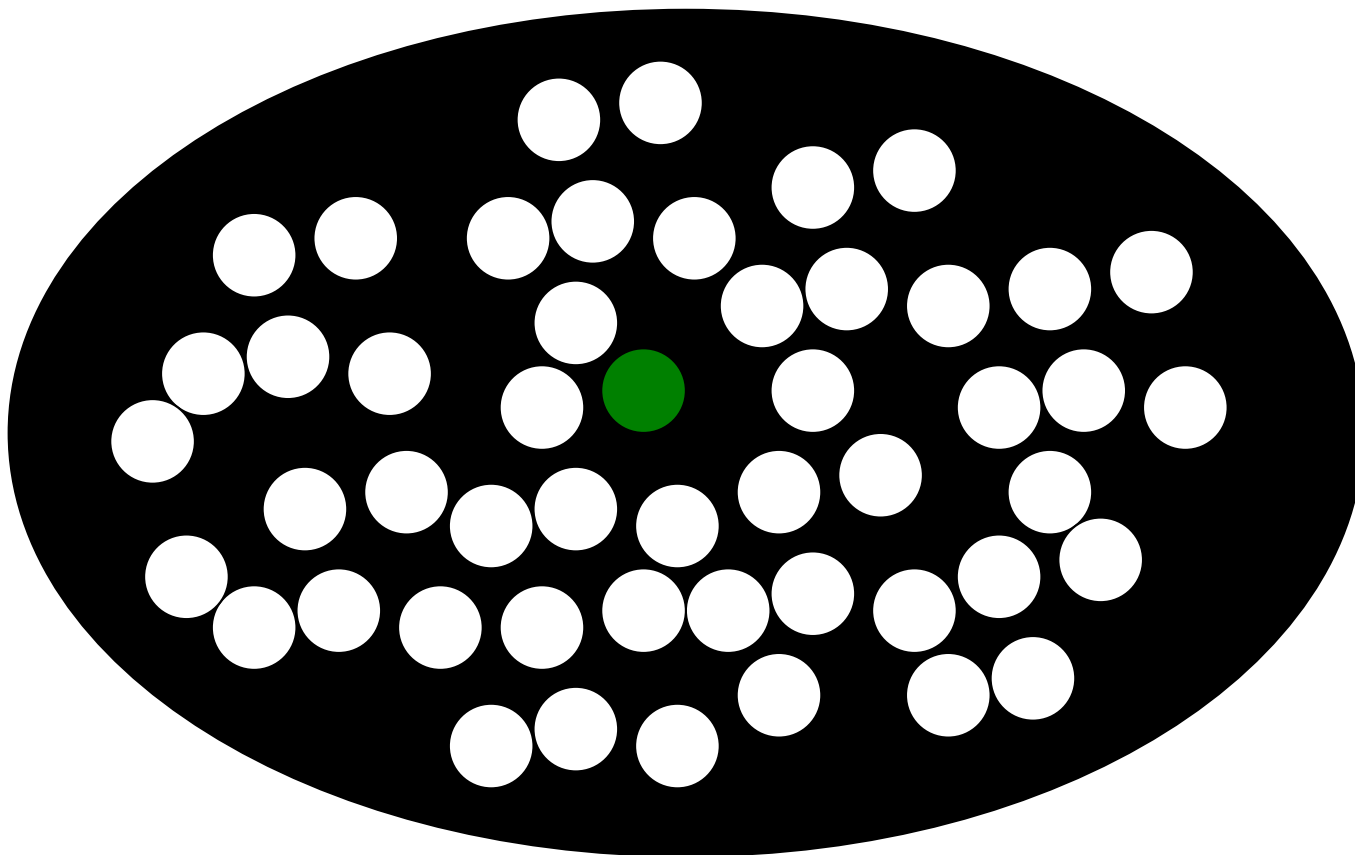
# 突然変異(mutation)と置換 (substitution)



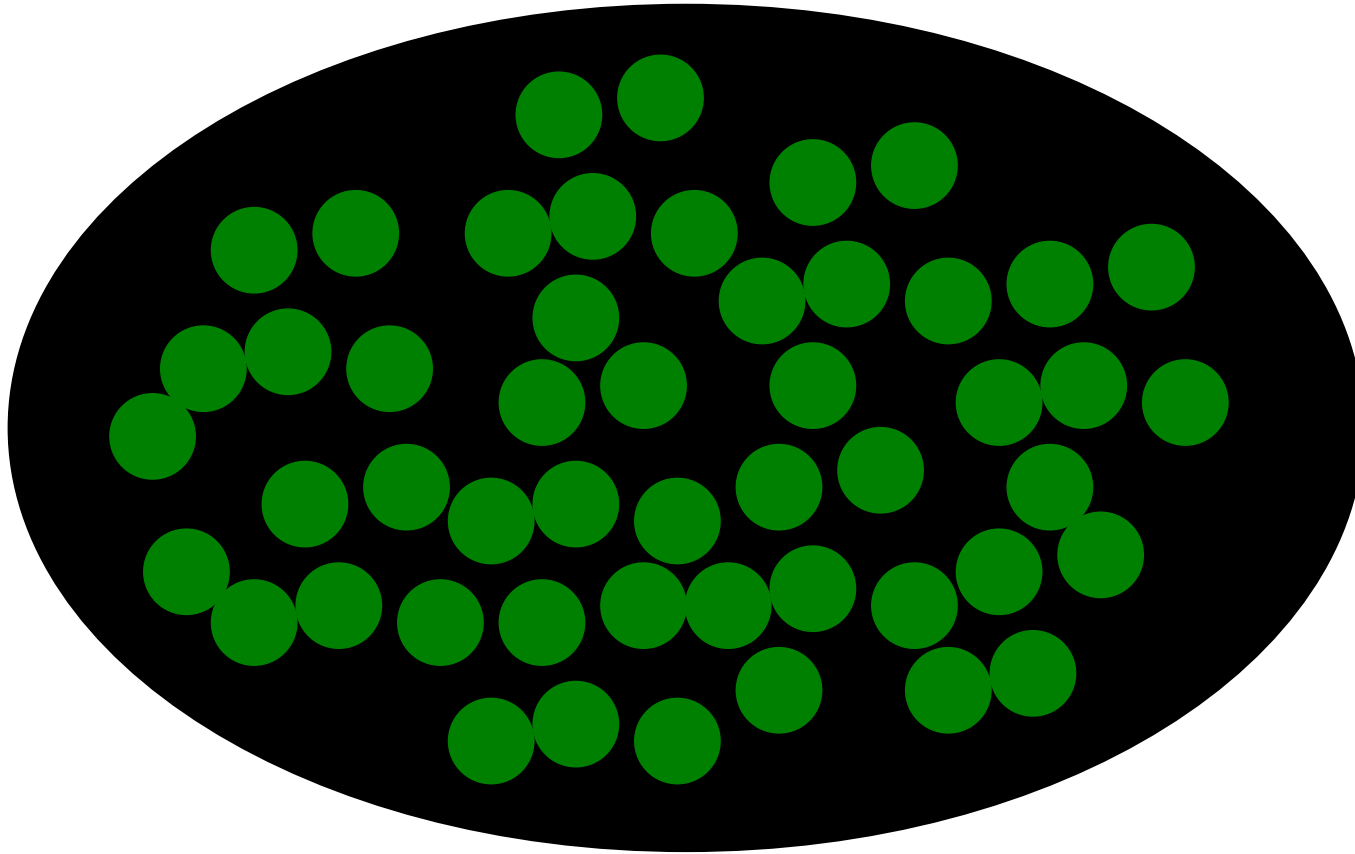
中立な突然変異の頻度はランダムに浮動し  
確率的に集団中に固定 (選択的に中立)

固定にいたる途中の過程      多型(polymorphism)

# 突然変異(mutation)と置換 (substitution)



# 突然変異(mutation)と置換 (substitution)



個体に生じた突然変異が集団全体に広まること **固**

突然変異が固定されること **置**

## The industrial melanism mutation in British peppered moths is a transposable element

Arjen E. van't Hof<sup>1\*</sup>, Pascal Campagne<sup>1\*</sup>, Daniel J. Rigden<sup>1</sup>, Carl J. Yung<sup>1</sup>, Jessica Lingley<sup>1</sup>, Michael A. Quail<sup>2</sup>, Neil Hall<sup>1</sup>, Alistair C. Darby<sup>1</sup> & Ilik J. Saccheri<sup>1</sup>

Discovering the mutational events that fuel adaptation to environmental change remains an important challenge for evolutionary biology. The classroom example of a visible evolutionary response is industrial melanism in the peppered moth (*Biston betularia*): the replacement, during the Industrial Revolution, of the common pale *typica* form by a previously unknown black (*carbonaria*) form, driven by the interaction between bird predation and coal pollution<sup>1</sup>. The *carbonaria* locus has been coarsely localized to a 200-kilobase region, but the specific identity and nature of the sequence difference controlling the *carbonaria*-*typica* polymorphism, and the gene it influences, are unknown<sup>2</sup>. Here we show that the mutation event giving rise to industrial melanism in Britain was the insertion of a large,

involved in wing pattern development or melanization. By extending the association mapping approach to a larger population sample and more closely spaced genetic markers (see Methods), we narrowed the *carbonaria* candidate region to about 100 kb (Fig. 1a). The candidate region resides entirely within the span of one gene — the orthologue of *Drosophila cortex* (*cort*), the only known function of which is as a cell-cycle regulator during meiosis<sup>11</sup>. In *B. betularia*, *cortex* consists of eight non-first exons, multiple alternative first exons (of which only two, 1A and 1B, are strongly expressed in developing wing discs), and a very large first intron (Fig. 1b).

The rapid spread of *carbonaria* gave rise to strong linkage disequilibrium<sup>2</sup>, such that many sequence variants are associated with the *carbonaria* phenotype. This poses a challenge for isolating the specific

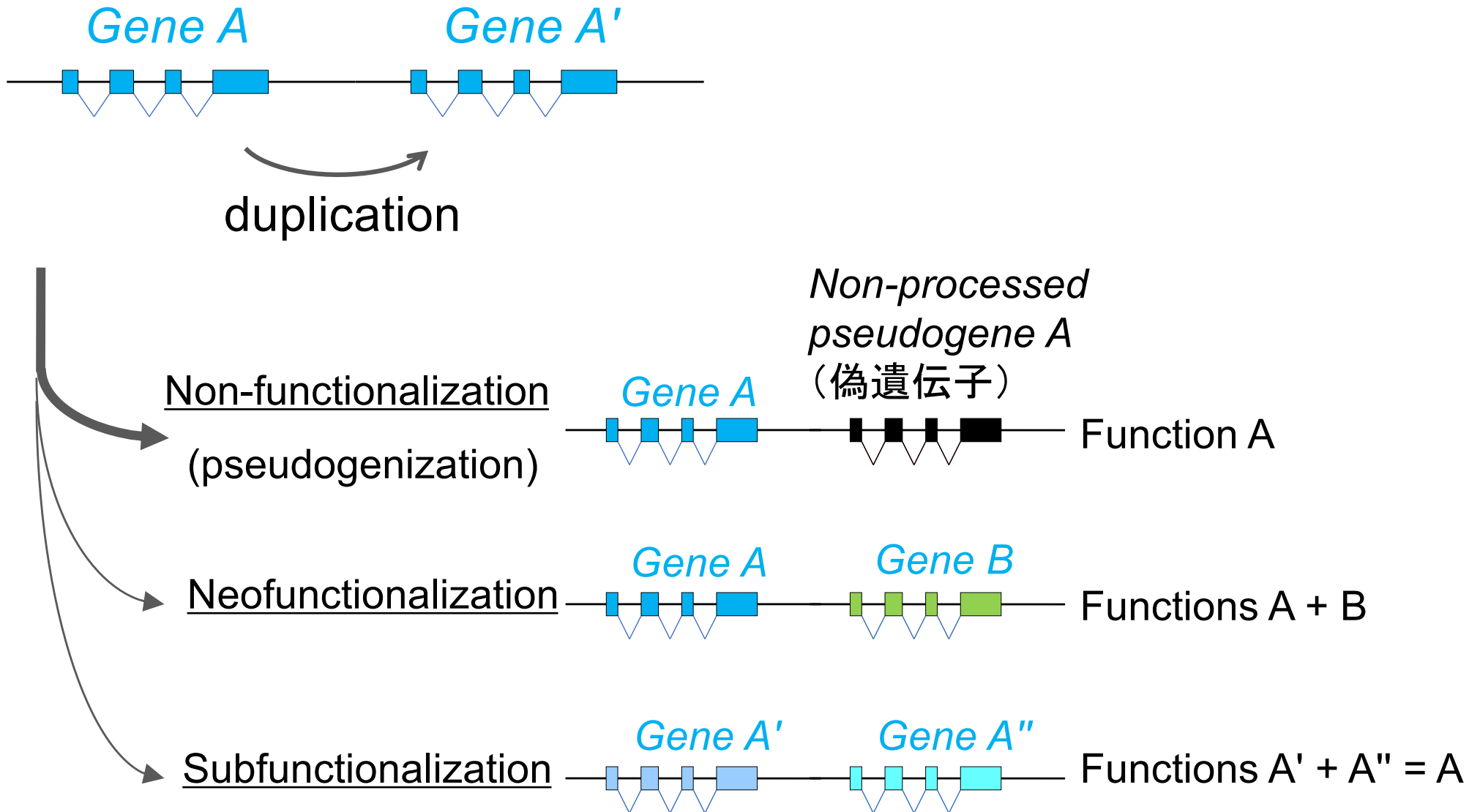
オオシモフリエダシャクの17番染色体 cortex遺伝子  
ショウジョウバエでは減数分裂時の細胞周期の制御に関与  
黒い蛾のcortex 遺伝子第一イントロンにトランスポゾン

