- Q1: Find the secret message in the sentences.
  - Why is DNA necessary? Life is very special and cells are great things. Significantly advanced life science is easily mistaken as magic.
- Q2: Can I delete a “w” letter in these sentences?

Session 4:Synthetic/constructive experiment of life and evolution of genome

Daisuke Kiga

ELSI /

Dept of computational intelligence and systems science,

Tokyo Institute of Technology

Back ground: biochemistry

Why 4 nucleotide and 20 amino acids?

For resurrection of proteins before commonotes,
engineered genetic code less than 20 amino acids is important

Previous works: engineering of genetic code
xanthine-biniding aptamer
autonomous DNA computing,
artificial genetic network in living cell

Acknoledgements

遺伝暗号の改変

- 21番目のアミノ酸
  - 理研　SSBC （横山グループ）
- 19アミノ酸遺伝暗号
  - 木賀研
    - 内山正彦（NEDO研究員）
    - 小林晃大
    - 網蔵和晃
- 無細胞タンパク合成
  - 木川隆則
- アミノ酸組成分析
  - 理研 堂前直・中山洋・益田晶子
- 質量分析
  - 東大　福沢世傑・児玉公一郎
- 結晶構造解析
  - 東大　荒磯裕平・濡木理
- NEDO産業技術助成事業 (-2009.3)
- JSPS科研費若手（A) (-2009.3), (-20011.3)
- JSPS科研費挑戦的萌芽 (-2013.3)

人工遺伝子ネットワーク

- 東工大　木賀研
  - 鮎川翔太郎
  - 畑敬士
  - 石松愛
- 東工大　山村雅幸
- 東大　陶山研
  - 陶山明
  - 庄田耕一郎
  - 瀧ノ上正浩
  - 李 泳薫
- 東大　萩谷昌己
- JSPS科研費特定「分子プログラミング」 (-2007.3)
- JSPS科研費若手（B) (-2007.3)
- JST先端計測(-2009.3)
- JSPS科研費挑戦的萌芽(-2009.3) (-2010.3)
- JSTさきがけ(-2012.3)
- JSPS科研費若手（A) (-2014.3)
- JSPS科研費新学術領域「合成生物学」計画研究(-2016.3)

赤字:研究代表

contents

- Synthetic/Constructive approach in life science
  - proof by synthesis
  - construction of another life

- Simplified genetic code
  for reconstitution of ancient protein
  - code and protein with less than 20 amino acids

- How to engineer genetic code
  - detailed results

contents

- Synthetic/Constructive approach in life science
  - proof by synthesis
  - construction of another life

- Simplified genetic code
  for reconstitution of ancient protein
  - code and protein with less than 20 amino acids

- How to engineer genetic code

Understanding of System

analytical approach

constructive approach

What I cannot create, I do not understand. -R. Feynman

Life is a multi layer system

analytical approach

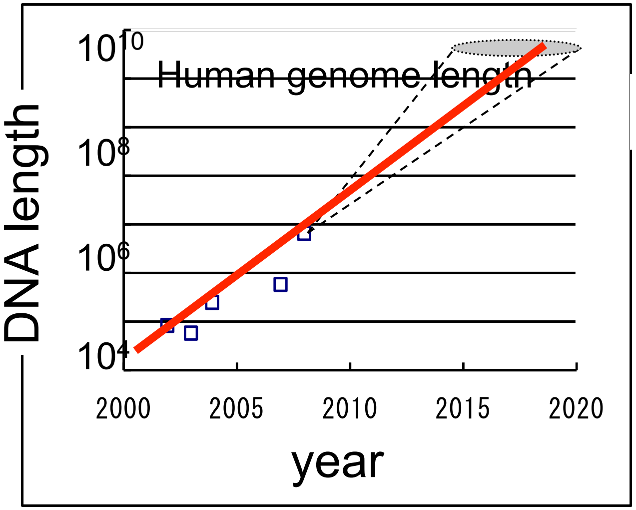
• ● atom	building block	bio macromolecule	subsystem of cell	cell	tissue	individual
• ● Carbon • ● Oxygen Nitrogen	nucleotide amino acid lipid	DNA RNA protein	gene network protein synthesis Photosynthesis vesicle			

Recent Developments

- Comprehensive analysis: genome, proteome….
- Preparation of biomolecules

Developments in preparation of biomolecules

Moor’s Low in DNA synthesis: exponential growth



Kiga and Yamamura, New Generation Computing, 26 (2008), p347-364

2 JULY 2010 VOL 329 SCIENCE 52

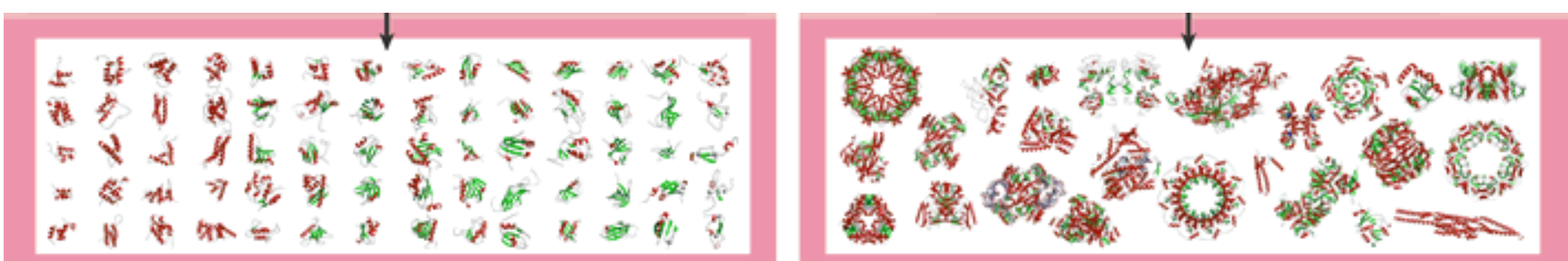
Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

J. Craig Venter[1,2*]

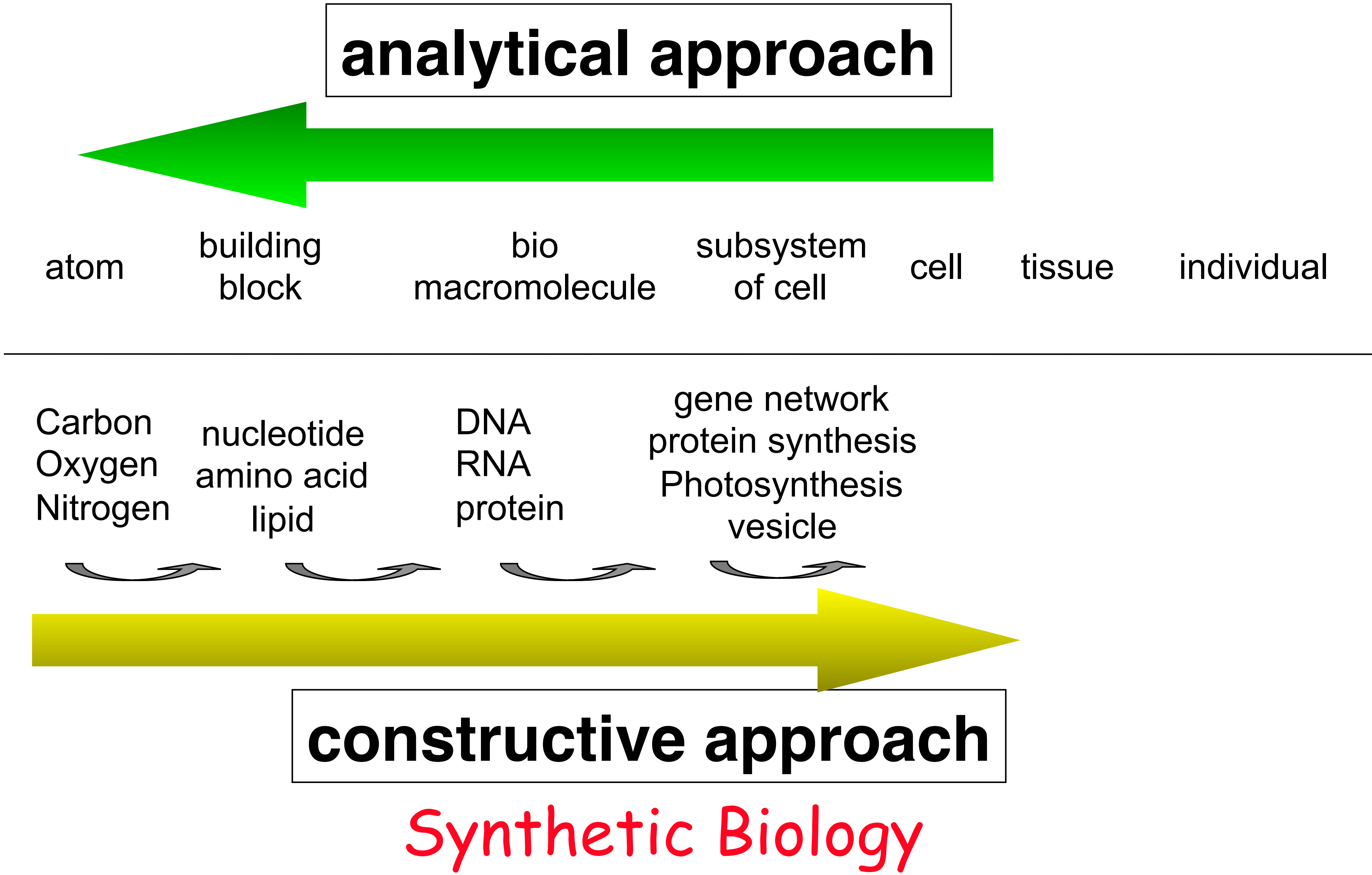
2015-2020 artificial DNA
with human genome length!?

