

生物における 情報表現変換

真栄城 哲也

概要

- 遺伝情報と機能情報
- 遺伝コード
- 標準遺伝コードと例外遺伝コード
- 仮説の検証
- おまけ

遺伝情報の流れ

世代間の伝達

DNA : 4種類の塩基の配列

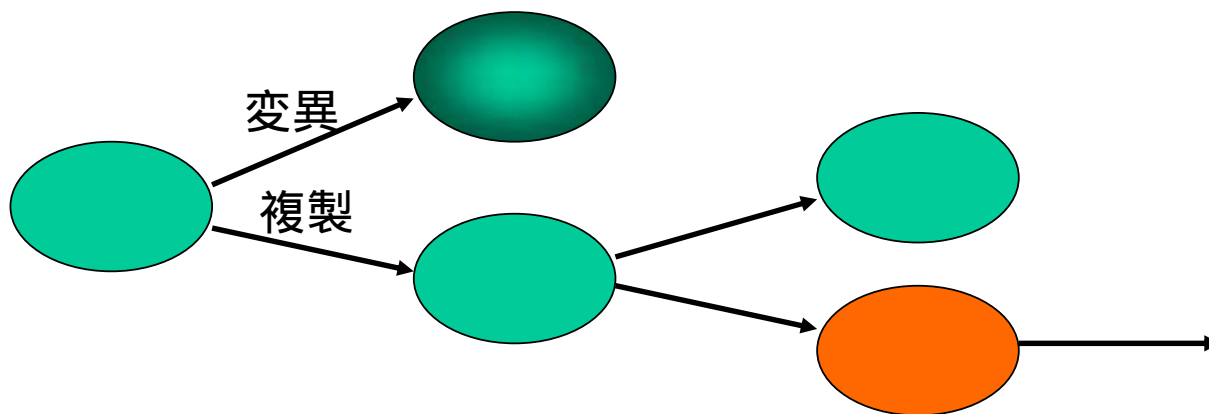
A, T, G, C

相補的
A T
G C

2重らせん

...AGCGGCCAGGATGCTTT...

...TCGCCGGTCCTACGAAA...



蛋白質

蛋白質の役割

- 1- 酵素反応
- 2- 物質の輸送や保存
- 3- 運動(筋肉など)
- 4- 免疫
- 5- 神経信号の生成と伝達
- 6- 成長と発生の制御

20種類のアミノ酸

Ala, Arg, Asn, Asp, Cys,
Gln, Glu, Gly, His, Ile,
Leu, Lys, Met, Phe, Pro,
Ser, Thr, Trp, Tyr, Val

蛋白質の形 = 機能

蛋白質の立体構造

アミノ酸配列 $\xrightarrow{\text{折り畳み}}$ 立体構造

アミノ酸の物理化学的特性

極性 親水
疎水

ヘモグロビンA	-Val-His-Leu-Thr-Pro-Glu-Glu-Lys-	親水
正常		GAG
変異体		GTG
ヘモグロビンS	-Val-His-Leu-Thr-Pro-Val-Glu-Lys-	
		疎水

遺伝情報の流れ

デジタル

DNA 情報
A, T, G, C



mRNA
A, U, G, C



遺伝コード

蛋白質 機能
20種類

アナログ

標準遺伝コード

64 コドン → 21表現型 (アミノ酸20種類 + ストップコドン)

UUU } UUC } Phe	UCU } UCC } UCA } Ser	UAU } UAC } Tyr	UGU } UGC } Cys
UUA } UUG } Leu	UCA } UCG } Ser	UAA } stop UAG } stop	UGA } stop UGG } Trp
CUU } CUC } CUA } Leu	CCU } CCC } CCA } Pro	CAU } CAC } His	CGU } CGC } Arg
CUA } CUG } Leu	CCA } CCG } Pro	CAA } CAG } Gln	CGA } CGG } Arg
AUU } AUC } Ile	ACU } ACC } Thr	AAU } AAC } Asn	AGU } AGC } Ser
AUA } AUG } Met	ACA } ACG } Thr	AAA } AAG } Lys	AGA } AGG } Arg
GUU } GUC } Val	GCU } GCC } Ala	GAU } GAC } Asp	GGU } GGC } Gly
GUA } GUG } Val	GCA } GCG } Ala	GAA } GAG } Glu	GGA } GGG } Gly