羽化時に発現が上昇するが、トランスポゾン挿入によるさらに発現が上昇

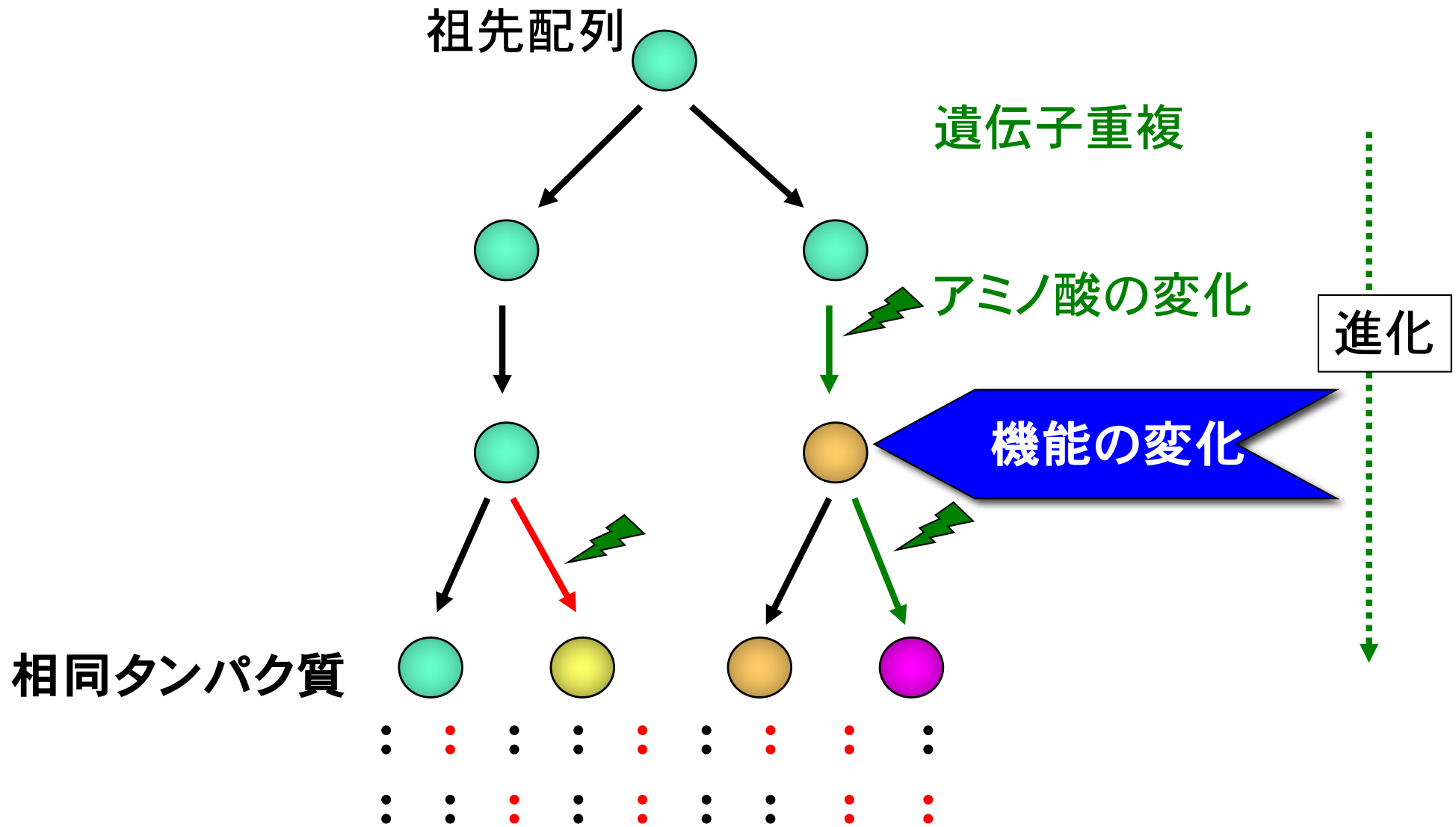
まだ、この遺伝子が工業暗化をもたらすことの機構までは解明されていない

遺伝子重複による相同遺伝子の形成

# Evolutionary fate and functional consequence



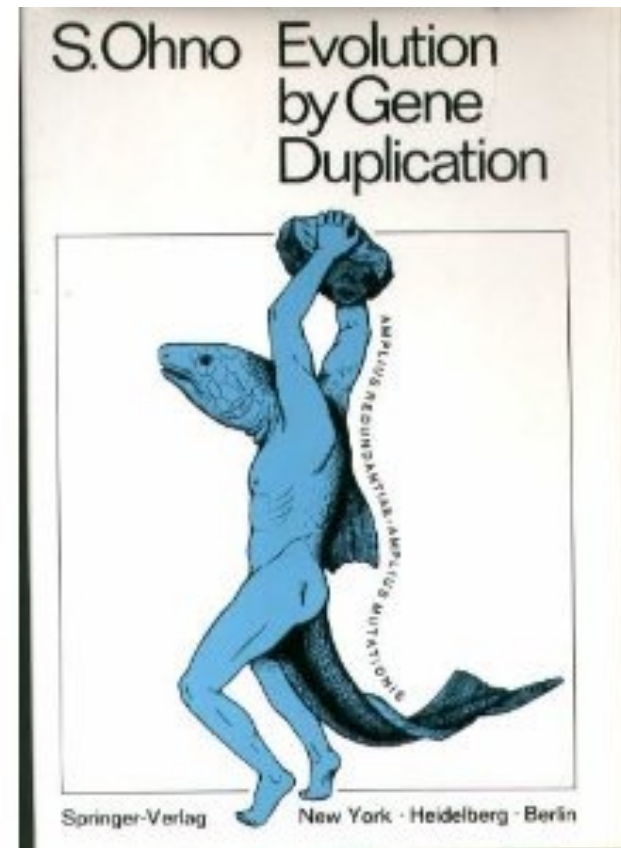
# 遺伝子重複



1つの祖先から、機能の多様化を反映した複数のタンパク質が生成

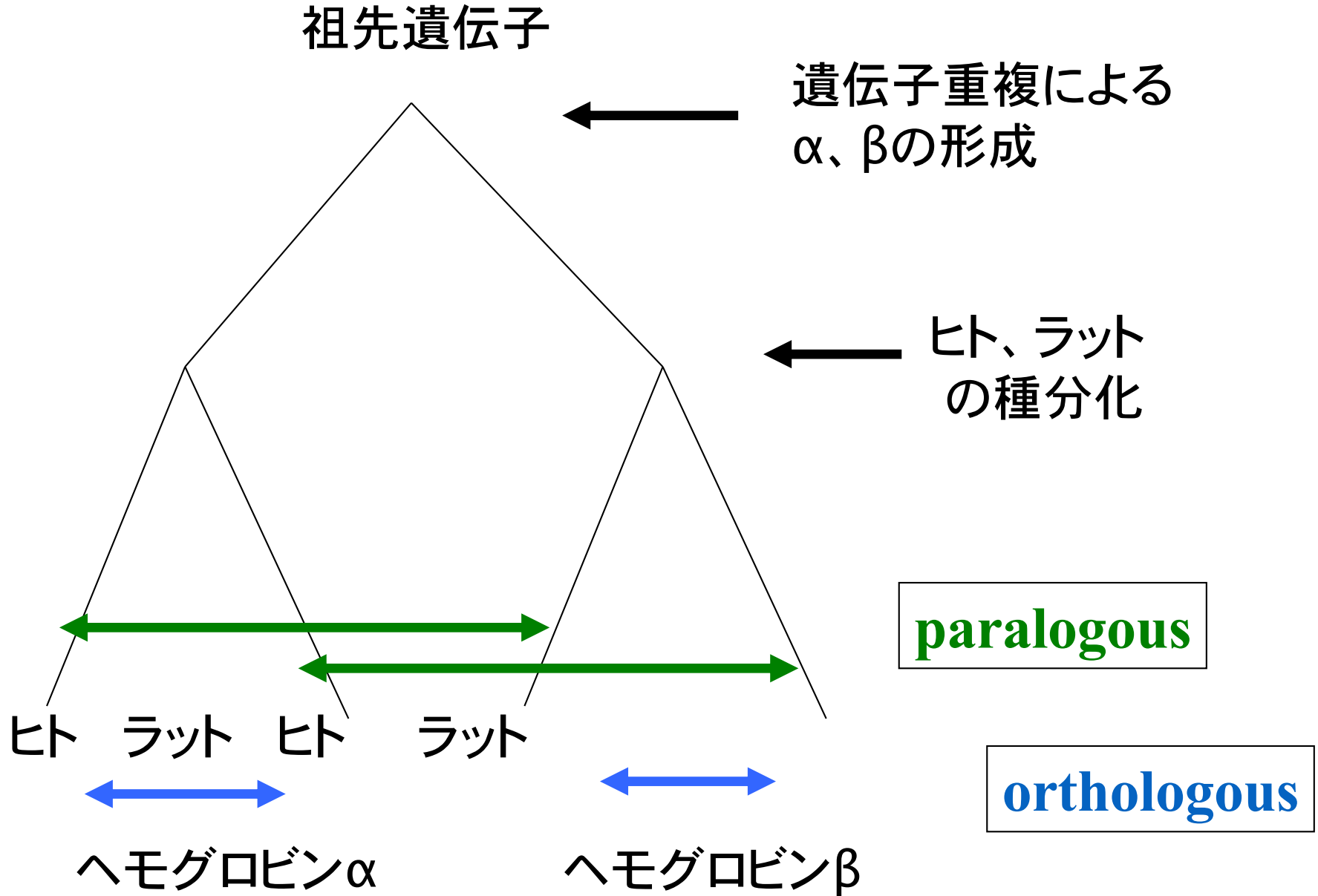
# Susumu Ohno

(1928 - 2000)



## Evolution by Gene Duplication (1970)





相同  
(homologous, homolog)  
共通祖先から分岐

orthologous, ortholog  
種分化に伴い分岐

paralogous, paralog  
遺伝子重複に伴い分岐

# 種系統樹と遺伝子系統樹

分子系統樹

種系統樹 (Species Tree)

オーソロガスな遺伝子のみ使用  
生物の進化的関係を反映

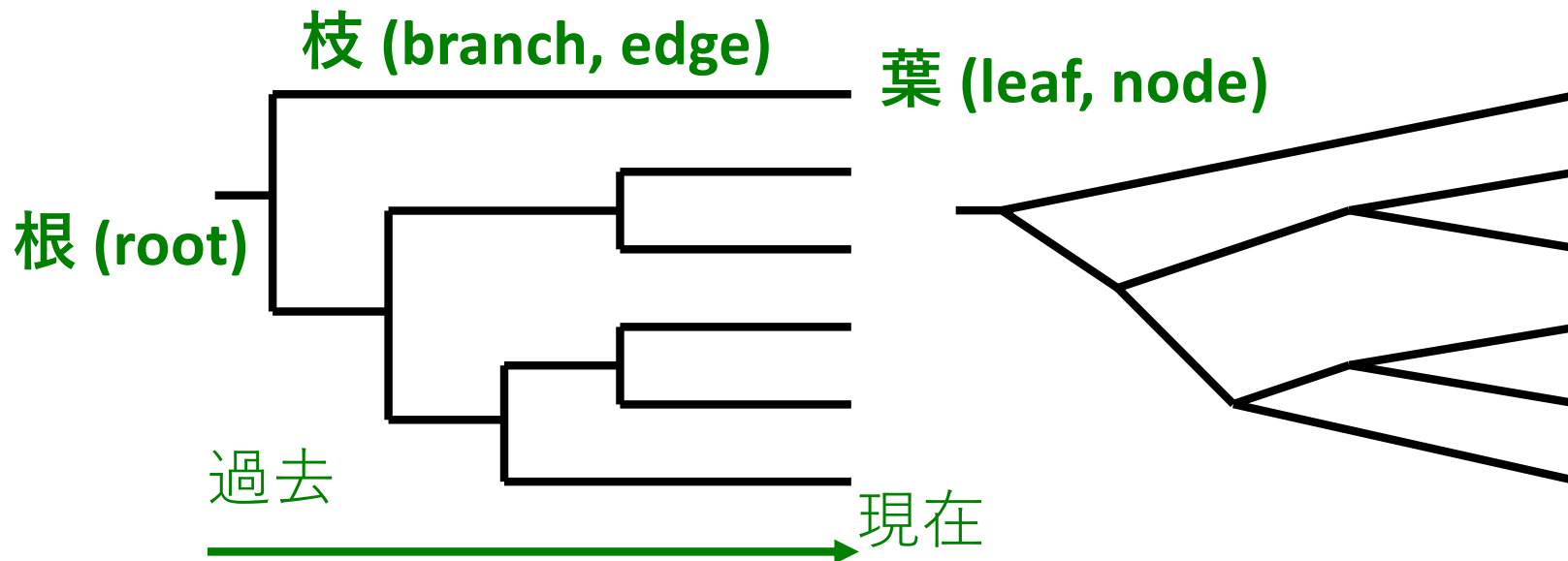
遺伝子系統樹 (Gene Tree)

パラロガスな遺伝子もオーソロガスな遺伝子  
も混在。遺伝子（タンパク質）の分子進化の  
歴史を反映

枝の先（葉）に生物種などが配置される

枝の長さは生物種間の何らかの「形質」の違いを反映した量を表す（形態の差異、遺伝子の有無、遺伝子中の塩基置換など）  
もし、それらの違いが時間と共に増大するならば、枝の長さは近似的に時間に比例

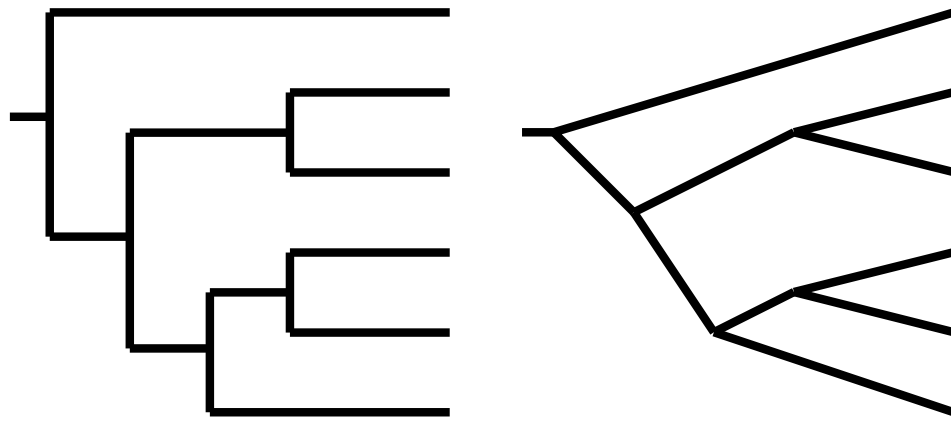
## 有根樹(Rooted tree)の表現例



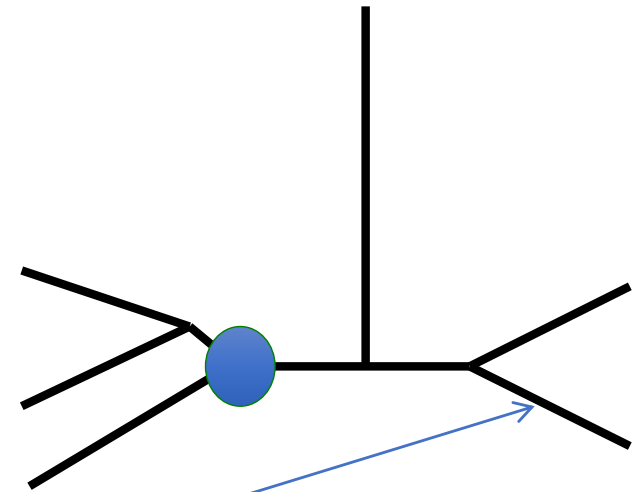
なお、系統樹に含まれる操作単位（生物、遺伝子）を OTU (Operational Taxonomic Unit) とよぶ。

# 有根樹と無根樹

有根樹  
(Rooted tree)



無根樹  
(Unrooted tree)

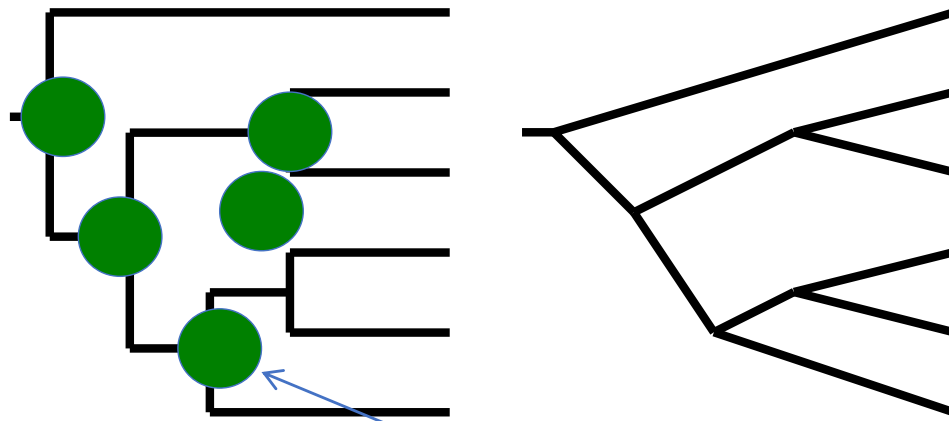


Edge (枝)