Preparation of proteins

Protein 3000: Structural genomics

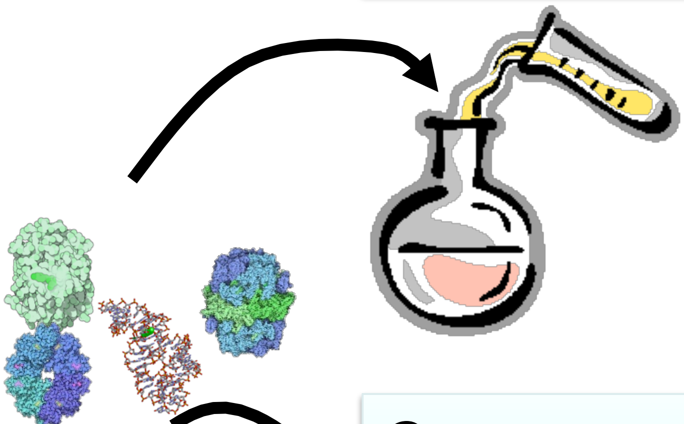


Natural / unnatural combinations can be realized



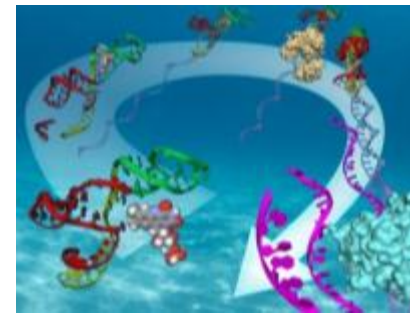
synthetic molecular system

Mix proteins and DNA/RNA



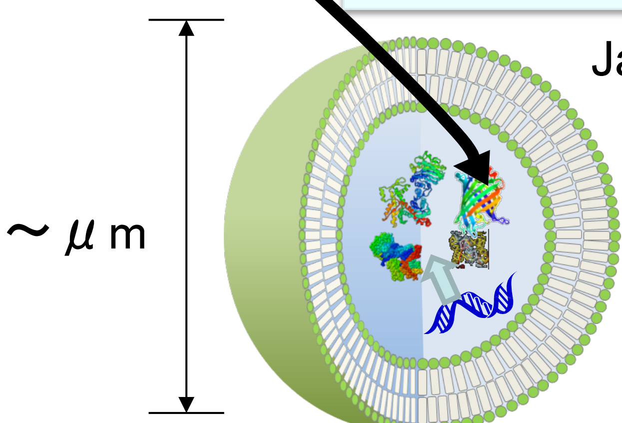
In vitro protein production
DNA computation
/molecular programming

aptazyme system



ChemComm 2012
cover figure

Can we synthesize a cell?



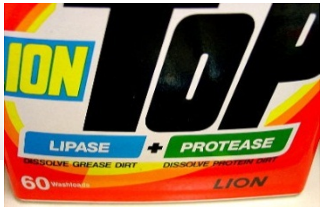
Japanese Society for Cell Synthesis Research

President -2008 Tetsuya Yomo
President -2012 Daisuke Kiga

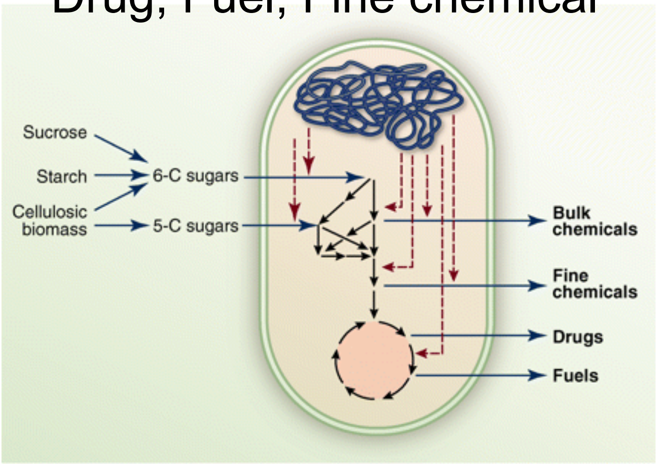
Objectives of Synthetic biology

- Engineering
  - protein engineering
  - metabolic engineering
  - tissue engineering
- Science

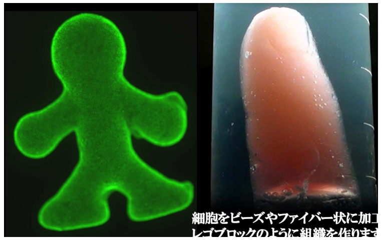
washing-powder



Drug, Fuel, Fine chemical

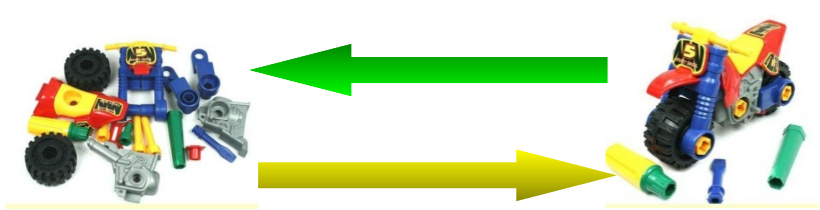


combination of cells



Objectives of Synthetic biology

- Engineering



Process for robust system

- Science
  - Reconstitution: proof-by-synthesis
    - ATPase, RNA Pol, Ribosome
    - PURE Translation system
    - Regulatory circuit
  - Construction of “another life”

Objectives of Synthetic biology

- Engineering

- Science
  - Reconstitution:

Process for robust system

Natural Complicated system

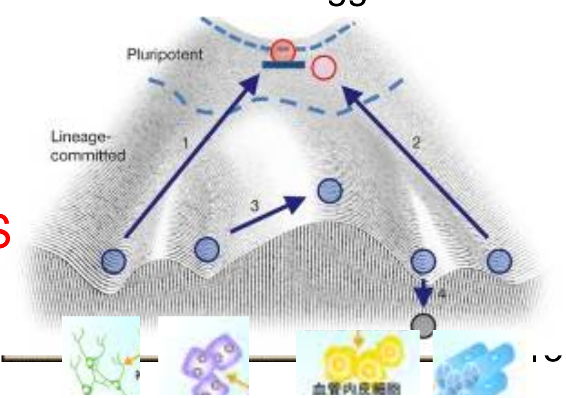
Waddington 1957

Epigenetic landscape

fertilized egg

one phenotype
↓
multi phenotypes

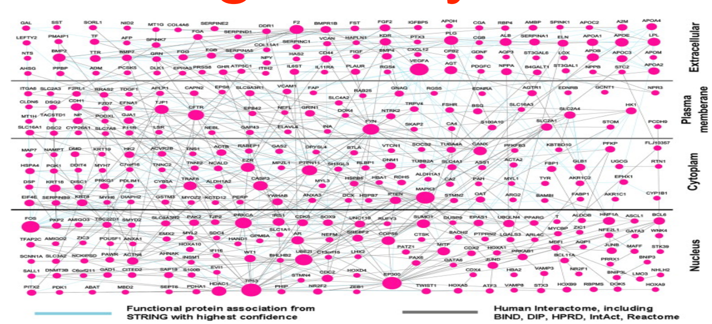
iPS



S Yamanaka *Nature* (2009)

differentiated cells

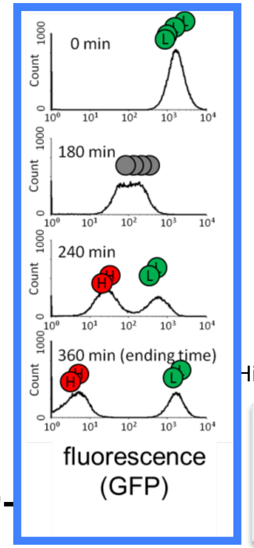
- Regulatory circuit

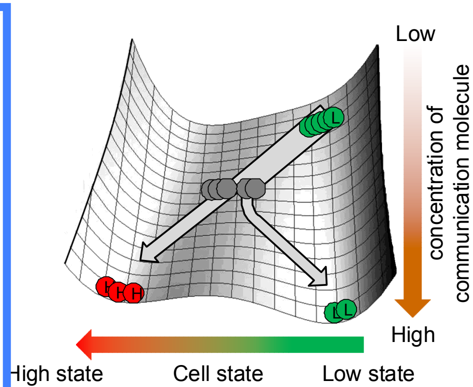


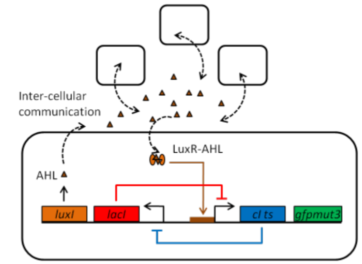
Molecular Interaction Network in Human Organogenesis
Fang et al , Dev Cell 2010