自然が用いた符号化の基準は？

標準遺伝コード

Standard Genetic Code (SGC)

- ほとんどの種で共通
- 表現型に1～6個のコドンが割り当てられている
- コドンの第3位置の変異は表現型を変化させにくい
- Ser 以外の表現型のコドンは点突然変異で遷移可能

UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA stop UAG stop	UGU } Cys UGC } UGA stop UGG Trp
CUU } Leu CUC } CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }
AUU } Ile AUC } AUA } AUG Met	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }

従来の仮説

- 偶然の産物 (Crick)
- コドンの点突然変異に対するアミノ酸の物理化学的性質の変化が最小 (Haig91, DiGiulio94)

問題(パズル)

コドン：文字数4、長さ3 $4 \times 4 \times 4 = 64$

表現型：21種類(アミノ酸20種類+ストップコドン)



大きさ64の符号空間に21種類の記号を符号化する

64 \longrightarrow 21

通信理論

どのような遺伝コードの構造(コドン - 表現型の対応)が
生物の生存に最適か？

変わりにくさ × 変わりやすさ

相反する要求

- 定常環境で生存する

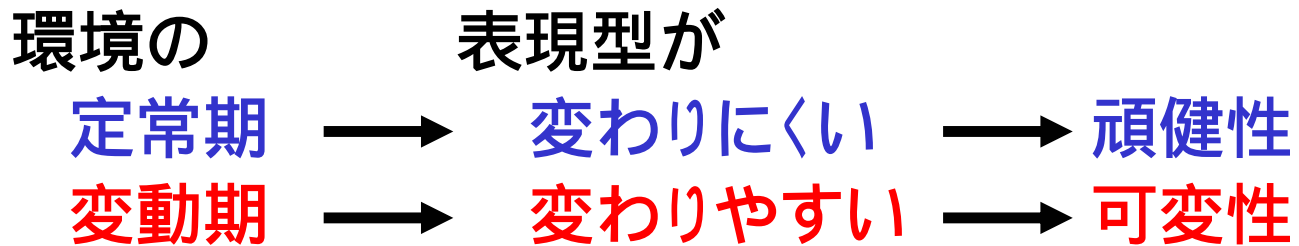
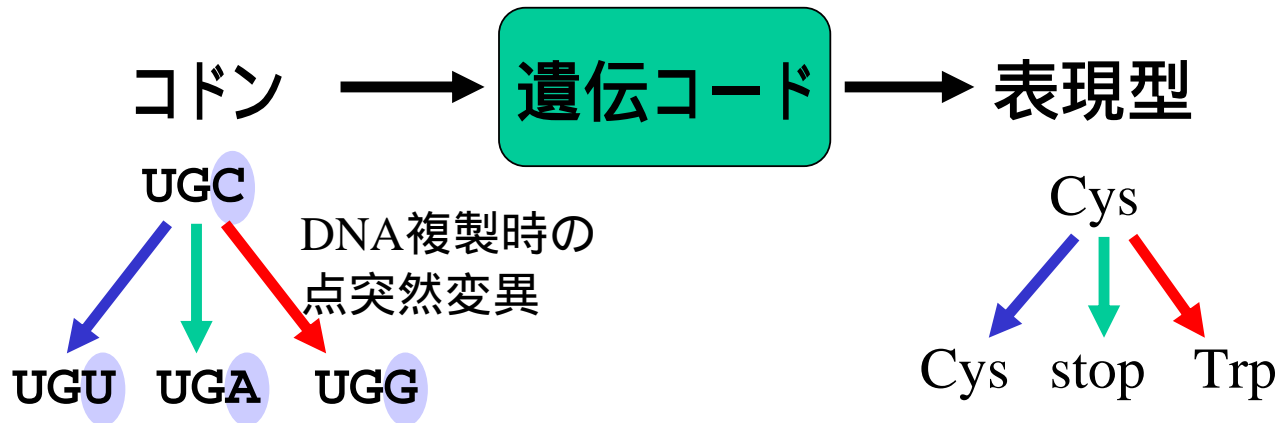
DNA複製時の誤りが蛋白質に影響しにくい
生存率が高い