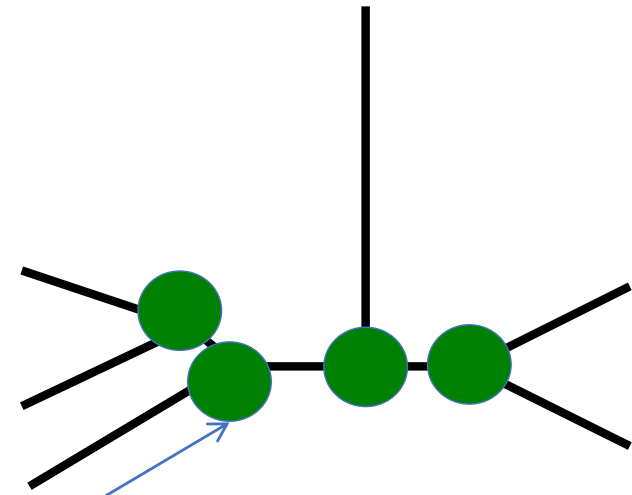
Node 節 (Internal node、内部節)  
(External node、外部節 = leaf = 葉)

# 有根樹と無根樹

有根樹  
(Rooted tree)



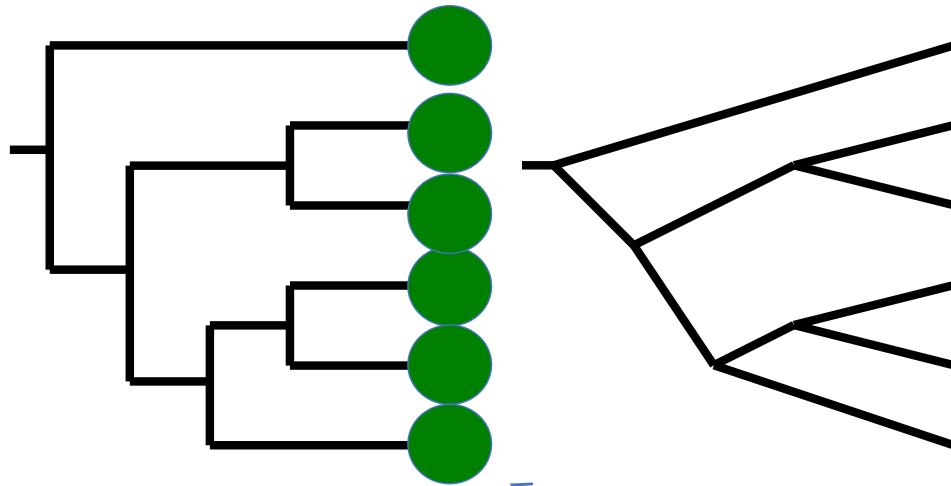
無根樹  
(Unrooted tree)



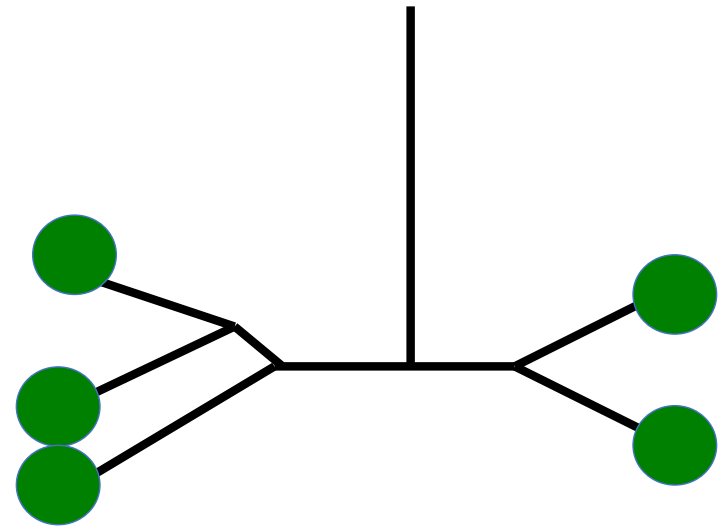
Node 節 (Internal node、内部節)  
(External node、外部節 = leaf = 葉)

# 有根樹と無根樹

有根樹  
(Rooted tree)



無根樹  
(Unrooted tree)



Node 節 (External node、外部節 = leaf = 葉)

# 配列アラインメント (sequence alignment)

進化の過程でのINDELを考慮しながら、相同な配列の間の対応する塩基（あるいはアミノ酸）を同じ位置に並べる操作あるいは、その操作によってできたもの。

INDELに対応してギャップ(gap)とよばれる空記号を挿入し位置をずらして、塩基やアミノ酸を対応づける。

通常、動的計画法(dynamic programming algorithm)や、そのバリエーションを用いて、配列間の類似度が高くとなるようにアラインメントが構築される。



## マルチプルアライメント (multiple alignment)

分子系統樹を構築するには、まず複数本の相同配列についてのマルチプルアライメントを作成する。

# リゾチームのアミノ酸配列

>LYC\_HUMAN **ヒト** Lysozyme C

MKALIVLGLVLLSVTVQGKVFERCELARTLKRLGMDGYRGISLANWMCLAKWESGYNTRATNYNAGDRST  
DYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAKRVVRDPQGIRAWVAWRNRCQNRDVR  
QYVQGC

>LYC1\_BOVIN **ウシ** Lysozyme C 1

MKALIIILGFLFLSVAVQGKVFERCELARTLKRLGMDGYRGISLANWLCCLKWESSYNTKATNYNPGSEST  
DYGIFQINSKWWCNDGKTPNAVDGCHVSCSELMENDIAKAVACAKQIVSEQGITAWVAWKSHCRDHDVSS  
YVEGCTL

>LYC\_CHICK **ニワトリ** Lysozyme C

MRSLLILVLCFLPLAALGKVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTD  
YGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCACKIVSDGNGMNAWVAWRNRCKGTDVQA  
WIRGCRL

>LYC2\_ONCMY **マス** Lysozyme C II

MRAVVVLLLVAVASAKVYDRCELARALKASGMDGYAGNSLPNWVCLSKWESSYNTQATNRNTDGSTDYGI  
FQINSRYWCDDGRTPGAKNVCGIRCSQLLTADLTVAIRCAKRVVLDPNGIGAWVAWRLHCQNQDLRSYVA  
GCGV

>LYC\_BOMMO **カイコ** Lysozyme

MQKLIIFALVVLCVGSEAKTFTRCGLVHELKHKHGFEEENLMRNWVCLVEHESSRDTSKTNTNRNGSKDYGL  
FQINDRYWCSKGASPGKDCNVKCSDLLTDDITKAAKCAKKIYKRHRFDWYGWKNHCQGSPLPDISSC

>LYSP\_DROME **ハエ** Lysozyme P

MKAFLVICALTLTAVATQARTMDRCSLAREMSKLGVPRDQLAKWTCIAQHESFRTGVVGPANSNGSNDY  
GIFQINNKYWCKPADGRFSYNECGLSCNALLTDDITNSVKCARKIQRQQGWTAWSTWKYCSGSLPSINSC  
F

# FASTA フォーマット

>LYC\_HUMAN ヒト Lysozyme C

```
MKALIVLGLVLLSVTVQGKVFERCELARTLKRLGMDGYRGISLANWMCLAKWES  
GYNTRATNYNAGDRSTDYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIA  
DAVACAQRVVRDPQGI RAWVAWRNRCQNRDVRQYVQGC
```

“>” で始まる注釈行（通常、配列名）と  
改行後に、1文字表記で塩基あるいはアミノ酸の配列を記す書式を  
**FASTAフォーマット**といい、分子系統解析ばかりでなく、バイオインフォ  
マティクス分野では配列の書式としてよく使われている。

前ページのように、一つのファイルに複数のFASTAフォーマットの  
配列がおさめられている場合は、**マルチFASTAフォーマット**とよぶ。

# リゾチームのアミノ酸配列 マルチプル・アラインメント

二次構造情報を重ねて表示  $\alpha$  helix  $\beta$  strand

ヒト  
ウシ  
マス  
ニワトリ  
カイコ  
ハエ

CLUSTAL W (1.81) multiple sequence alignment

```

LYC_HUMAN      -MKALIVLGLVLLSVTVQGKVFERCELARTLKRLGMDGYRGISLANWMCLAKWESGYNTR
LYC1_BOVIN     -MKALIIILGFLFLSVAVQGKVFERCELARTLKKLGLDGYKGVSLANWLCLTKWESSYNTK
LYC2_ONCMY     ----MRAVVVLLLVAVASAKVYDRCELARALKASGMDGYAGNSL PNWVCLSKWESSYNTQ
LYC_CHICK      -MRSLLIILVLCFLPLAALGKVFGRCELAAAMKRHGLDNRYRGYSIGNWVCAAKFESNFNTQ
LYC_BOMMO      -MQKLIIFALVVLCVVGSEAKTFTRCGLVHELKHKHGFEEENL---MRNWNWVCLVEHESSRDTS
LYSP_DROME     MKAFLVICALTLTAVATQARTMDRCSLAREMSKLGVPRDQ---LAKWTCIAQHESFRTG

```

: . . . . . \*\* \* . : \* . : \* \* : \*\* . \*

```

LYC_HUMAN      ATNYNAGDRSTDYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAQRVVR
LYC1_BOVIN     ATNYNPGSISTDYGIFQINSKWCNDGKTPNAVDGCHVSCSELMENDIAKAVACAQIIVS
LYC2_ONCMY     ATNRN-TDGSTDYGIFQINSRYWCDDGRTPGAKNVCGIRCSQLLTADLTVAIRCAKRVVL
LYC_CHICK      ATNRN-TDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIIVS
LYC_BOMMO      KTNTN-RNCSKDYGLFQINDRYWCS--KGASPGKDCNVKCSDLLTDDITKAAKCAKKIYK
LYSP_DROME     VVG PANSNGSNDYGIFQINNKYWCKPADGRFSYNECGLSCNALLTDDITNSVKCARKIQR

```

.. . \* .\*\*\*::\*\*\*::\*\*\*. . . \* : \* . \* : ::: : \*\*:::

```

LYC_HUMAN      DPQGIRAWVAWRNRQNIRDVRQYVQGCYV
LYC1_BOVIN     E-QGITAWVAWKSHCRDHDVSSYVEGCTL
LYC2_ONCMY     DPNGIGAWVAWRLHCQNIQDLRSYVAGCGV
LYC_CHICK      DGNGMNAWVAWRNRCKCTDVQAWIRGCRLE
LYC_BOMMO      R-HRFDAWYGWKNHCQGS--IPDISSC--
LYSP_DROME     Q-QGWTAWSTWK-YCSGS--IPSI NSCF-

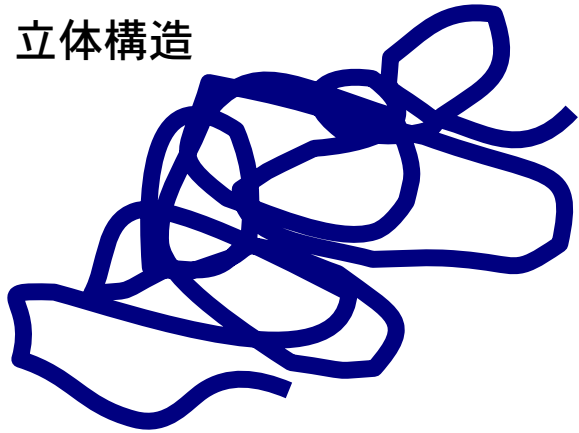
```

: \*\* \* : \* ! \_ \_ | : . \*

アミノ酸配列



立体構造



## ヒト・リゾチームのアミノ酸配列

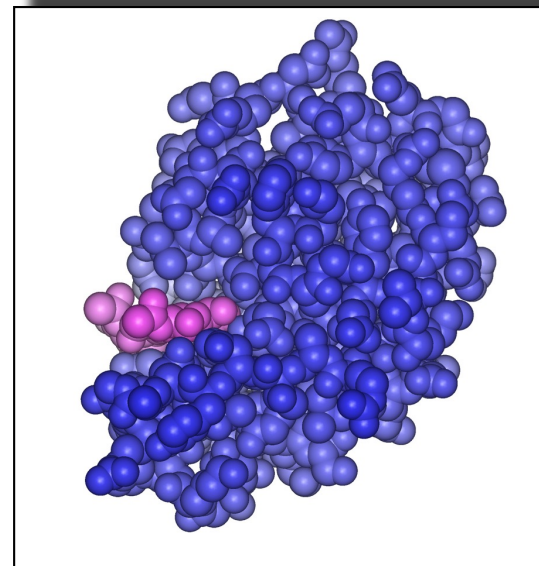
(一文字表記)

>LYC\_HUMAN

MKALIVLGLVLLSVTVQGKVFERCELARTL  
KRLGMDGYRGISLANWMCLAKWESGYNTRA  
TNYNAGDRSTDYGIFQINSRYWCNDGKTPG  
AVNACHLSCSALLQDNIADAVACAKRVVRD  
PQGIRAWVAWRNRCQNRDVRQYVQGCGV