Objectives of Synthetic biology

Kiga Lab
PNAS (2011)
vol108, 17969-







– Reconstitution:

$$\frac{dy_i}{dt} = \frac{\alpha_y}{1+\left(x_i/K_x\right)^{nx}} \cdot \frac{z^{nz}}{K_z^{nz}+z^{nz}} - d_y \cdot y_i$$

$$\frac{dx_i}{dt} = \frac{\alpha_x}{1+\left(y_i/K_y\right)^{ny}} - d_x \cdot x_i$$

$$\frac{dz}{dt} = \frac{N_{tot}}{N} \sum \lambda \cdot x_i - d_z \cdot z$$

Wet experiment & Modeling

Synthetic Simple system
in living cell/test tube

Process for
robust system

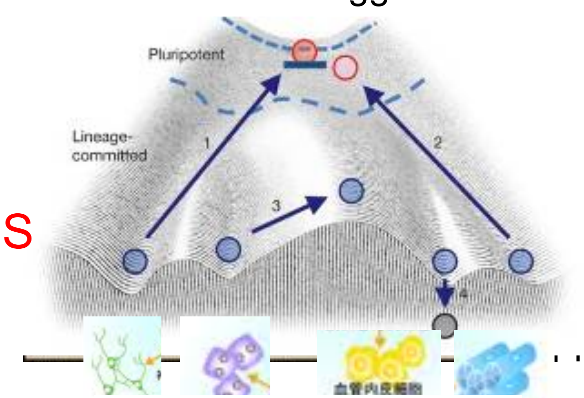
Natural Complicated system

Waddington 1957

Epigenetic landscape
fertilized egg

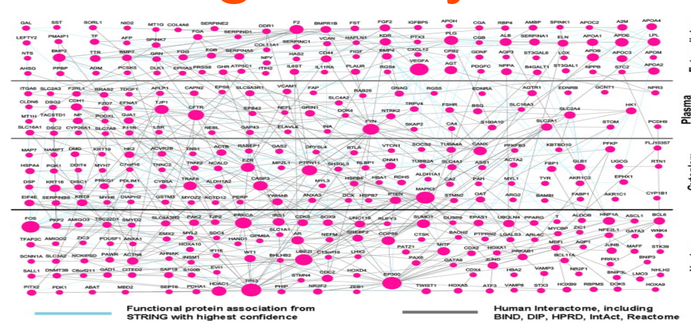
one
phenotype
↓
multi
phenotypes

iPS



differentiated cells

• Regulatory circuit



Molecular Interaction Network in Human Organogenesis
Fang et al , Dev Cell 2010

Objectives of Synthetic biology

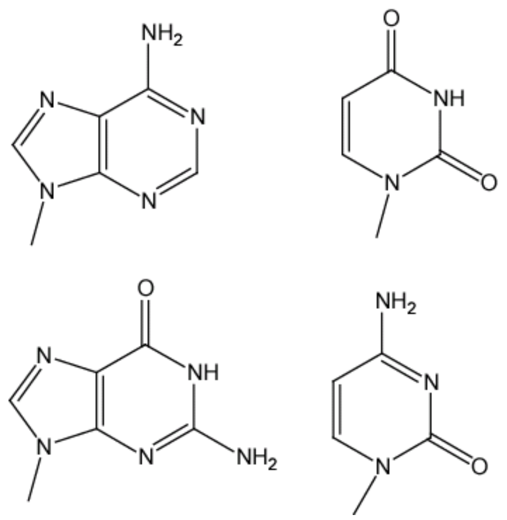
- Engineering
  - protein engineering
  - metabolic engineering
  - tissue engineering

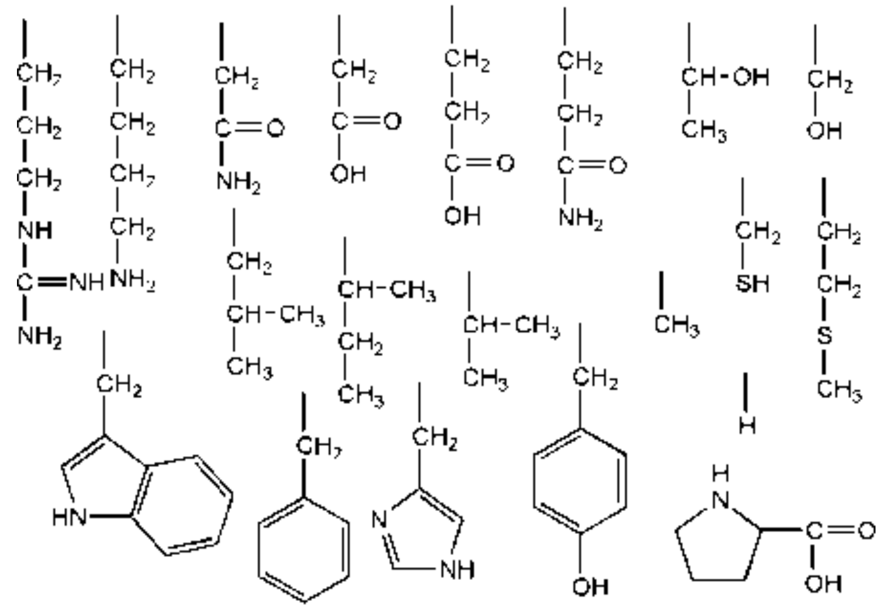
- Science
  - Reconstitution:
    proof-by-synthesis
    - ATPase, RNA Pol, Ribosome
    - PURE Translation system
    - Regulatory circuit
  - Construction of
    "another life"
    - other "natural " systems
      made of natural parts
    - what life could be
      in early earth or exoplanets

contents

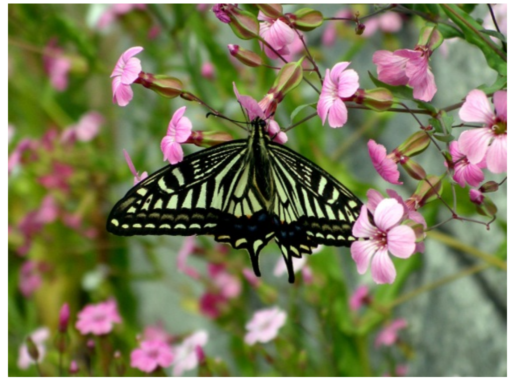
- Synthetic/Constructive approach in life science
- Simplified genetic code
  for reconstitution of ancient protein
  - code and protein with less than 20 amino acids
- How to engineer genetic code

Why 4 nucleotide and 20 amino acids?





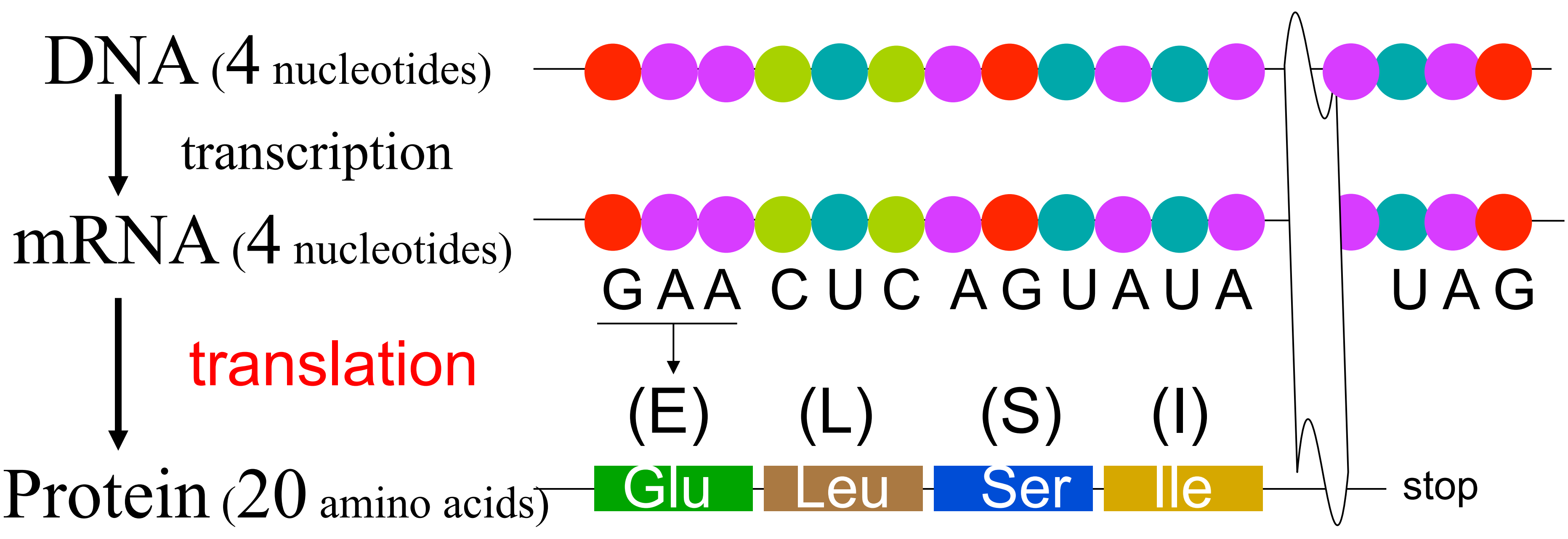
Only one solution for one function?



many
“alternative forms”

Universal property in Present Life is Important for Life?