機能が変化しにくい → 進化しない

- 進化する、環境変動に適応する

機能を変化させる必要性

変わりにくさと変わりやすさが同時に必要



自然が用いた符号化の基準は？

点突然変異に対する可変性と頑健性の両立

情報理論

頑健性

通信理論

ノイズに強い符号化

可変性

暗号理論

解読に強い暗号化
鍵の類似度と暗号化情報の類似度

コドン：文字数4、長さ3 $4 \times 4 \times 4 = 64$

表現型：21種類(アミノ酸20種類+ストップコドン)



大きさ64の符号空間に21種類の記号を符号化する

どのような遺伝コードの構造(コドン - 表現型の対応)が
生物の生存に最適か？

例外遺伝コード

標準遺伝コードから進化(派生)

標準遺伝コードとの違い:

コドンと表現型の対応が違う

酵母のミトコンドリアの
例外遺伝コード

UGA=Trp

AUA=Met

CUU,CUC,CUA,CUG=Thr

例外遺伝コードを持つ遺伝系:

- ミトコンドリア
- 単細胞生物(単純な)

高い突然変異確率

短いゲノム(DNA)

UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA stop UAG stop	UGU } Cys UGC } UGA stop UGG } Trp
CUU } CUC } CUA } CUG } Thr	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }
AUU } AUC } Ile AUA } AUG } Met	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }

従来の仮説

新たに割り当てられる表現型は偶然に決まる

(Jukes & Osawa 88)

例外遺伝コードの出現

コドンの再割り当て

遺伝コードが変化すると
その生物の機能が変わる
下手すると死ぬ

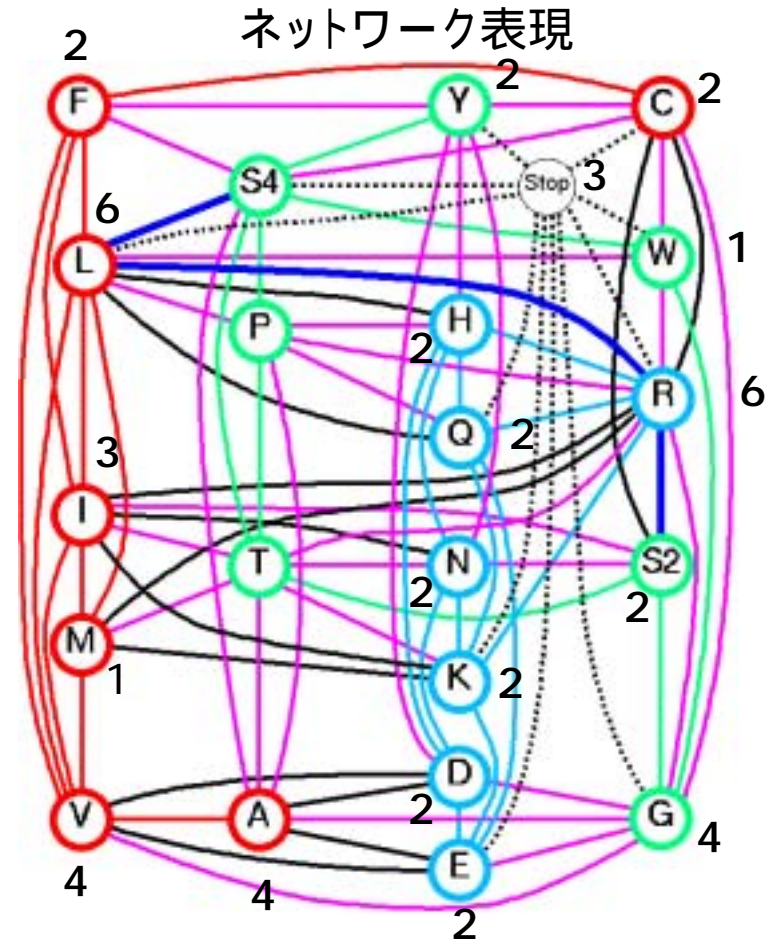
Frozen accident theory

	コドン使用の偏り	GC圧
GAU } Asp GAC }	Asp-Asp-Glu-Glu-Asp	未使用コドン
GAA } Glu GAG }	GAUGAUGAAGAAGAU → GAC; GAG GACGACGAGGAGGAC → GAU; GAA	