148 残基

## ヒト・リゾチームの立体構造



# 置換や挿入/欠失は、一次構造中不均一に生じる

機能的制約と構造的制約（疎水コア、モチーフ）

モチーフとは、相同配列間で強く保存された短い領域をさす。

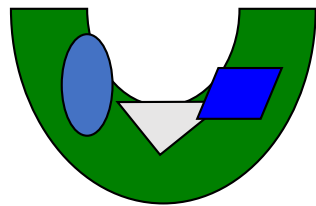
モチーフの保存は、負の選択によって維持される。



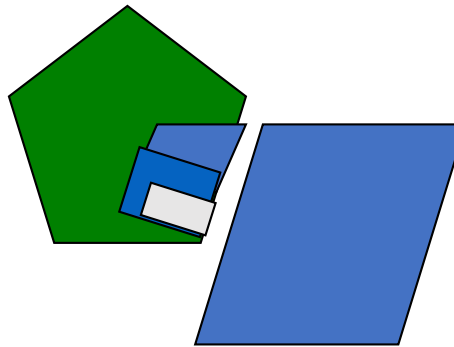
Motif A

Motif B

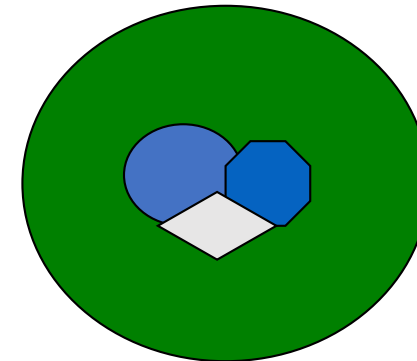
Motif C



酵素の活性中心



多量体の  
会合面



疎水コア

側鎖のパッキングが密なため。表面近傍は逆に置換や挿入/欠失が生じやすい。

保存的な領域は一次構造的には分散していても、三次構造上は近接している場合が多い。

形質

アフリカゾウ



耳  
形質状態  
大きい三角形

牙  
形質状態  
雌雄共に長く発達



アジアゾウ



形質状態  
小さい四角形

形質状態  
短い

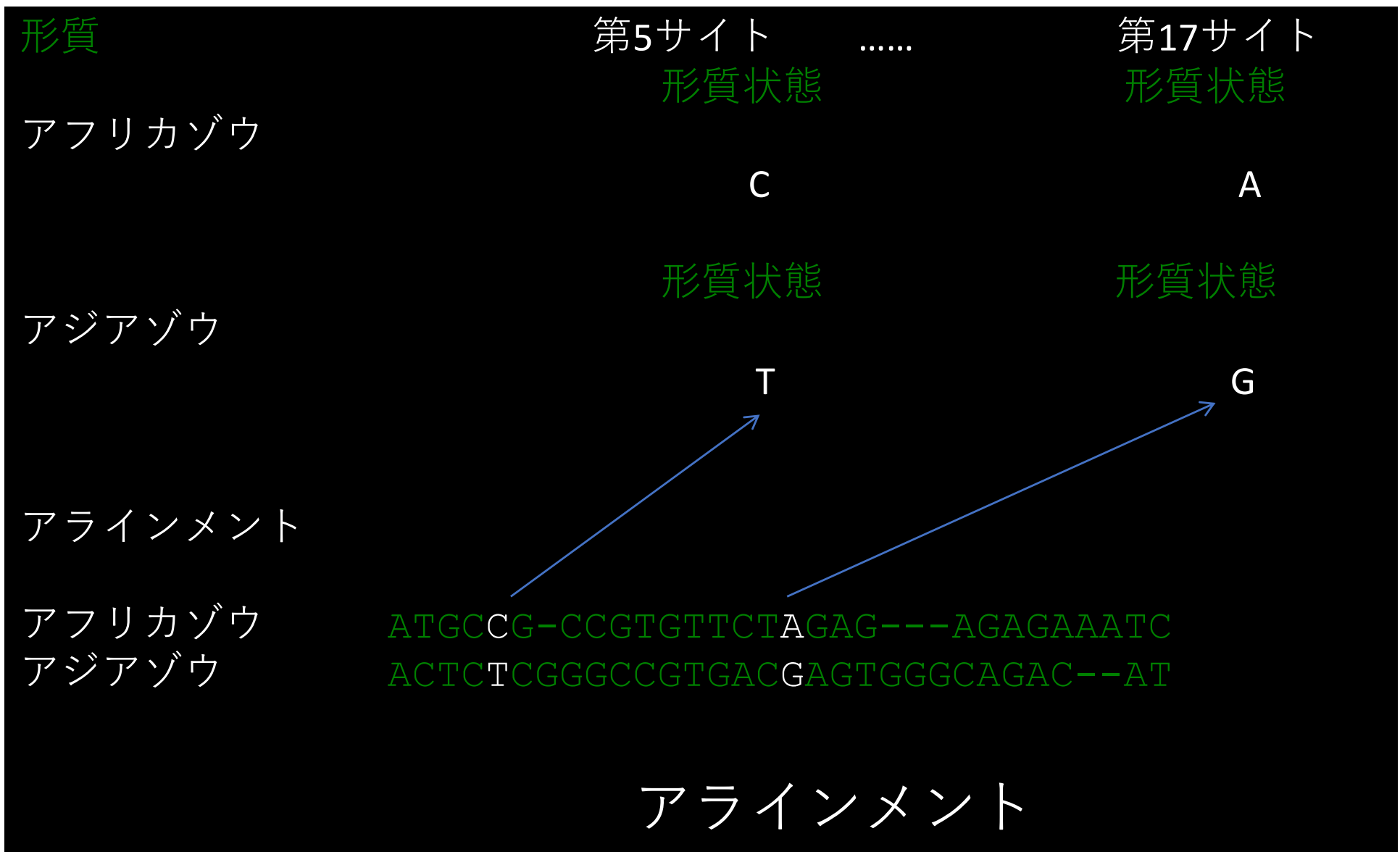


4つの種A～Dについて、5つの形質の形質状態行列

	形質1	形質2	形質3	形質4	形質5
種A	a	b	c	d	e
種B	a+	b	c	d	e
種C	a+	b+	c	d	e
種D	a+	b+	c+	d+	e+

例：形質1は目の色。aは黒、a+は青





前回、数量分類学で説明した形質のコード化が、配列データの場合、アラインメント上の一文字表記で完了しているとみなせる。

# リゾチームのアミノ酸配列 マルチプル・アラインメント

二次構造情報を重ねて表示  $\alpha$  helix   $\beta$  strand 

ヒト  
ウシ  
マス  
ニワトリ  
カウ  
ハエ

CLUSTAL W (1.81) multiple sequence alignment

```

LYC_HUMAN      -MKALIVLGLVLLSVTVQGVFERCELARTLKRLGMDGYRGIISLANWMCLAKWESGYNTR
LYC1_BOVIN     -MKALIIILGFLFSLVAVQGVFERCELARTLKKLGLDGYKGVSLANWLCLTKWESSYNTK
LYC2_ONCMY     ----MRAVVVLLLVAVASAKVYDRCELARALKASGMDGYAGNSL PNWVCLSKWESSYNTQ
LYC_CHICK      -MRSLLILVLCFLPLAALGKVFGRCELAAAMKRHGLDNYRGYSIGNWVCAAKFESNFNTQ
LYC_BOMMO      -MQKLIIFALVVLCVVGSEAKTFTRCGLVHELKHKHGFEEENL---MRNWVCLVEHESSRDTS
LYSP_DROME     MKAFLVICALTLTAVATQARTMDRCSLAREMSKLGVPRDQ---LAKWTCIAQHESSFRTG

```

```

      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *
      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *
      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *
      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *
      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *
      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *
      :      . .      . . .      * * * .      :      * .      :      : * *      :      * * .      *

```

```

LYC_HUMAN      ATNYNAGDPRSTDYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAKRVVR
LYC1_BOVIN     ATNYPGSEISTDYGIFQINSKWCNDGKTPNAVDGCHVSCSELMENDIAKAVACAKQIVS
LYC2_ONCMY     ATNRN-TDQSTDYGIFQINSRYWCDDGRTPGAKNVCGIRCSQLLTADLTVAIRCAKRVVL
LYC_CHICK      ATNRN-TDQSTDYGILQINSRWCNDGRTPGSRNLCNIPCSALLSSDITASVNC AKKIIVS
LYC_BOMMO      KTNNTN-RNCSKDYGLFQINDRYWCS--KGASPGKDCNVKCSDLLTDDITKAAKCAKKIYK
LYSP_DROME     VVGPA NSNGSNDYGIFQINNKYWCKPADGRFSYNECGLSCNALLTDDITNSVKCARKIQR

```

```

LYC_HUMAN      DPQGIRAWVAWRNRCQNRDVRQYVQGCGV
LYC1_BOVIN     E-QGITAWVAWKSHCRDHDVSSYVEGCTL
LYC2_ONCMY     DPNGIGAWVAWRLHCQNRQDLRSYVAGCGV
LYC_CHICK      DGNGMNAWVAWRNRCKCTDVQAWIRGCRL
LYC_BOMMO      R-HRFDAWYGWKNHCQGS--IPDISSC--
LYSP_DROME     Q-QGWTAWSTWK-YCSGS--IPSI NSCF-

```

リゾチームのアラインメントの2段目の一部

<b>LYC_HUMAN</b>	<b>ATNYNAGD</b>
<b>LYC1_BOVIN</b>	<b>ATNYNPGS</b>
<b>LYC2_ONCMY</b>	<b>ATNRN-TD</b>
<b>LYC_CHICK</b>	<b>ATNRN-TD</b>
<b>LYC_BOMMO</b>	<b>KTNTN-RN</b>
<b>LYSP_DROME</b>	<b>VVGPANSN</b>

リゾチームのアラインメントの2段目の一部

	61	62	63	64	65	66	67	68	
ヒト		A	T	N	Y	N	A	G	D
ウシ		A	T	N	Y	N	P	G	S
マス		A	T	N	R	N	-	T	D
ニワトリ	A	T	N	R	N	-	T	D	
カイコ		K	T	N	T	N	-	R	N
ハエ		V	V	G	P	A	N	S	N

形質はアラインメントのサイト番号

形質状態はそのサイトにおけるアミノ酸

# 主な分子系統樹推定法

## (1) 距離行列法

- 近隣結合法 (Neighbor Joining Method)

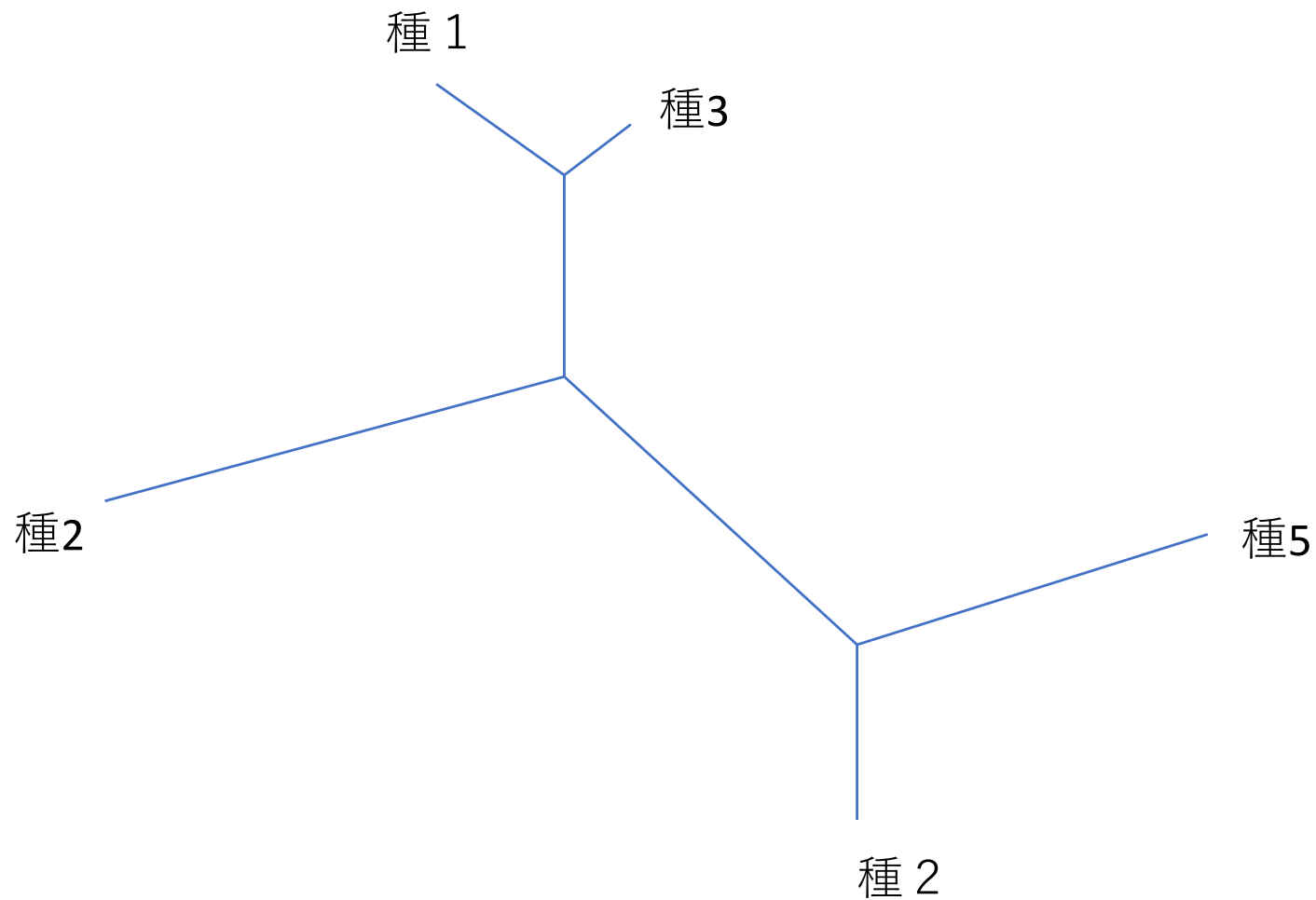
## (2) 形質状態法

- 最節約法 (Maximum Parsimony Method)

- 統計的方法

- 最尤法 (Maximum Likelihood Method)

- ベイズ法



無根系統樹がえられる

近隣結合法では進化速度の一定性は仮定されていないので、  
一つの内部節から分岐した外部節への枝の長さが異なることに注意

分子系統解析では、系統樹は無根系統樹として作成される。  
根の導入は、外群(outgroup)を用いてなされる



進化速度の一定性が成立する場合 (=分子時計が成立している場合) を除き、注目するグループの根を決められない

多くの場合、分子時計の一定性は仮定できない

## 外群の利用

研究対象である群：内群 (ingroup)

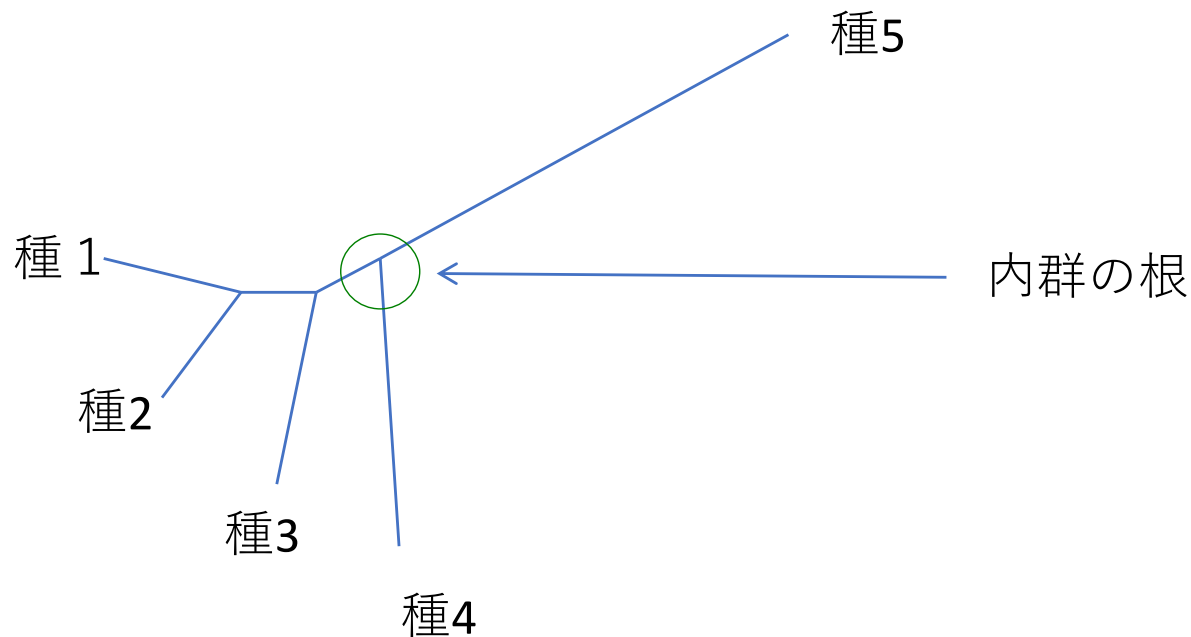
内群に対して遠い関係であることがわかっているもの：外群(outgroup),

外群を含めて系統樹を構築

全体としては無根系統樹だが、内群の根(root)を決めることができる。

今、種1~4（哺乳類）の系統関係を調べたい

根を導入するために種5(爬虫類)を外群として導入





# 最尤法・ベイズ推定プログラム

PHYLIP (<http://evolution.genetics.washington.edu/phyli.html>)

最尤法、最節約法、距離行列法など

PAUP\* (<http://paup.csit.fsu.edu>)

最尤法、最節約法、距離行列法など

Molphy (<http://bioweb.pasteur.fr/seqanal/interfaces/prot.nucml.html>)

最尤法

RAxML (<http://sco.h-its.org/exelixis/web/software/raxml/index.html>)

最尤法

MEGA (<http://www.megasoftware.net>)

最尤法、距離行列法

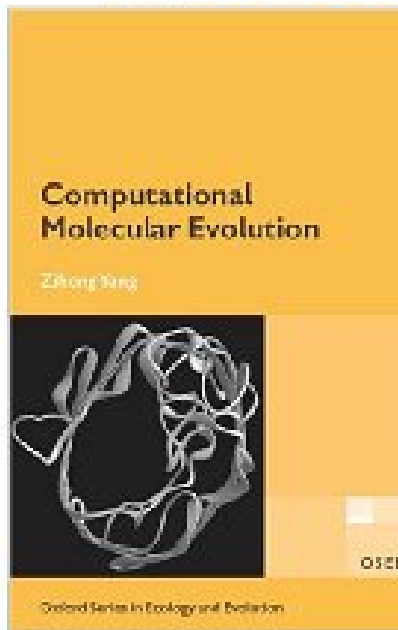
MrBayes (<http:// mrbayes.csit.fsu.edu/index.php>)