- Replication, Metabolism, Containment
- Carbon-based polymer
  - L-amino acids, D-ribose
- Succession of DNA through generations
- DNA=Blue Print, Protein=Functional Molecule



The Universal genetic code

from *E. coli* to Human

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

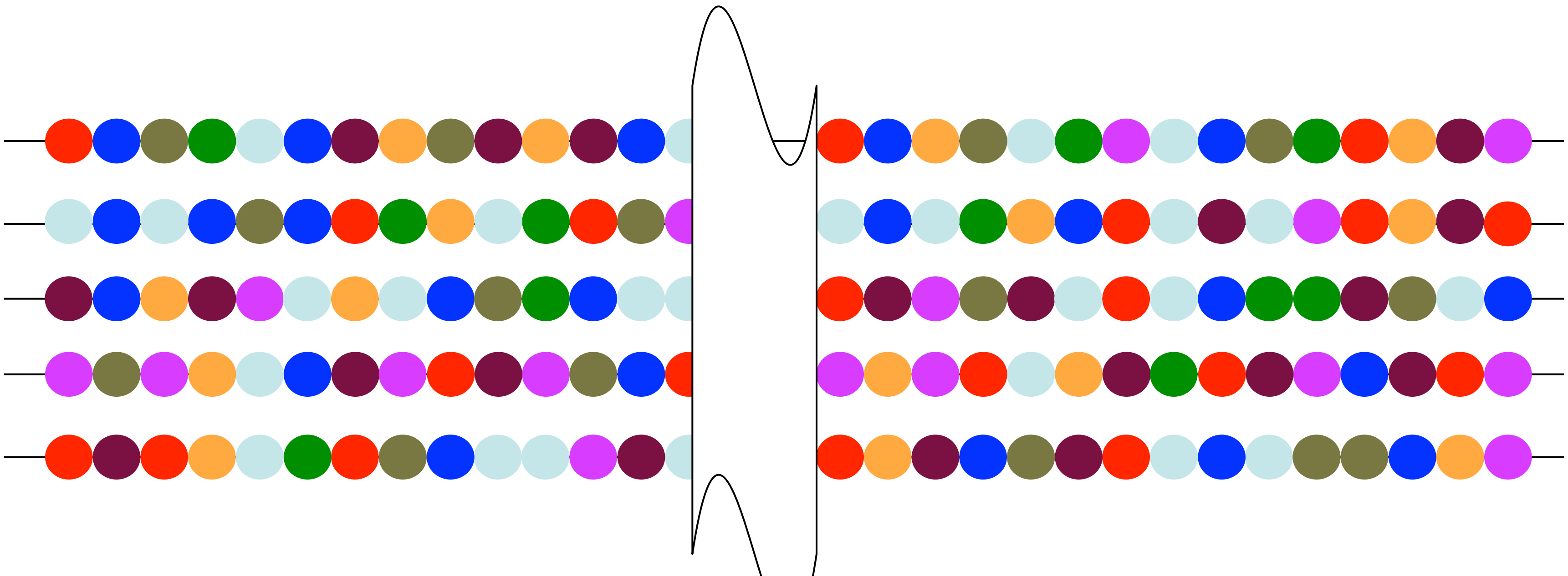
Nature creates little part of "What life could be"

Sequence variety for a small protein (200 amino acids length)

$$20^{200} \fallingdotseq 10^{260}$$

atoms in the Universe $10^{70} \sim 10^{80}$

age of the Universe 10^{16} (sec)



Is set of natural 20 amino acids the only solution ?

The Univerasl Genetic Code

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

～2005:　all of the natural codes contains 20 aa.

Why 20 amino acids in the Universal code?
can we construct other codes?

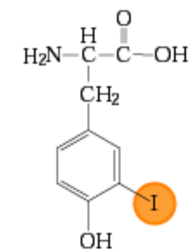
The universal genetic code **20**

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

21

Expanded genetic code

UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC		UCC		UAC		UGC	
UUA	Leu	UCA		UAA	Stop	UGA	Stop
UUG		UCG		UAG	21st aa	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	Lys	AGA	Arg
AUG	Met	ACG		AAG		AGG	
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC		GCC		GAC		GGC	
GUA		GCA		GAA	Glu	GGA	
GUG		GCG		GAG		GGG	



3-iodo tyrosine

Kiga et al., *PNAS* 2002

“Simplified” genetic code

19

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

any mRNA produces protein w/o Trp

Kiga Lab
Nucleic Acids Research
2012, featured article

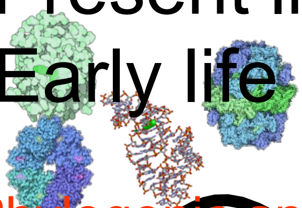
synthetic molecular system

Mix proteins and DNA/RNA

Subsystem of early life

Present life: 20 amino acid

Early life : Less than 20

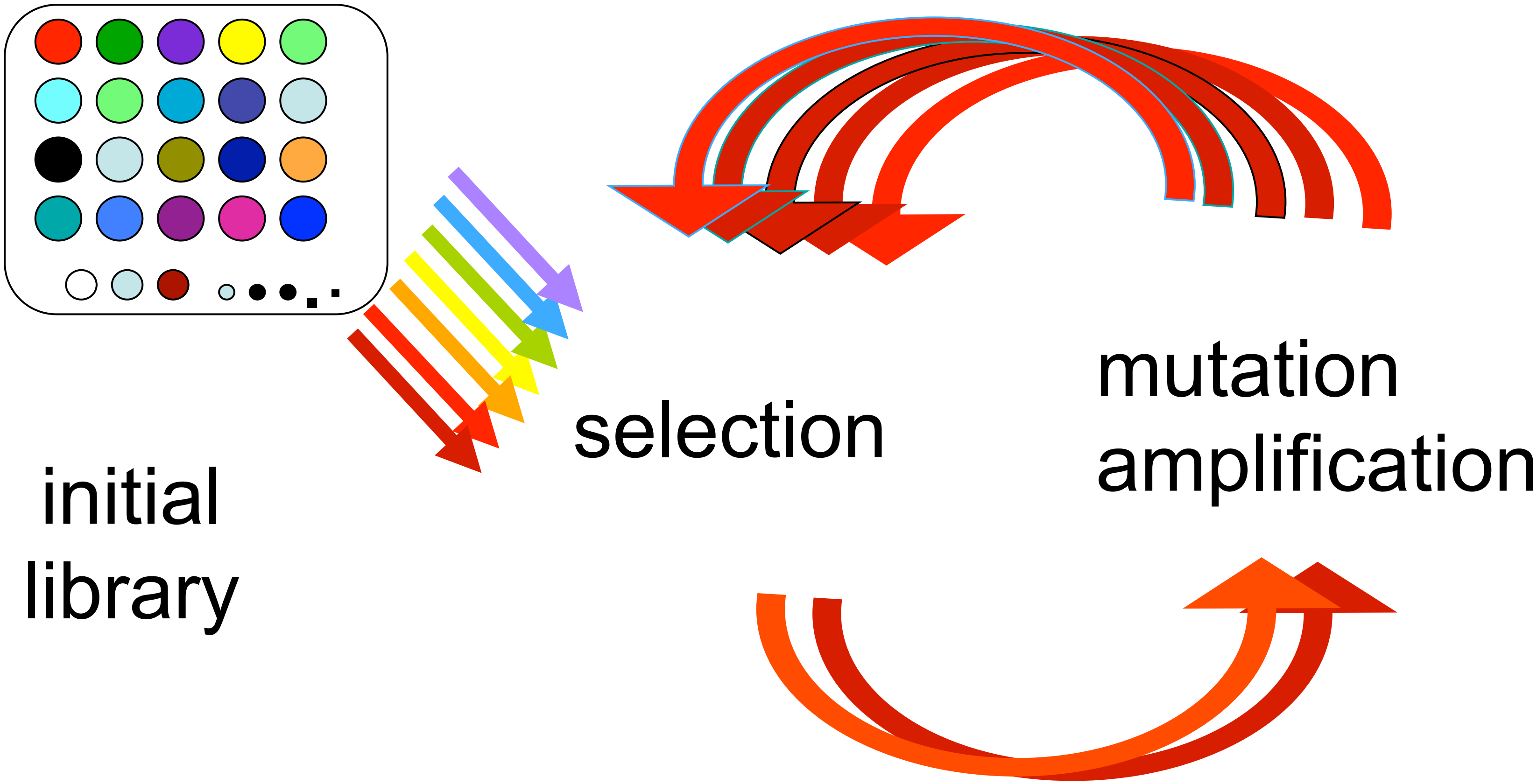


/Phylogenic analysis of components
for protein production shows possibility of
organisms with less than 20 AA

/Chemical evolution experiments
can produce not all 20 amino acids
(exlude non-canonical amino acids
/accept contamination of non-canonical AA)

Can proteins with less than 20 amino acids provide activities for life system?