DNA(ゲノム)が短いほど未使用コドンが生じやすい

標準遺伝コードの構造

UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA stop UAG stop	UGU } Cys UGC } UGA stop UGG Trp
CUU } Leu CUC } CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }
AUU } Ile AUC } AUA } AUG Met	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }



ノード内とノード間の遷移確率の平均

頑健性

最適率82.9%

可変性

最適率42.6%

頑健性

突然変異

蛋白質の機能が変わる・失われる

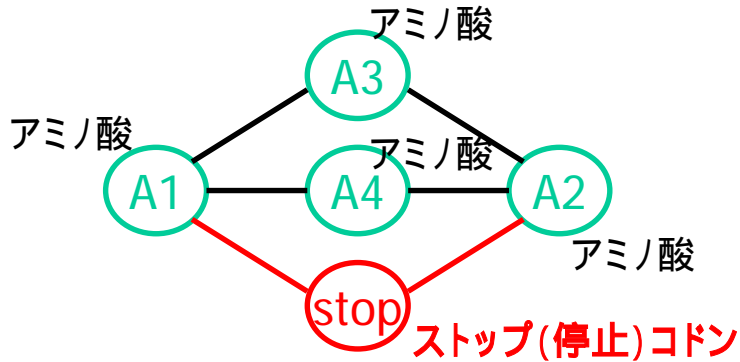
停止頑健性

蛋白質の長さの変化

より短い蛋白質



ほぼ確実な機能の損失



アミノ酸から別のアミノ酸に
変異する時の
ストップコドンが生じる確率

A1 A2

例外遺伝コード

15種類

Mitochondrion Yeast	UGA(stop) Trp AGR(Arg) Ser CUN(Leu) Thr
Mitochondrion Euscomycete Mitochondrion Cnidaria	UGA(stop) Trp
Mitochondrion Platyhelminth	UGA(stop) Trp UAA(stop) Tyr AAA(Lys) Asn AGR(Arg) Ser
Mitochondrion Nematoda Mitochondrion Protozoa Mitochondrion Arthropoda Mitochondrion Mollusca	UGA(stop) Trp AUA(Ile) Met AGR(Arg) Ser
Mitochondrion Echinodermata	UGA(stop) Trp AAA(Lys) Asn AGR(Arg) Ser
Mitochondrion Chlorophyceae	UAG(stop) Leu
Mitochondrion Scenedesmus obliquus	UAG(stop) Leu UCA(Ser) stop
Mitochondrion Echinodermata	UUA(Leu) stop