ベイズ推定を導入した系統樹推定

BEAST (<https://beast.community>)

ベイズ推定を導入した系統樹推定

分子系統解析の原理をもっと詳しく勉強したい人は次の本を読んでください。また秋学期の分子進化学でも説明します。



Oxford University Press



共立出版

# 3 ドメイン説

- (1) 五界説以前
- (2) 五界説
- (3) 3ドメイン説

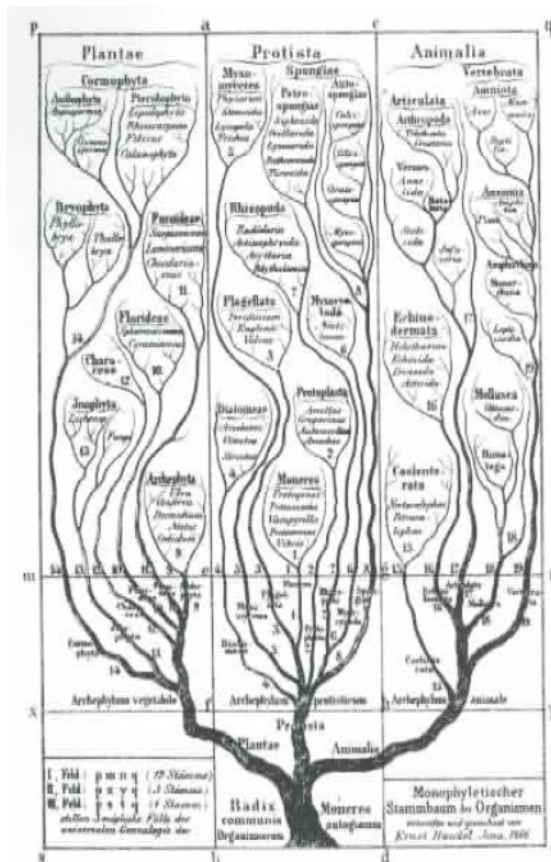
- (1) 五界説以前
- (2) 五界説
- (3) 3ドメイン説

～18世紀

アリストテレス（BC4世紀）からリンネの時代まで生物は動物界と植物界に分類されていた（二界説）

19世紀

ヘッケル（Ernst Haeckel, 1834-1919）は、ダーウィンの進化論を取り入れ、生物を一つの根をもつ系統樹で表現した。また、動物にも植物にも分類されない生物を原生生物界として、三界説を提案した。



その他にも古典的な二界説に代わる分類体系はいくつかも提案されたが、一般に使用されるほど定着するものは、ホイットカーの五界説をまたなければならなかった。

伊藤元己 (2012) 遺伝 66(1), 2-7.

藤田俊彦(2010) “動物の系統分類と進化” 裳華房 より

- (1) 五界説以前
- (2) 五界説
- (3) 3ドメイン説

# ロバート・ホイットカー

Robert Harding Whittaker, 1920-1980

- 細胞構造を基準として、**原核生物**と**真核生物**を分ける
- 原核生物は全体で一つの界（原核生物界 or モネラ界）
- **真核生物を4つの界に分類**

原生生物界 （プロチスタ界）

単細胞

動物界

植物界

菌界



多細胞

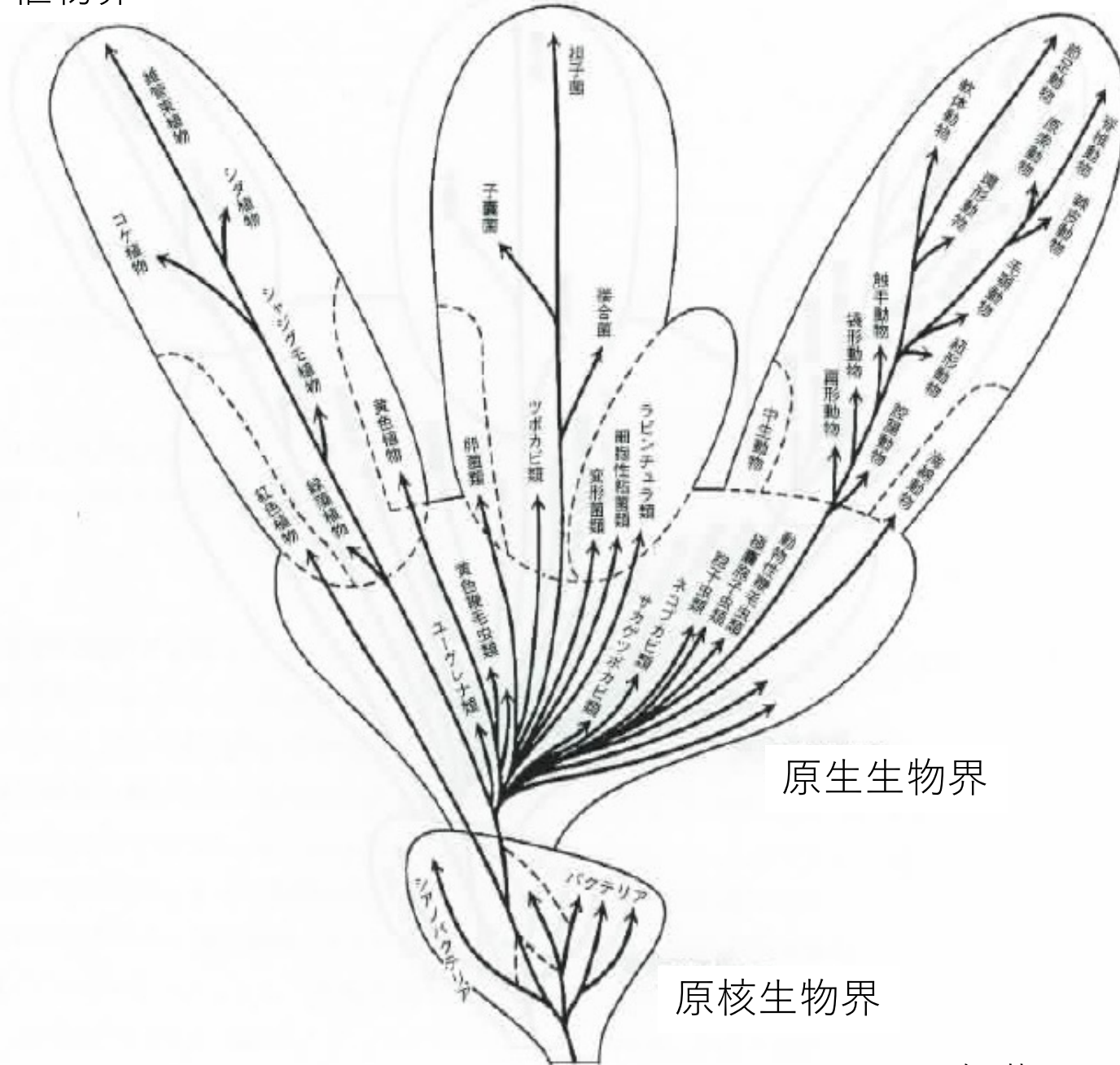
- 多細胞生物の3界はおおざっぱに、  
消費者、生産者、分解者に対応  
(ホイットカーが植物生態学者であることを反映)



植物界

菌界

動物界



原生生物界

原核生物界

伊藤元己 (2012) 遺伝 66(1), 2-7.

# ホイタッカーの五界説

藻類のほとんどを植物に分類

単細胞藻類を植物にいれるか、原生生物にいれるかは  
今の議論が分かれている

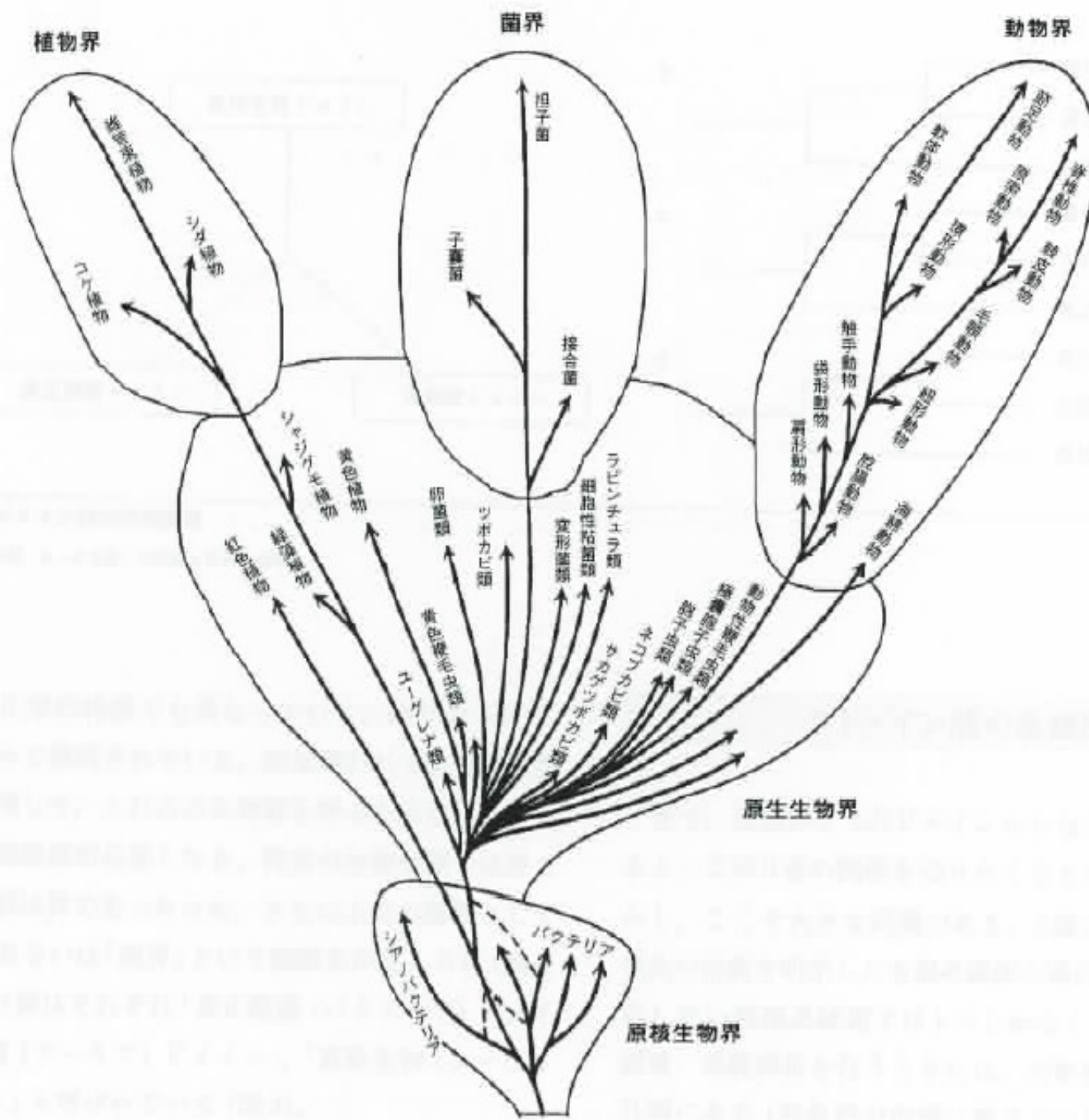
# マーギュリスとシュワルツによる五界説の修正

(Lynn Margulis & Karlene V. Schwartz)

※ マーギュリスは細胞内共生説の提唱者

植物界を陸上植物に限定し、藻類は原生生物界に分類

※ 原生動物の系統関係は現在も研究されている



伊藤元己 (2012) 遺伝 66(1), 2-7.

図2 マーグリス-シュワルツの五界説

図1を元にして、界の区切りをマーグリス-シュワルツの説 (Margulis & Schwartz, 1997) に基づいて変更した

その後、ウーズの古細菌の発見(後述)に伴う六界説(原核生物を真正細菌と古細菌に分割)、キャバリエ-スミスによる八界説や修正六界説(いずれも原生生物の分類の見直し)が出されたが、現在真核生物についてはは2005年の国際原生生物学会で発表された新しい分類体系が使用されている。

真核生物は界に相当する **6つのスーパーグループ** に分類される

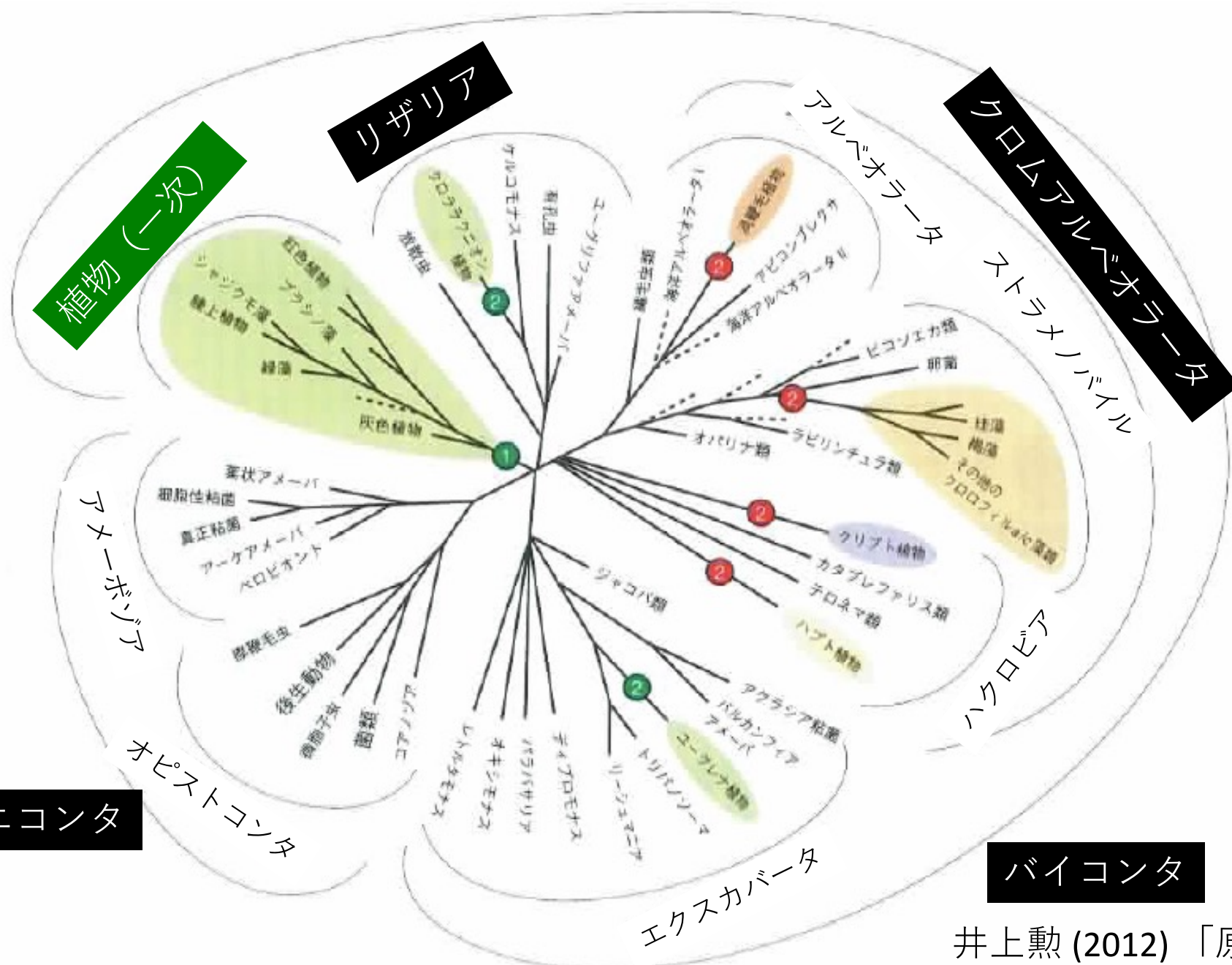
## ユニコンタ

**オピストコンタ** 動物、真菌類を含む  
**アメーボゾア** アメーバ、粘菌など

## バイコンタ

**アーケプラスチダ** 陸上植物、緑藻、紅色植物、灰色植物  
**エクスカバータ** ミドリムシ、トリパノゾーマなどが含まれる  
**リザリア** 有孔虫、放散虫など  
**クロムアルベオラータ** アルベオラータ：アピコンプレクサ(マラリア原虫など)、渦鞭毛虫、  
繊毛虫(ゾウリムシ、ツリガネムシ)など  
ストラメノバイル：不等毛藻(コンブ、ワカメ、珪藻など)  
ハプト植物  
クリプト植物