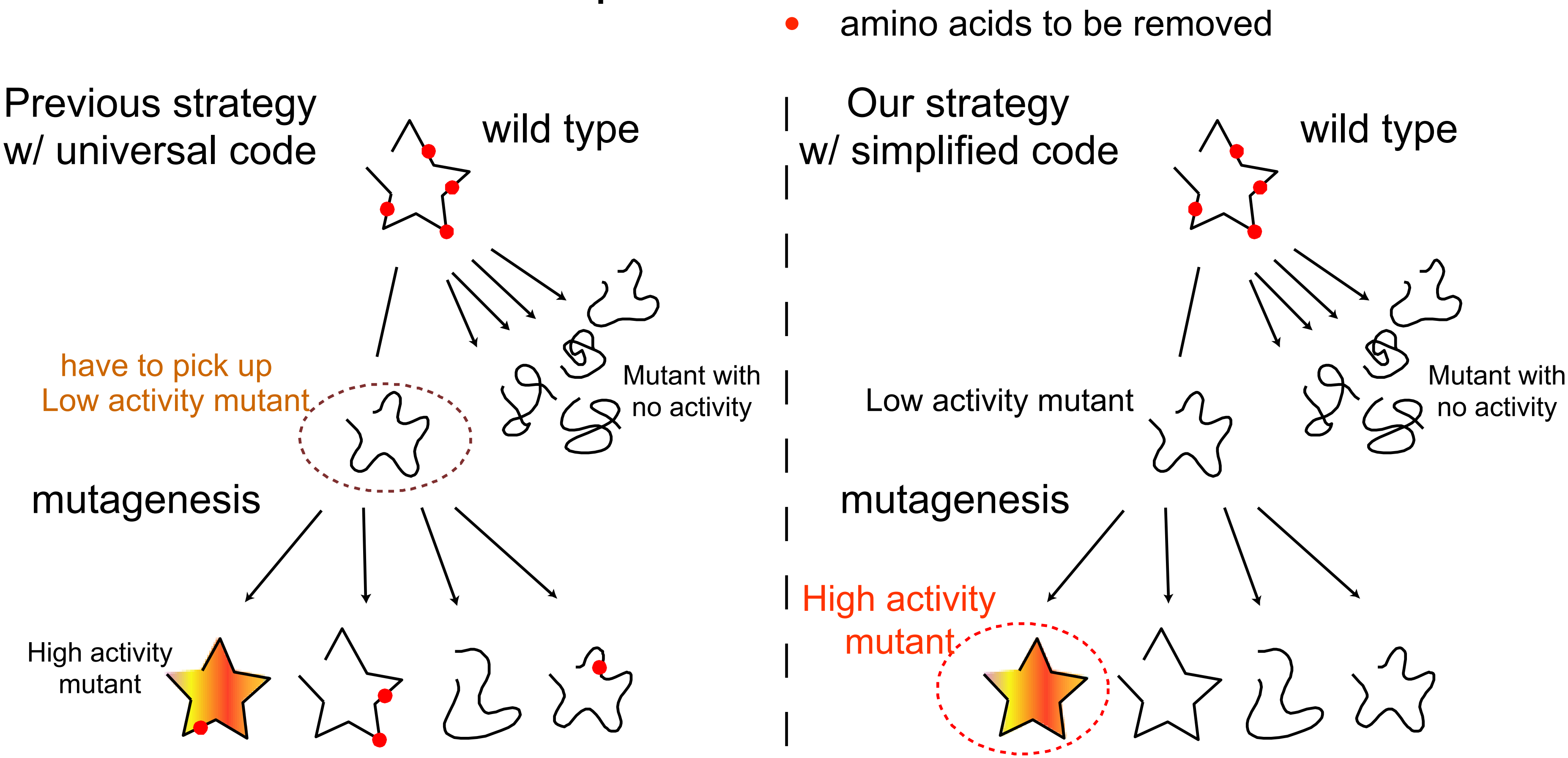
Artificial evolution of protein



Mutations are introduced to nucleotide sequence on DNA, not amino acid sequence on protein

Can proteins with less than 20 amino acids provide activities for life system?

Artificial evolution of protein with less than 20 amino acids



Cannot use evolutionary process

mutations on DNA restore codons
for amino acids to be removed

w/ simplified genetic code

Can use evolutionary process

amino acids to be removed never appear
even with mutations on DNA

contents

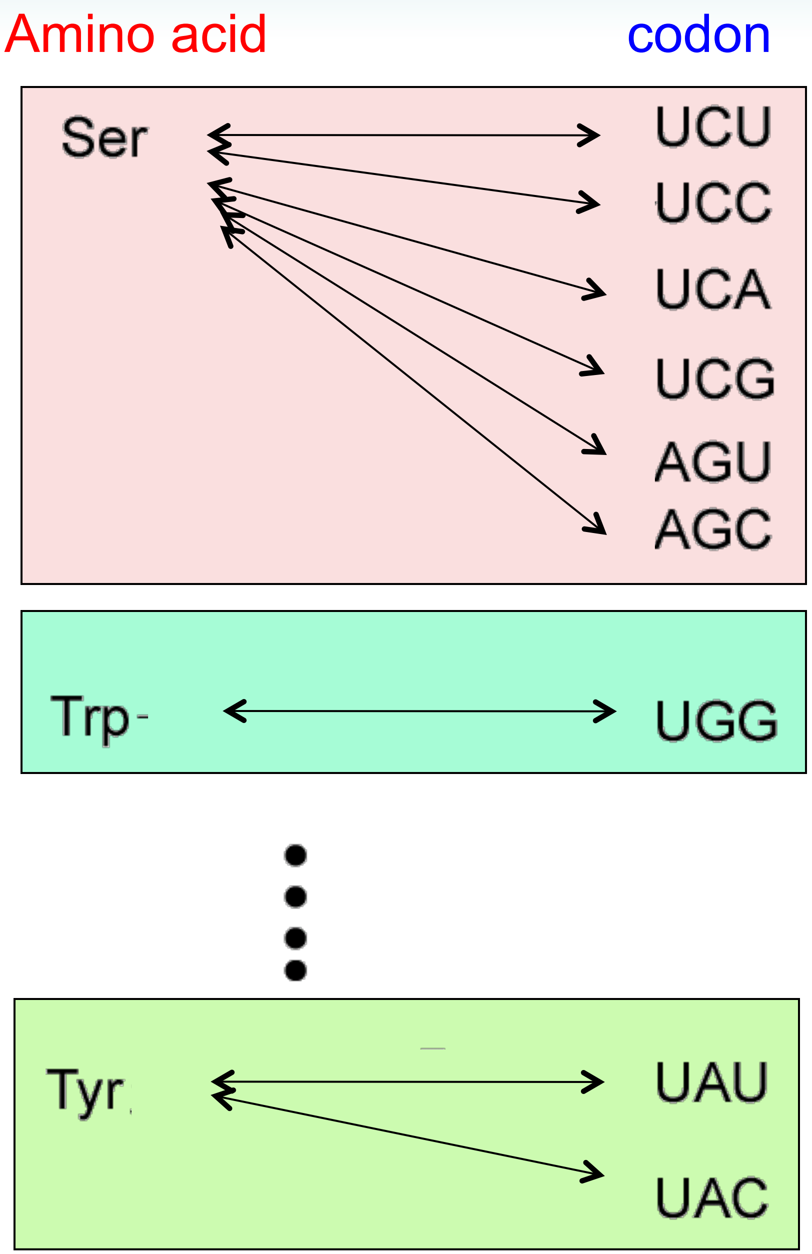
- Synthetic/Constructive approach in life science
- Simplified genetic code
  for reconstitution of ancient protein
  - code and protein with less than 20 amino acids
- How to engineer genetic code
  - detailed results