CUN = CUU/CUC/CUA/CUG

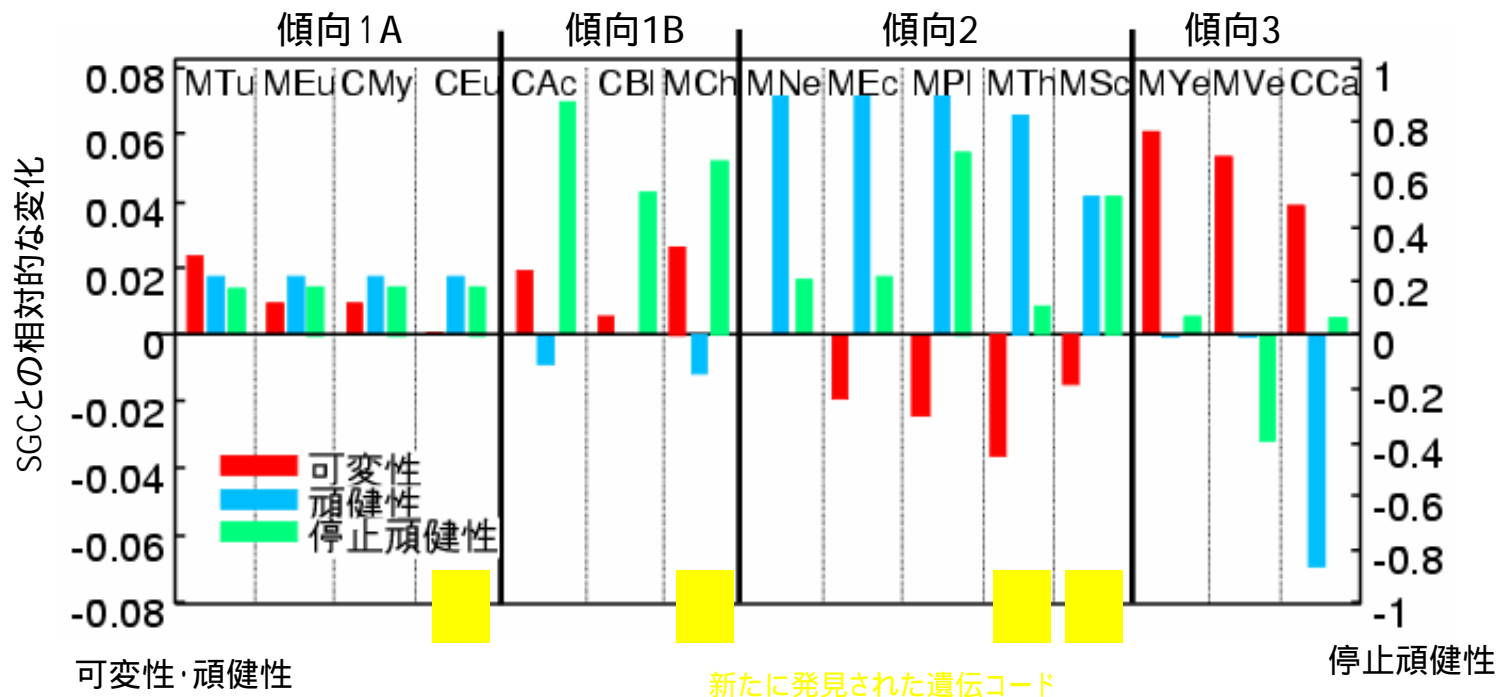
AGR = AGA/AGG

Mitochondrion Vertebrata	UGA(stop) Trp AUA(Ile) Met AGR(Arg) stop
Mitochondrion Tunicata	UGA(stop) Trp AUA(Ile) Met AGR(Arg) Gly
Mitochondrion Trematode	UGA(stop) Trp AAA(Lys) Asn AUA(Ile) Met
Mycoplasma Spiroplasma	UGA(stop) Trp
Euplote	UGA(stop) Cys
Acetabularia Tetrahymena Paramecium	UAR(stop) Gln
Blepharisma	UAG(stop) Gln
Candida	CUG(Leu) Ser