※ オピストコンタは、動物界と菌界を含むことから、**上界**とされる  
また、**6つのスーパーグループ**の関係について、さらなる細分化、あるいは統合を示唆する研究も多くある。詳細は原生生物の回で

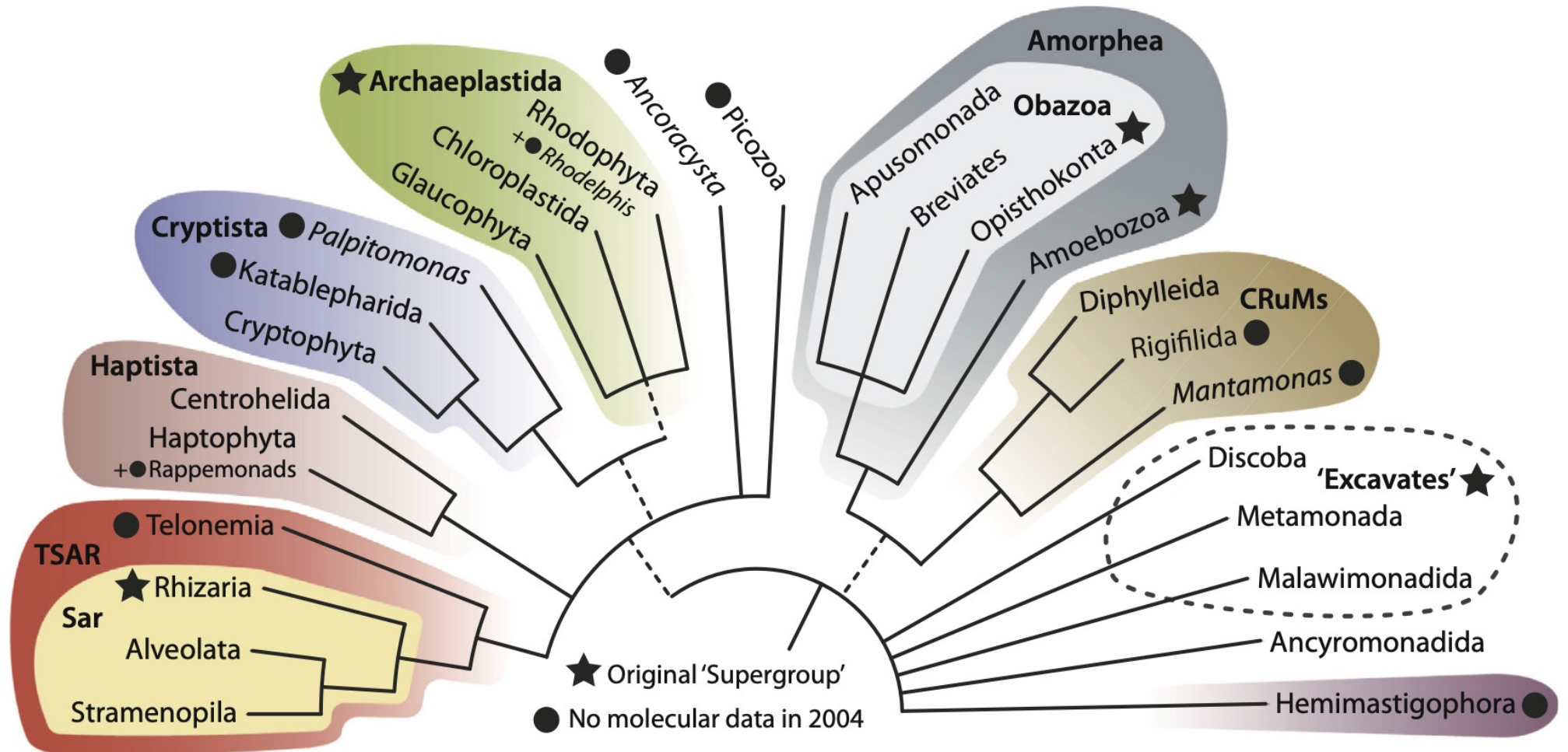


井上勲 (2012) 「原生動物」  
 遺伝 66, p439-444

# 真核生物の分子系統

ただし、紅色植物の二次共生はクロムアルベオラータの根元で起こったとある

Burki *et al.* (2020) "The New Tree of Eukaryotes" *Trends in Ecology and Evolution* **35**, 43-55.



- (1) 五界説以前
- (2) 五界説
- (3) 3ドメイン説

# カール・ウーズ

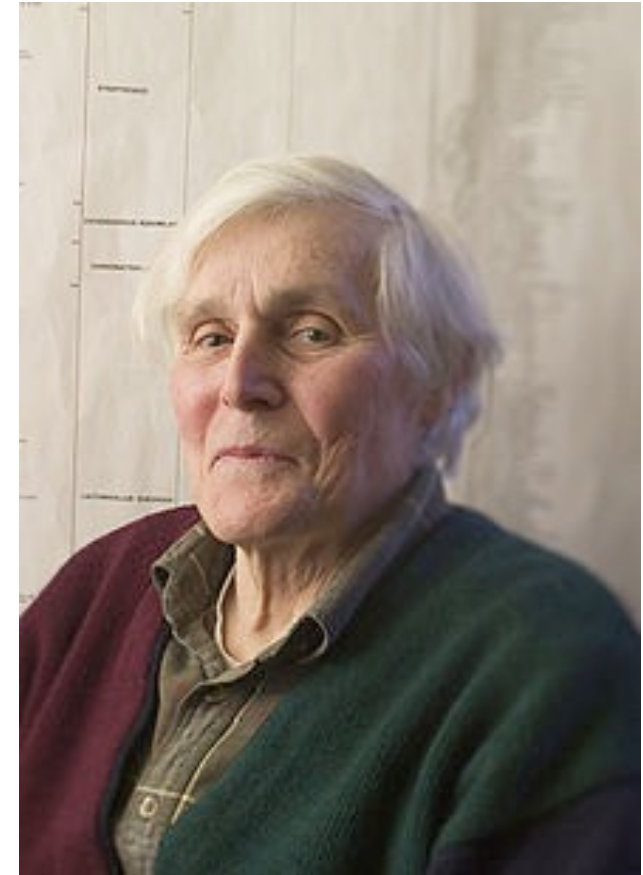
(Carl Recharad Woese, 1928-2012)