The Universal genetic code

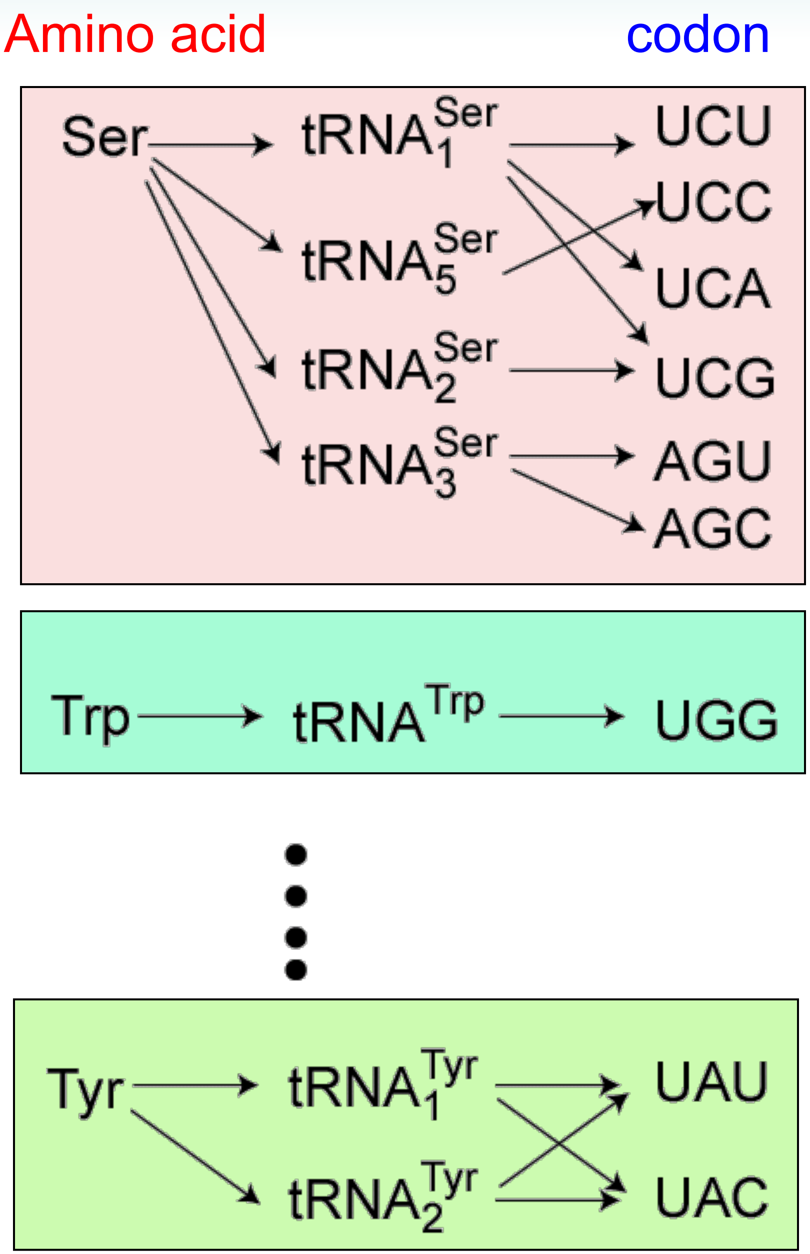
from *E. coli* to Human

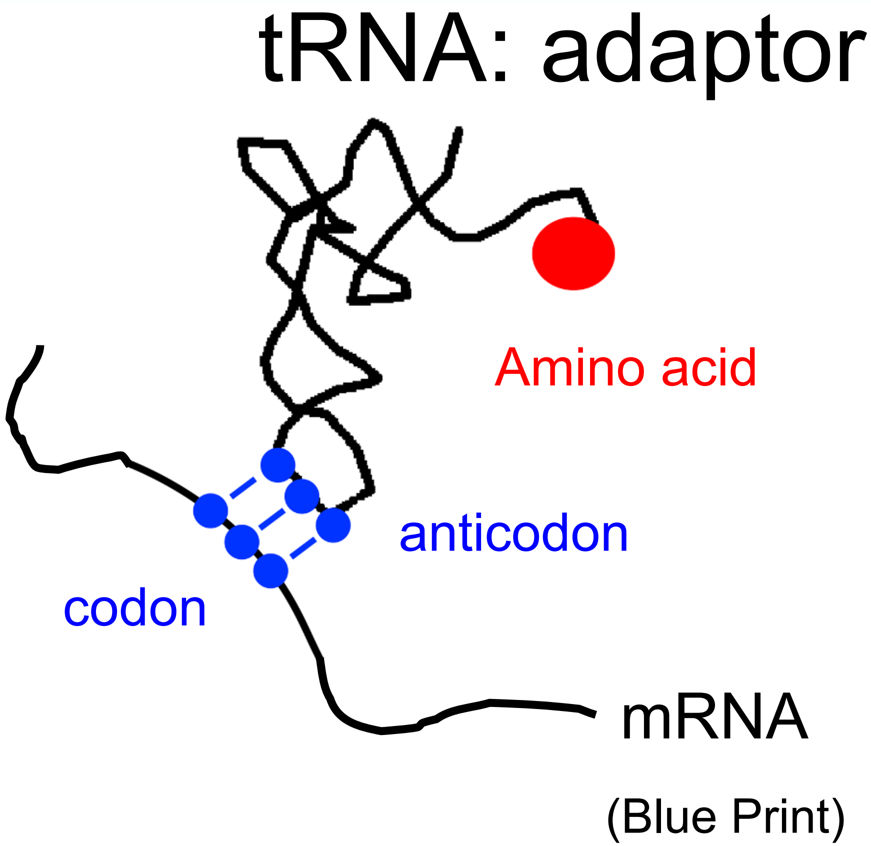
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC		UCC		UAC		UGC	
UUA	Leu	UCA		UAA	Stop	UGA	Stop
UUG		UCG		UAG	Stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	Lys	AGA	Arg
AUG	Met	ACG		AAG	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

Mechanism of the genetic code

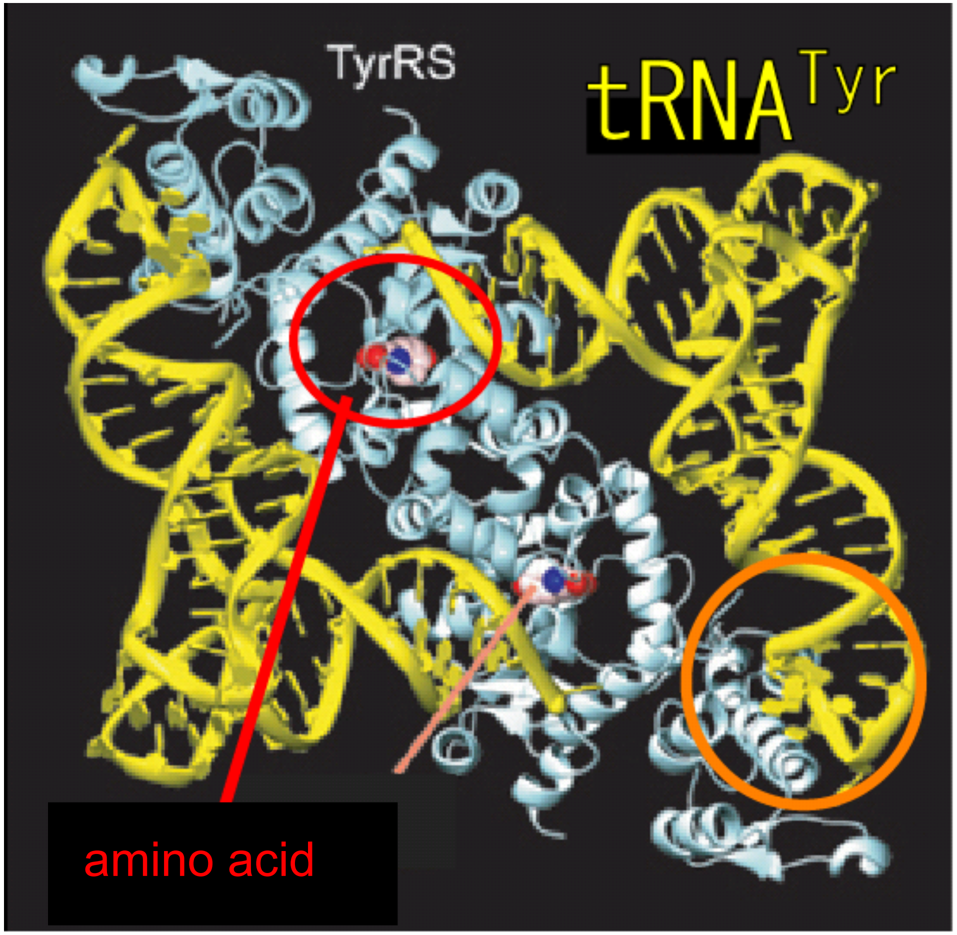


Mechanism of the genetic code





Aminoacyl-tRNA synthetase (aaRS) attaches amino acid to tRNA



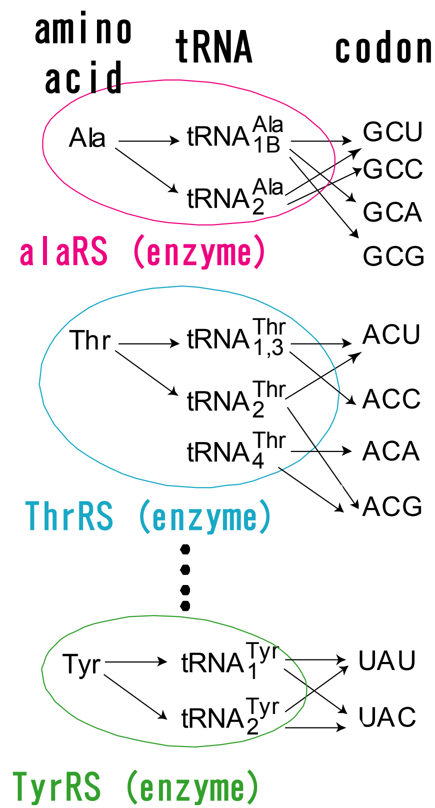
anticodon

Natural 20 exclusive groups

of amino acid, enzyme and tRNAs

The Univerasl Genetic Code

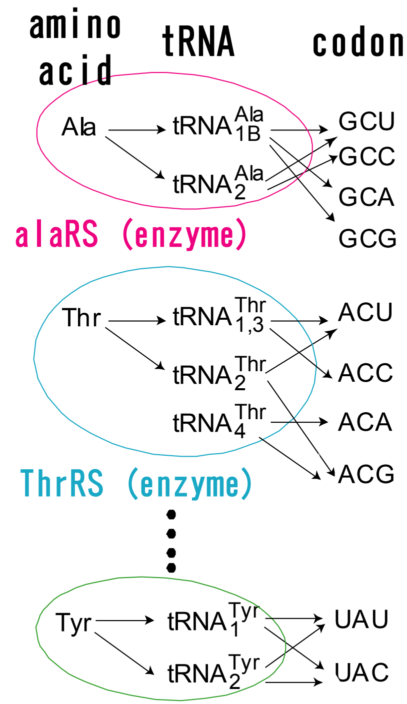
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC		UCC		UAC		UGC	
UUA	Leu	UCA		UAA	Stop	UGA	Stop
UUG		UCG		UAG	Stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	Lys	AGA	Arg
AUG	Met	ACG		AAG	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