ミトコンドリア
核

新たに発見された遺伝コード

例外遺伝コードの頑健性と可変性



傾向1: 頑健性と可変性

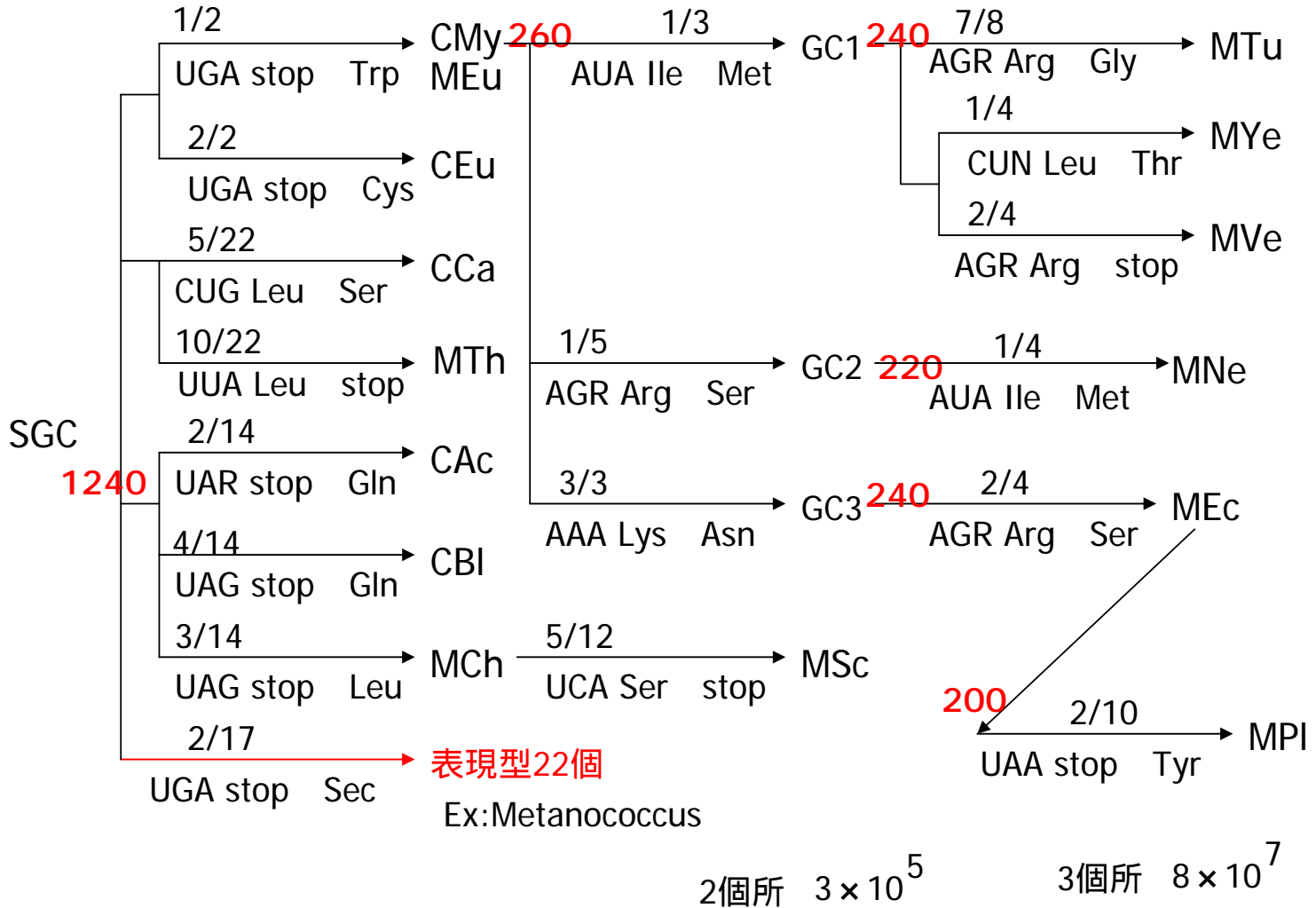
傾向2: 頑健性に偏り

傾向3: 可変性に偏り

傾向1A: 全て向上

傾向1B: 可変性と停止頑健性

例外遺伝コードの進化過程



頑健性と可変性

どちらの方が必要か？

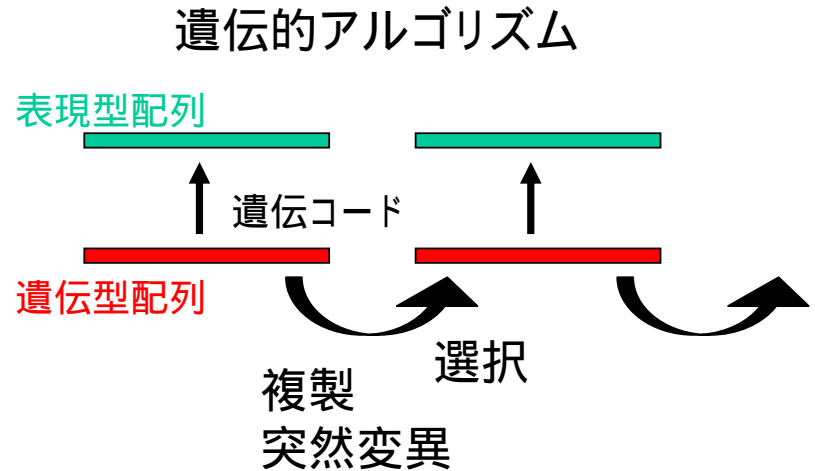
標準遺伝コードの最適率：頑健性82.9%、可変性42.6%

シミュレーション

集団1: 頑健性最大の遺伝コード

集団2: 可変性最大の遺伝コード

集団を競合
環境を変動



親に似ている(頑健性が高い)方が生存する確率が高い

頑健性が高い方が集団の生存率が高い

まとめ

遺伝コードの構造は可変性と頑健性の要求

例外遺伝コードの予測