rRNAを用いた分子分類や細胞の違いなどから、真正細菌とも真核生物とも異なる古細菌の存在を示唆し、3ドメイン体系を提唱した。

エルンスト・マイヤーらから激しい反発をうける

「こんな分類体系を思いつくのは生物学者ではない証拠だ」

受容には時間がかかった。





# Phylogenetic Tree of Life

真正細菌

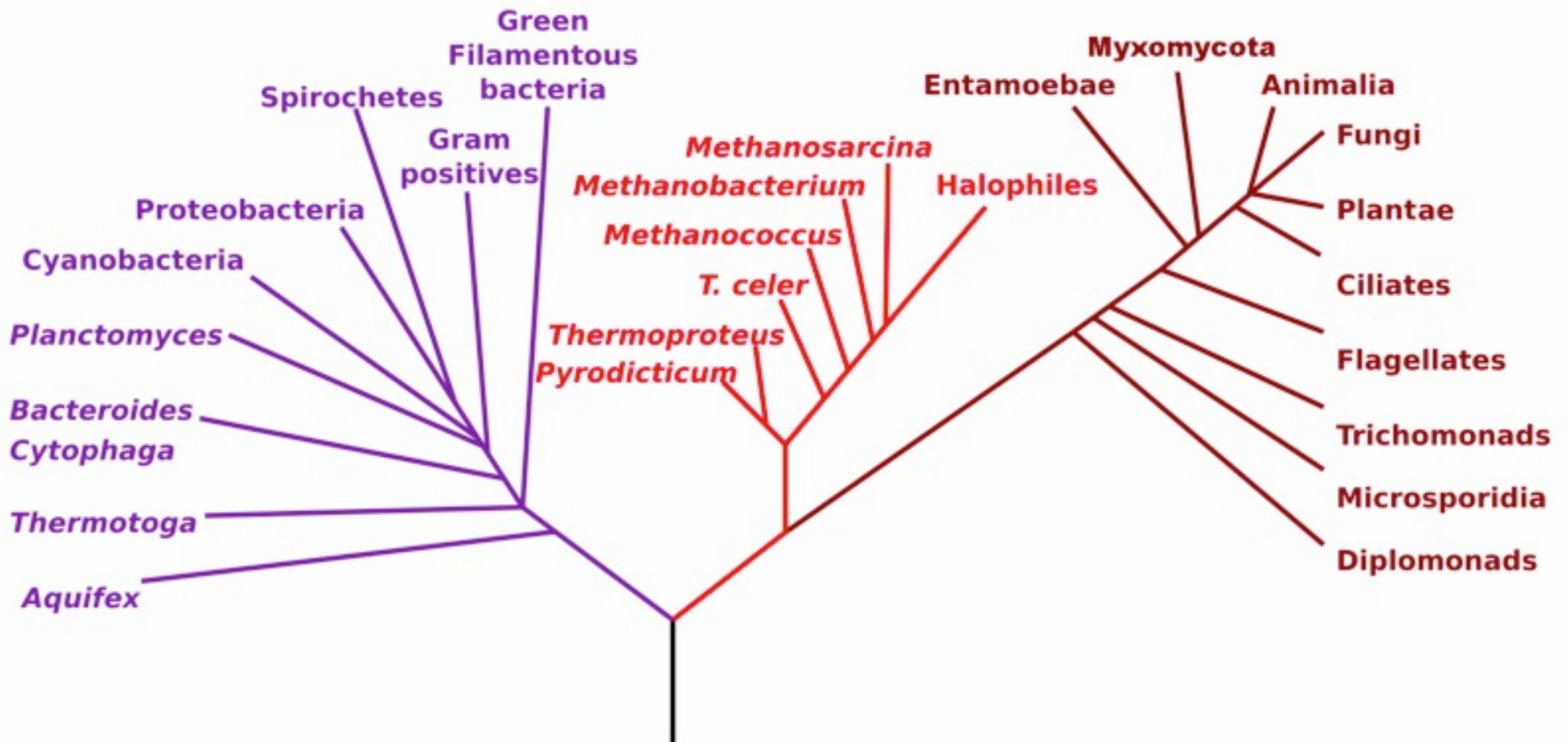
古細菌

真核生物

**Bacteria**

**Archaea**

**Eucarya**



<http://ja.wikipedia.org/wiki/ファイル:PhylogeneticTree.png>

古細菌も真正細菌も、原核細胞を持った原核生物  
残りは真核生物

従来あった動物界や植物界などを残して、これらの分類群を呼ぶために、上位の階級名として  
ドメイン (domain, 超界) がつくられた。

- 真正細菌 (バクテリア) ドメイン
- 古細菌 (アーキア) ドメイン
- 真核生物 (ユーカリア) ドメイン

従来の「界」はそのまま使用される・

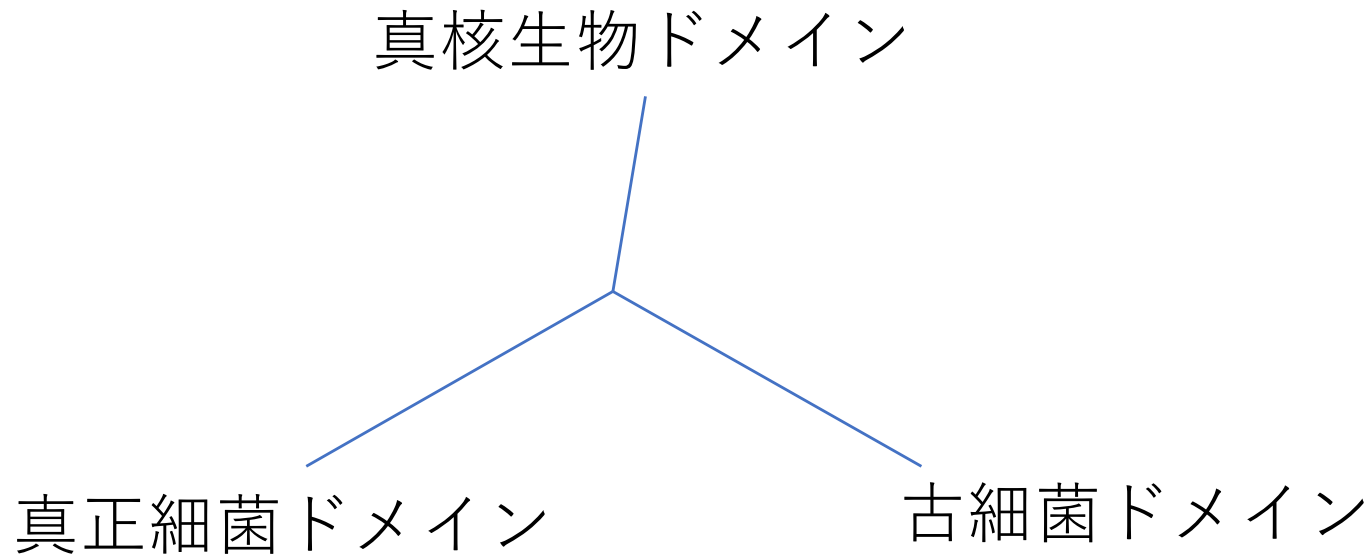
# 3つのドメインの間関係

3者についての無根系統樹

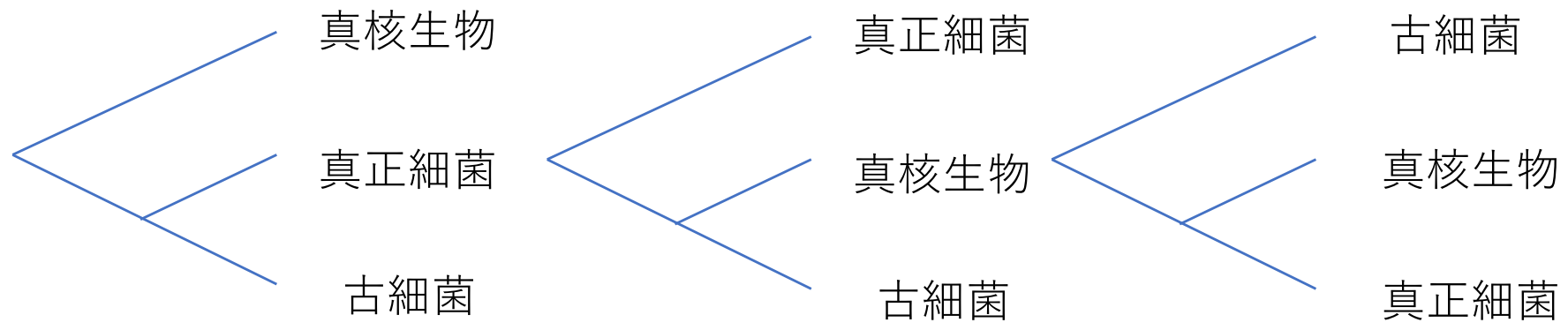
3通りの関係が考えられる。

通常は、外群を導入して「根 (= 共通祖先)」  
を決定

しかし、3者で地球上の生物は尽きており、  
これよりも遠い関係の生物はいない

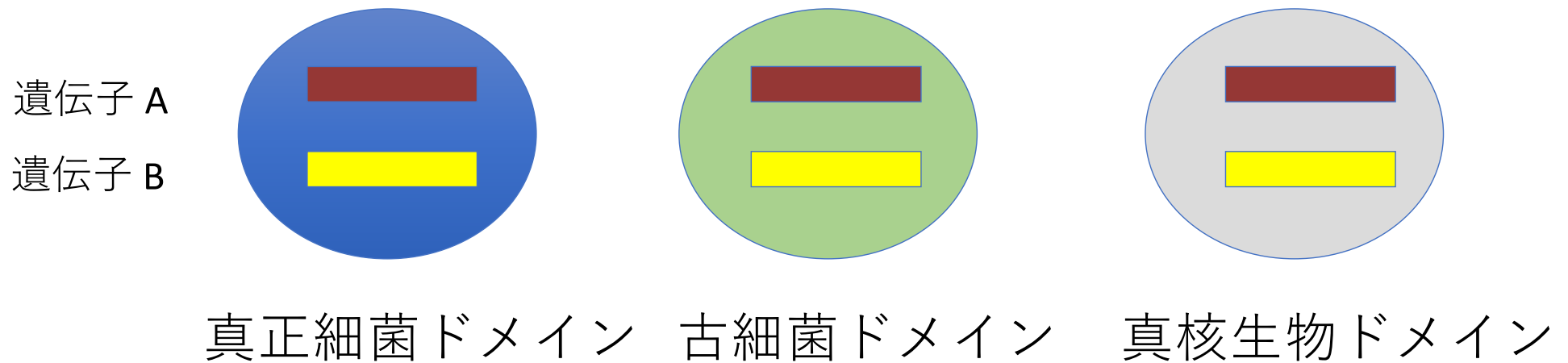


三者の関係の3つの仮説



遺伝子データとして、3つのドメイン全てにコードされたパラログを有する遺伝子を考える

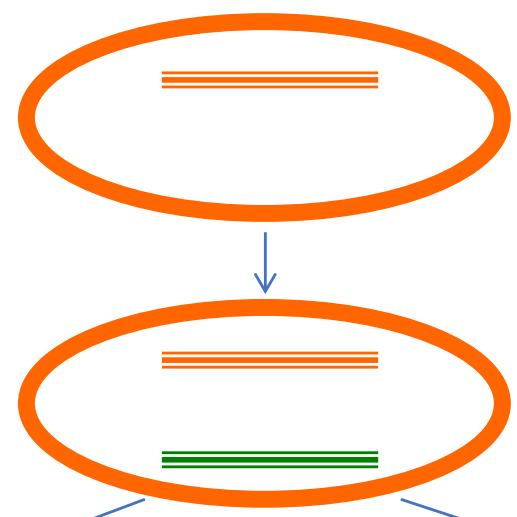
※ ミトコンドリアや葉緑体に由来するものを外す  
(細胞内共生によるものは真正細菌に近くなる)



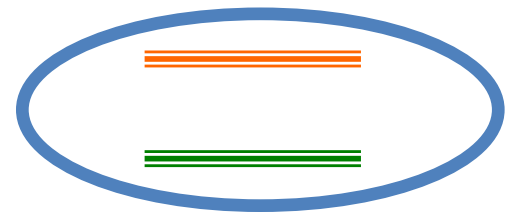
全てのドメインにパラログ (遺伝子Aと遺伝子B)があるということは、遺伝子A, Bはこれら3つのドメインが種分化する以前の祖先生物において遺伝子重複により生じたと考えられる。

遺伝子Aと遺伝子Bをお互いに外群として利用できる。

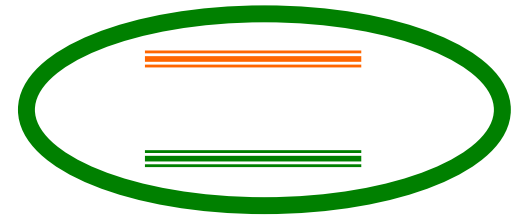
共通祖先で遺伝子重複した  
遺伝子の遺伝子系統樹を  
作成。互いを外群として  
利用することで3ドメインの  
根を決める



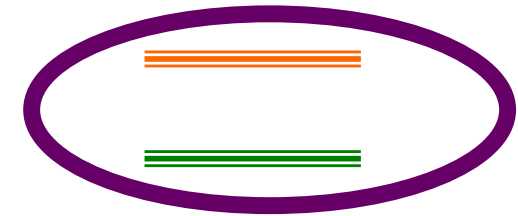
共通祖先での遺伝子重複  
古細菌、真正細菌、真核生物  
の分岐以前



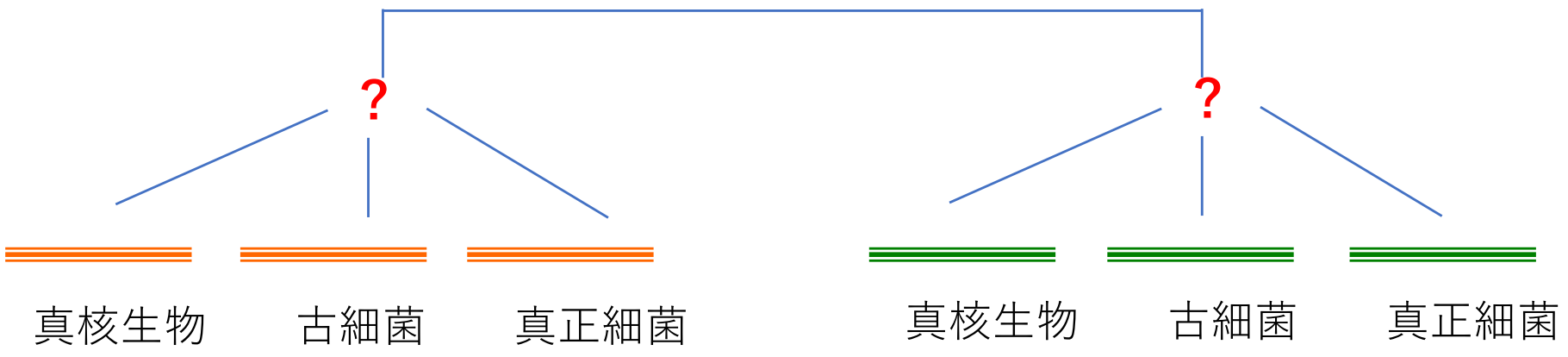
真核生物

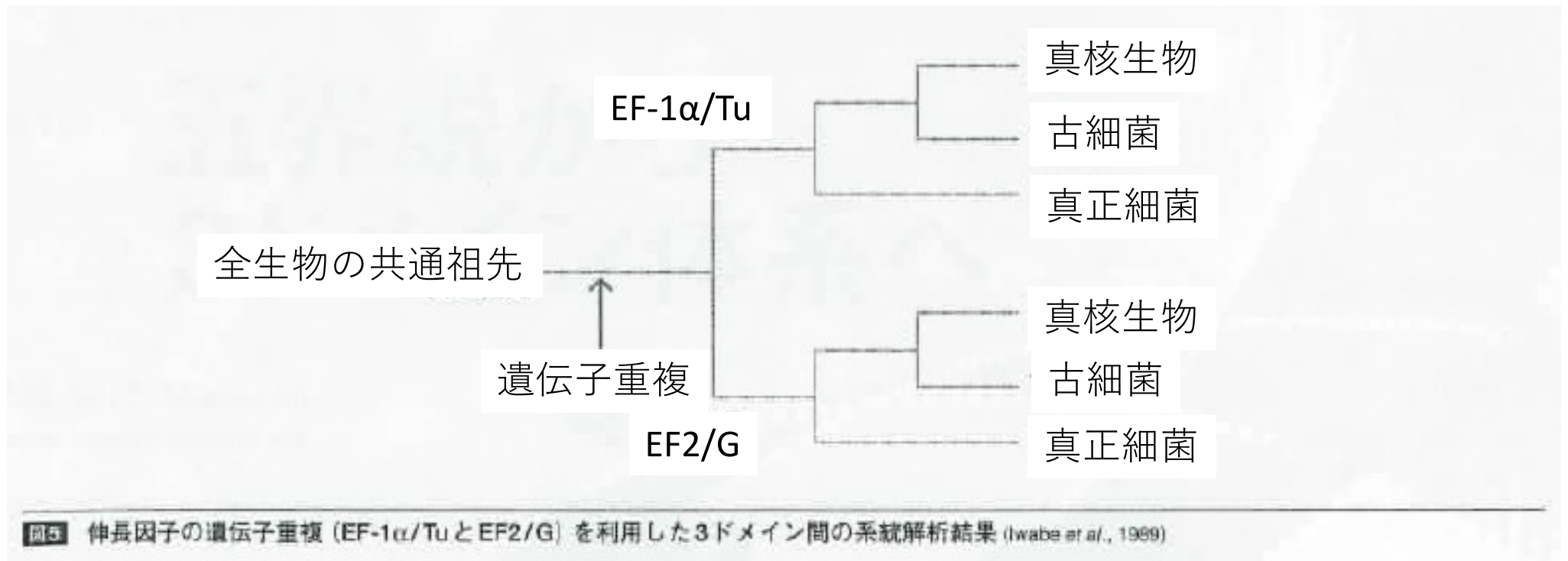


古細菌



真正細菌





外群としてパラログを利用して3つのドメインの根を決定したのは  
宮田隆のグループの研究 (Iwabe et al. 1989)

表 5.1 生物の分類体系の変遷

三界説 ハッケル 1866年	四界説 コーブランド 1956年	五界説 ホイタッカー 1969年	六界説 ウーズ・ フォックス 1977年	三ドメイン ウーズら 1990年	構成する生物群
原生生物界	モネラ界	モネラ界	真正細菌界	真正細菌ドメイン	細菌
			古細菌界	古細菌ドメイン	メタン生成細菌, 好熱好酸菌
植物界	原生生物界	原生生物界	原生生物界	真核生物ドメイン	藻類, 原生動物, 変形菌類
					菌界
植物界	植物界	植物界	植物界		コケ類, シダ類, 種子植物
動物界	動物界	動物界	動物界		無脊椎動物, 脊椎動物

藤田俊彦(2010) “動物の系統分類と進化” 裳華房 より



三ドメイン説  
からの逸脱

- (1) ニドメイン説
- (2) 巨大ウイルス 四ドメイン説
- (3) Rhizome of Life  
- Tree of Lifeからの逸脱 -



三ドメイン説  
からの逸脱

(1) ニドメイン説

(2) 巨大ウイルス 四ドメイン説

(3) Rhizome of Life

- Tree of Lifeからの逸脱 -



## 3 ドメイン説から2ドメイン説へ

ドメイン (domain) = 超界  
界より高次の分類体系  
superkingdom, empire とよばれていたこともある

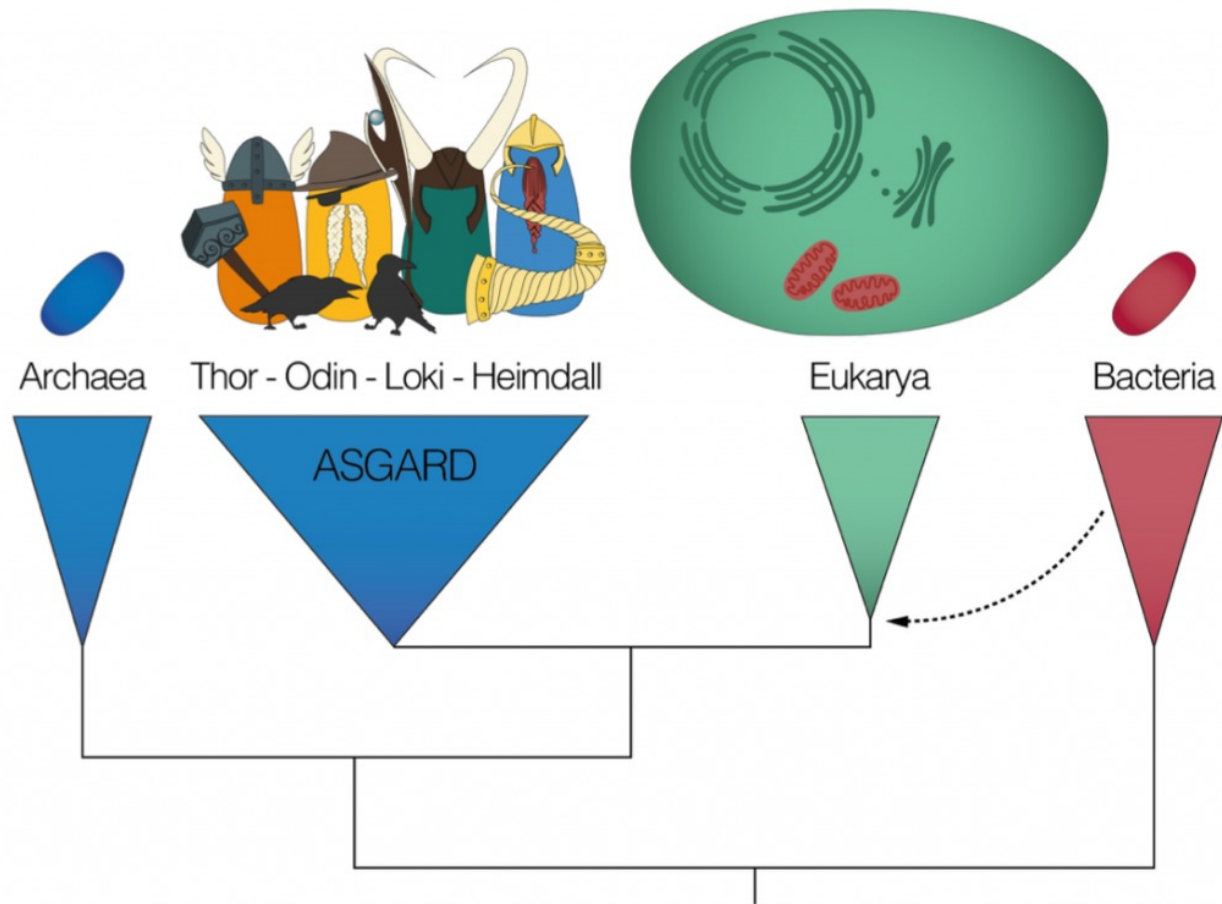
エオサイト説= 2ドメイン説

2ドメイン= 古細菌と真正細菌

※ 真核生物は古細菌に由来

Loki古細菌が見つかったロキ熱水噴出孔をはじめ、イエローストーン国立公園、竹富島の熱水噴出孔まで、世界中から集めた水中堆積物に含まれる全てのDNAの配列をメタゲノム解析により解読  
真核生物特異的と考えられていた多くの遺伝子を持ち、DNA配列上でもLoki古細菌と真核生物を埋める新しい古細菌を発見

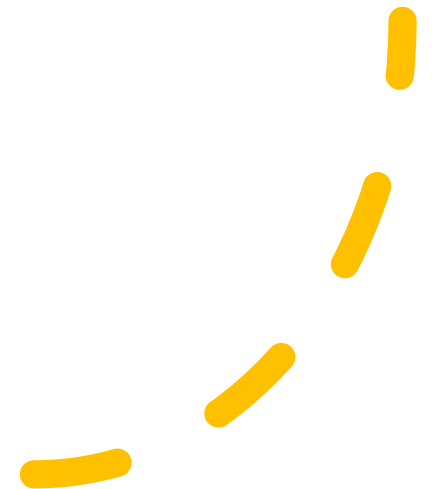
Loki, Thorともに北欧神話の神々の名前で、今回新たに発見された2種類もOrdinとHeimdallという北欧神話の名前がつけられた。そしてLoki,Thorを合わせた群を北欧の神々が棲む世界Asgardと命名