Genetic code with 21 amino acids by creation of additional group

an expanded genetic code

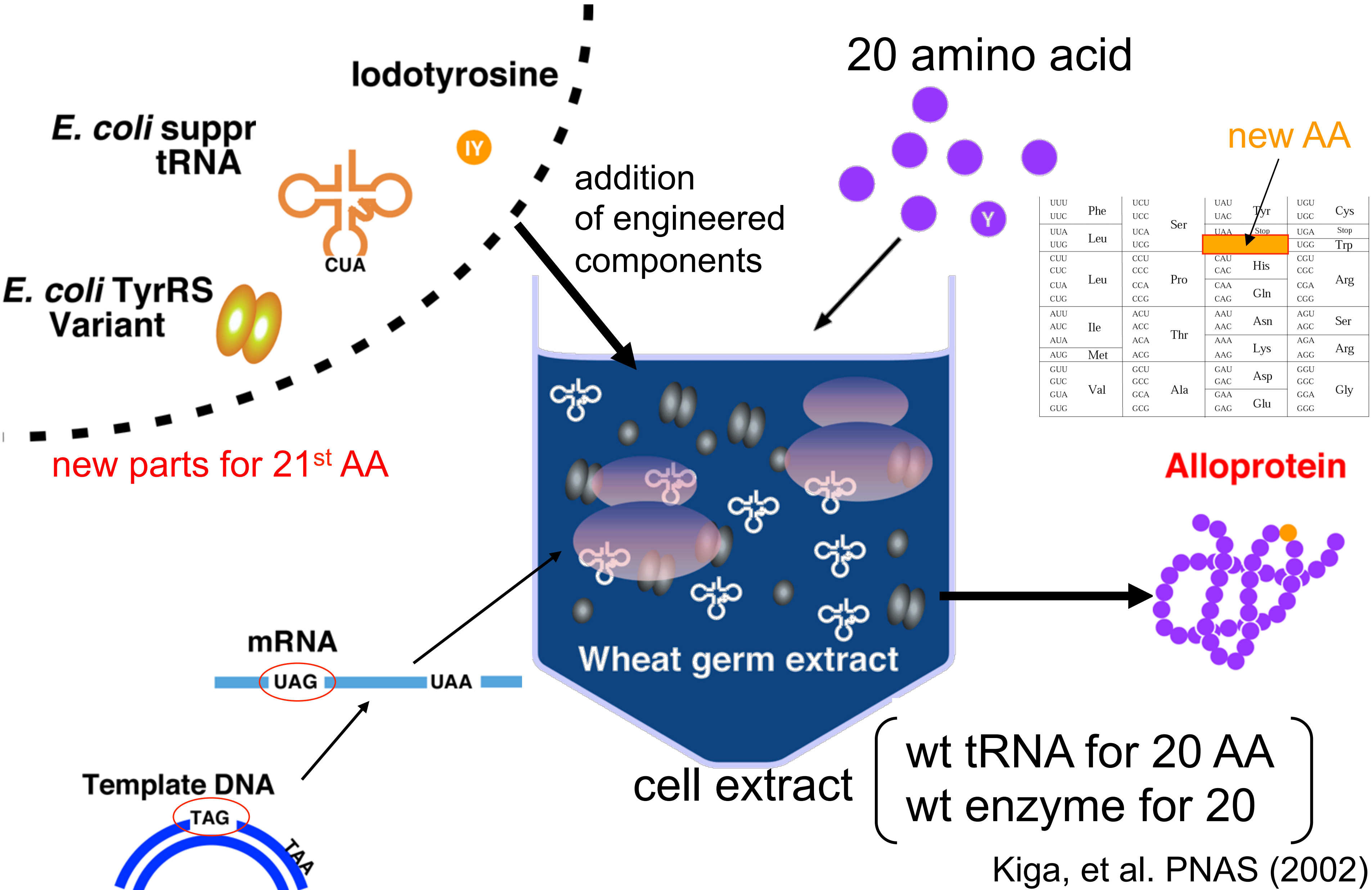
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC		UCC		UAC		UGC	
UUA	Leu	UCA		UAA	Stop	UGA	Stop
UUG		UCG		UAG	21st aa	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	Lys	AGA	Arg
AUG	Met	ACG		AAG		AGG	
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC		GCC		GAC		GGC	
GUA		GCA		GAA	Glu	GGA	
GUG		GCG		GAG		GGG	



new group without crosstalk

Kiga, et al. PNAS (2002)

Expansion of genetic code system by addition of engineered components for the 21st amino acid

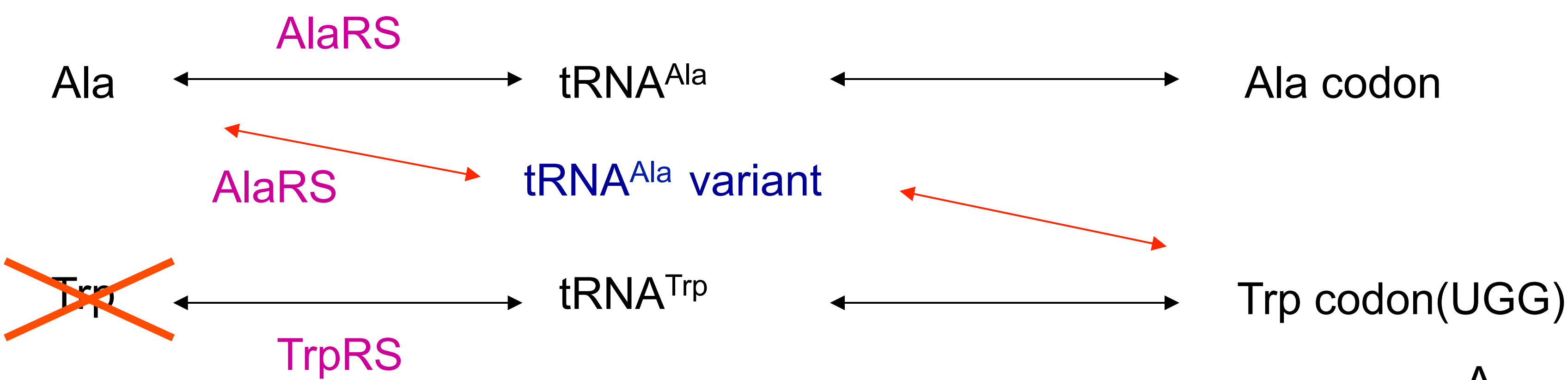


UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC		UCC		UAC		UGC	
UUA	Leu	UCA		UAA	Stop	UGA	Stop
UUG		UCG				UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	Lys	AGA	Arg
AUG	Met	ACG		AAG		AGG	
GUU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

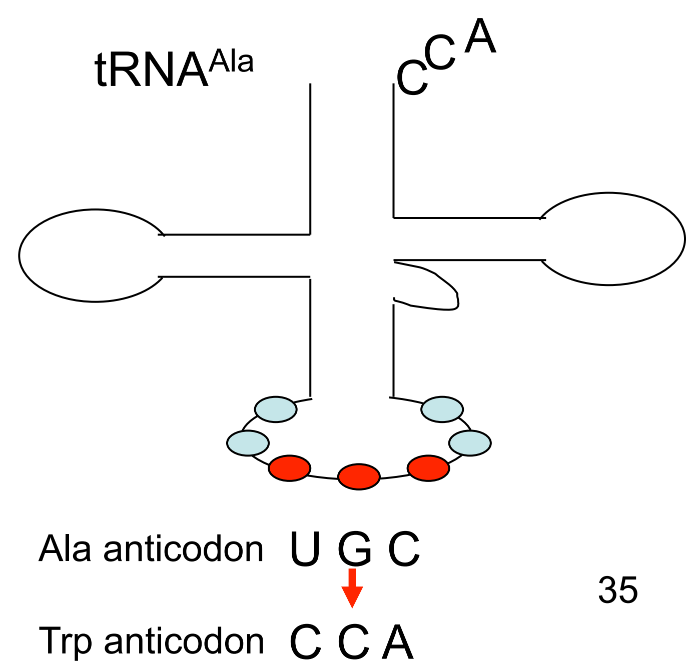
Kiga, et al. PNAS (2002)

Design of simplified genetic code in vitro

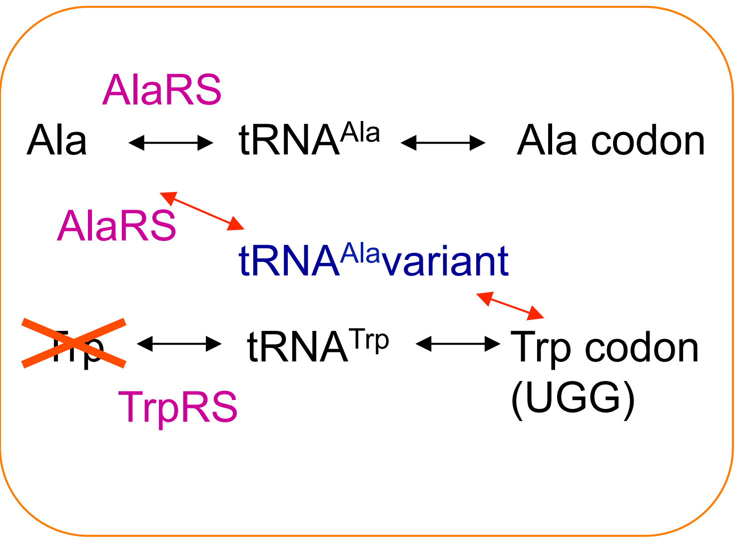
Reassign Trp codon to Ala



AlaRS doesn't recognize the anticodon loop to attach Ala.