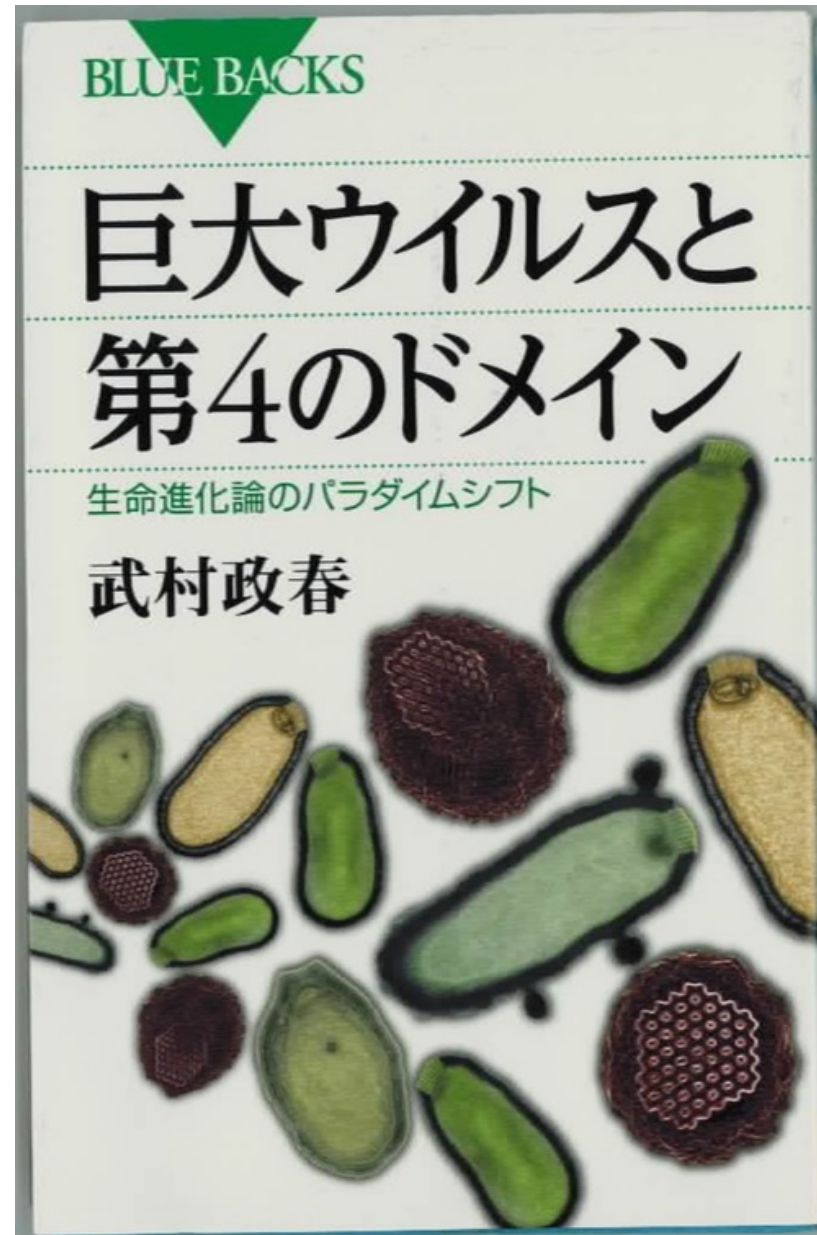
詳細は  
原核生物の回で

三ドメイン説  
からの逸脱

- (1) ニドメイン説
- (2) 巨大ウイルス 四ドメイン説
- (3) Rhizome of Life  
- Tree of Lifeからの逸脱 -



近年巨大ウイルスが次々に発見されてきており、それらを第四のドメインととらえる考え方もでてきている



三ドメイン説  
からの逸脱

- (1) ニドメイン説
- (2) 巨大ウイルス 四ドメイン説
- (3) **Rhizome of Life**  
- Tree of Lifeからの逸脱 -

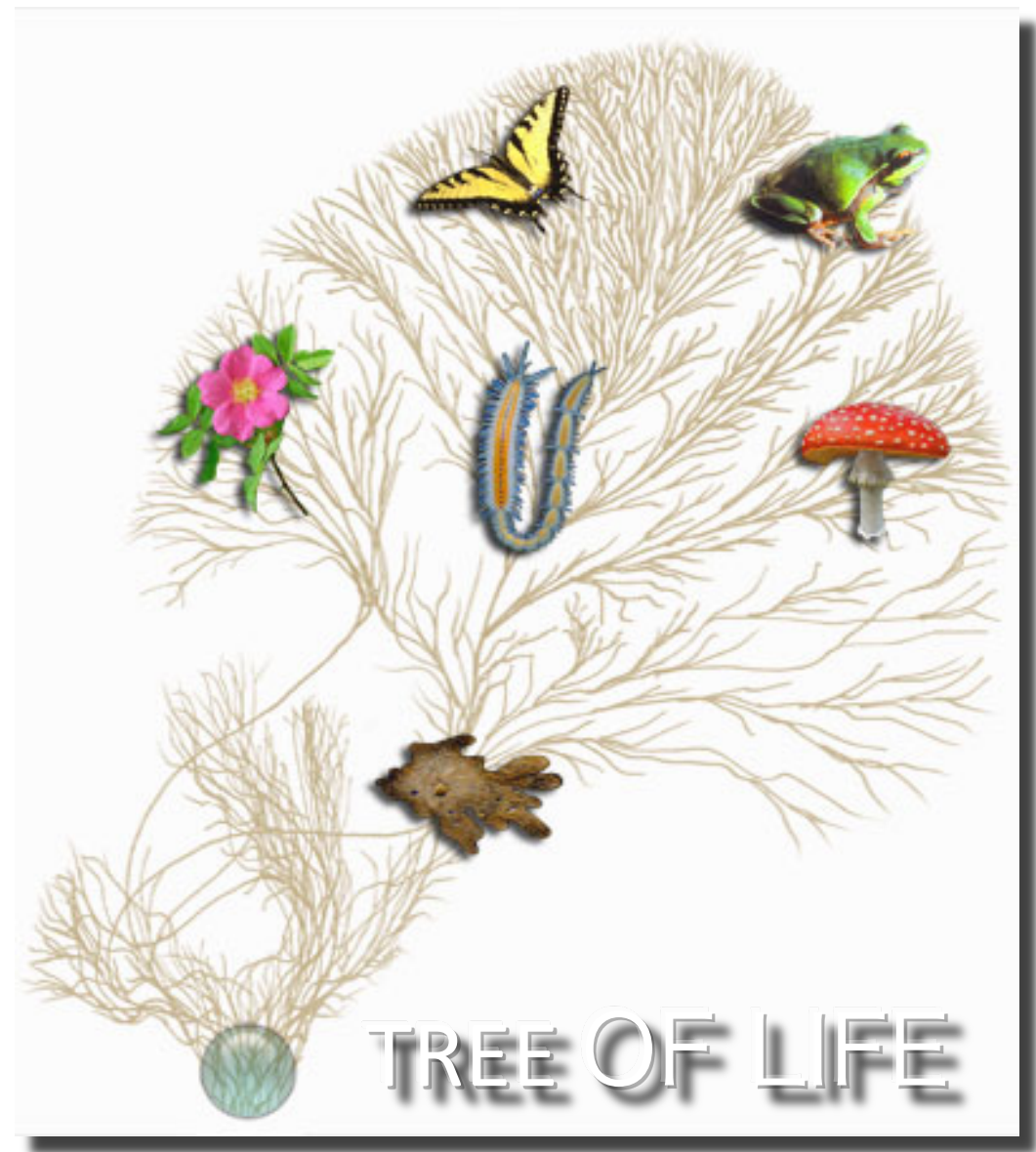




# 生命の木 Tree of Life

共通祖先からの分岐の  
イメージを伝えるための  
比喩としての「木」

ダーウィンのノート中に  
記述されていたもの

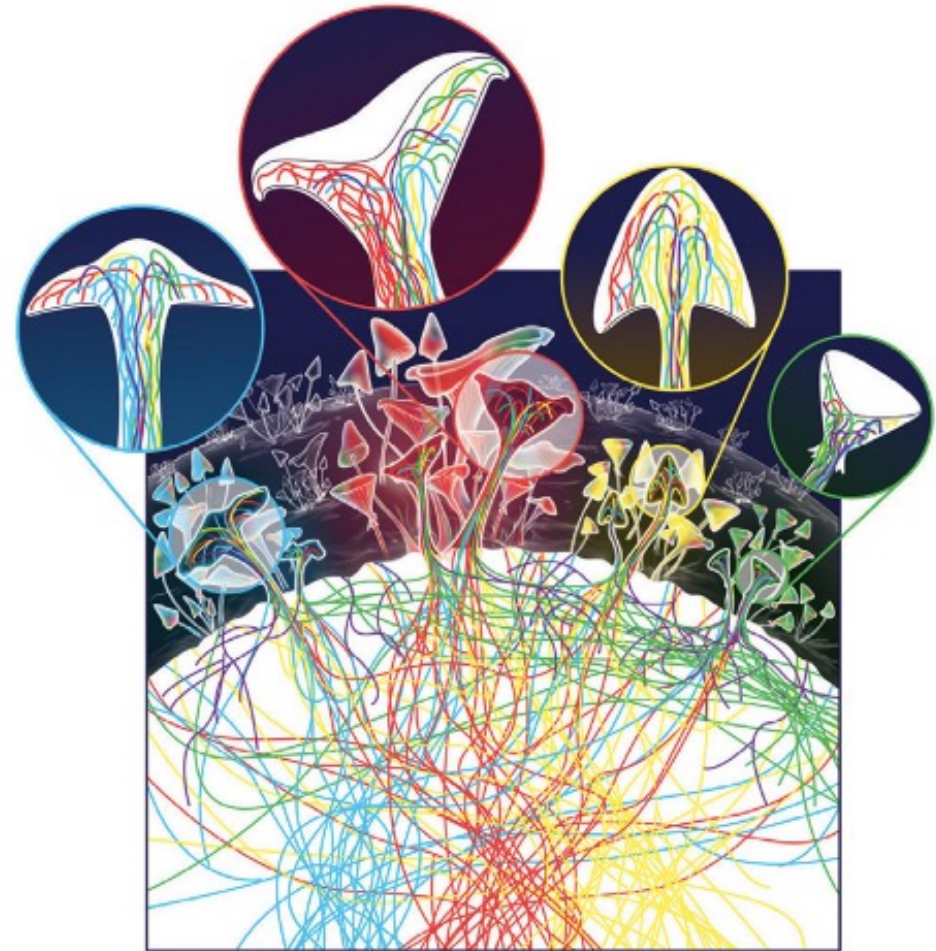


## The post-Darwinist rhizome of life

Raoult D. (2010) *Lancet* **375**, 104-105

### Rhizome of life

Roots of genes of living species are shown according to current classification of organisms: eukaryotes (red), bacteria (blue), viruses (green), archaea (yellow). In purple are genes without identified origin (ORFans). At the surface, in form of mushrooms, are the current species, containing mixture of genes. Colour of mushroom envelope is determined by origin of core genome of the species.



**リゾーム**は、地下茎を意味する**哲学用語**。ツリー（木）構造と対比して語られる。ツリー（木）が、ひとつの根を基礎とし、太い幹に支えられて多くの枝葉を成していく構造をもつものに対して、**リゾーム**は、構造全体の代謝を支える中心をもたず、地中を自在にのび広がって、様々な場所に生成の拠点を形成する。

### リゾーム 【(フランス)rhizome】

- 1 根茎。茎でありながら一見根のように見えるものの総称。竹・シダなどに見られる。
- 2 現代思想で、相互に関係のない異質なものが、階層的な上下関係ではなく、横断的な横の関係で結びつくさまを表す概念。



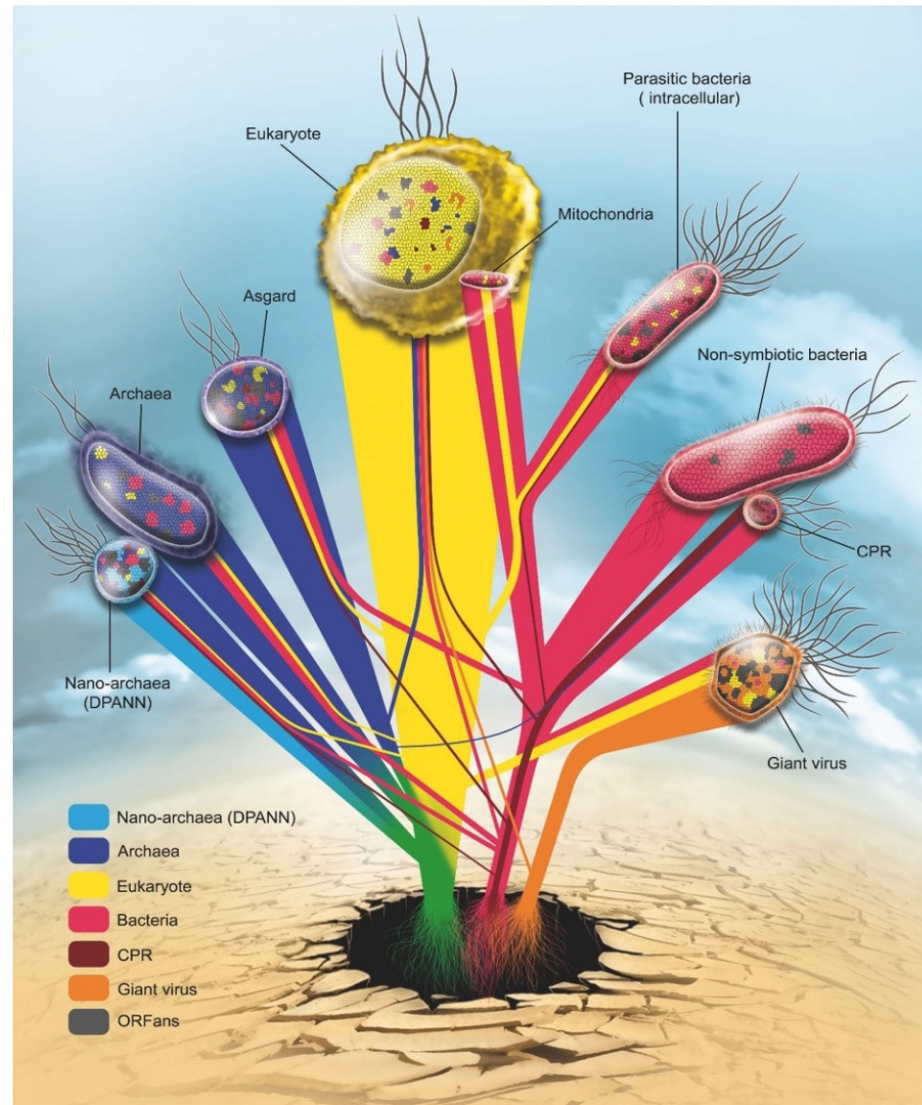
Rhizome 以外にネットワーク、ブッシュ(bush)と表現されることもある

ヨシ 根茎

<https://www.aflo.com/ja/contents/223905755> より

# Tree of Life から Rhizome of Life へ

Rhizome = 根茎  
地中あるいは地表を  
伸びる地下茎の一種



## Rhizome of Life:

Roots of each gene are represented according to the current reclassification of living organisms: Eukarya (yellow), Archaea (blue), Nano archaea (light blue), Bacteria (red), CPR (dark red) Giant virus (orange). In grey are genes without identified origin (ORFans).

Ibrahim *et al.* (2021) Rhizomal Reclassification of Living Organisms. *Int. J. Mol. Sci.* **22**, 5643. Fig.3より