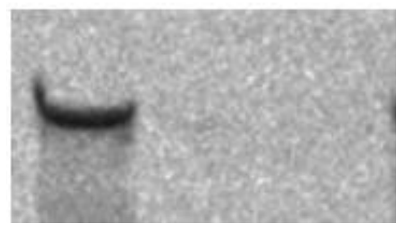


Sense codon "suppression" by $tRNA^{Ala}$ variant



^{14}C-labeled protein.

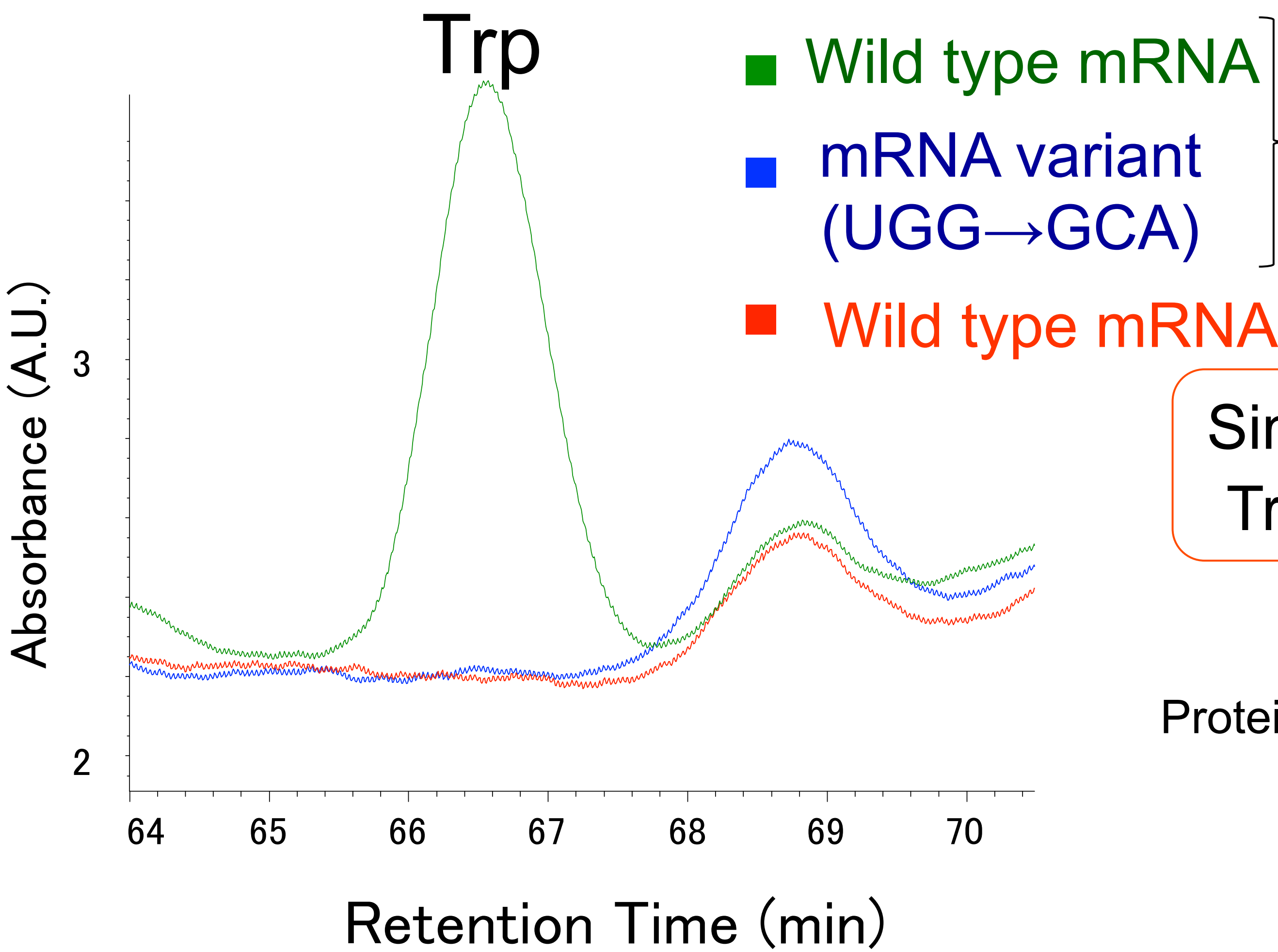
$tRNA^{Ala}$ variant (μM)	-	-	-
Trp	+	-	-
Trp-SA	-	-	+
Lane	1	2	3



Ala, not Trp, is likely to be inserted for UGG codons.

Next: amino acid composition analysis

Synthesized protein did not contain Trp.



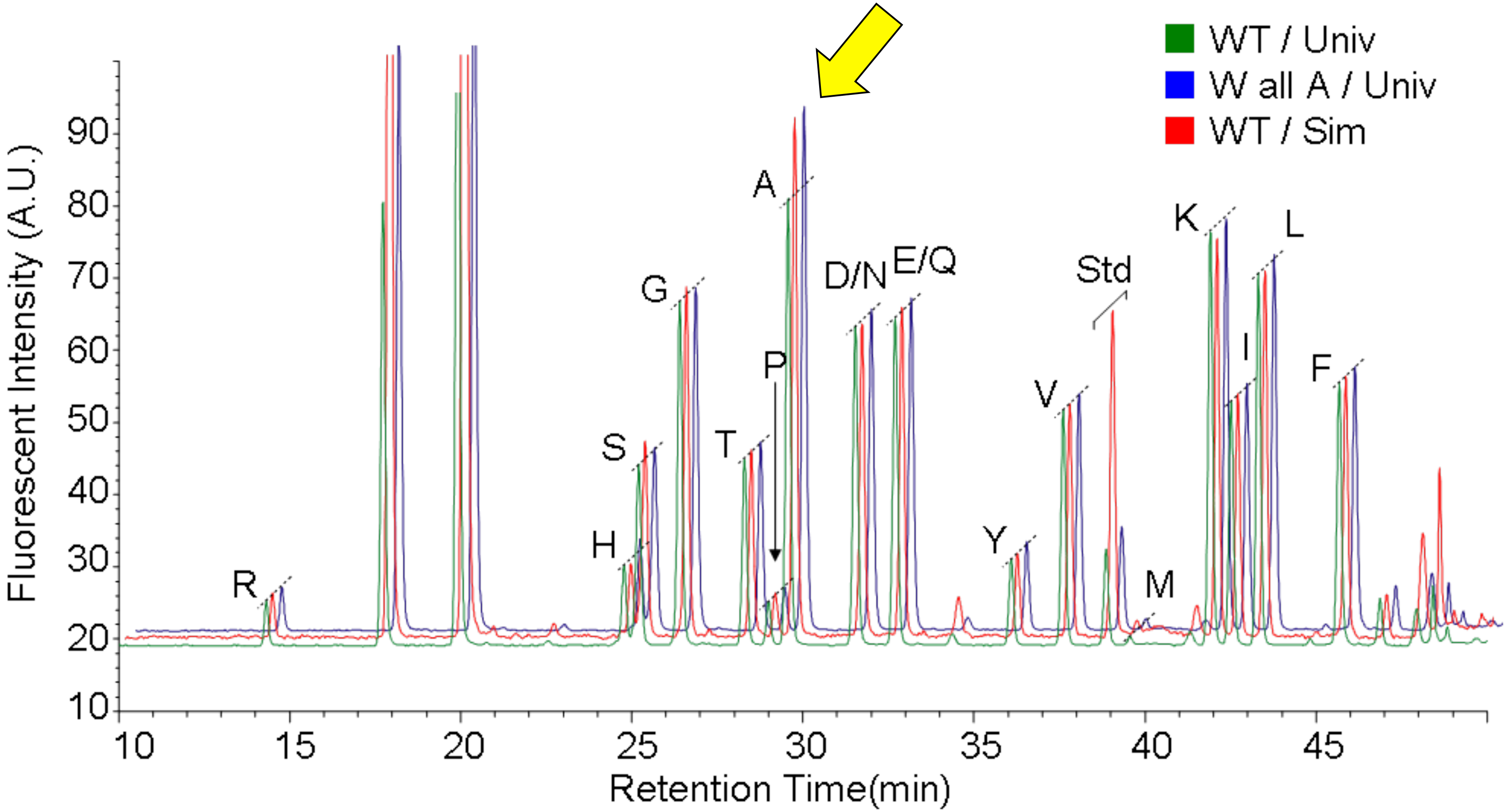
Universal code

Simplified code
Trp codon →Ala

Protein: MBP with 8 Trp

Ala was increased instead of Trp disappearance.

Amino acid composition analysis showed Ala and other amino acid's content.

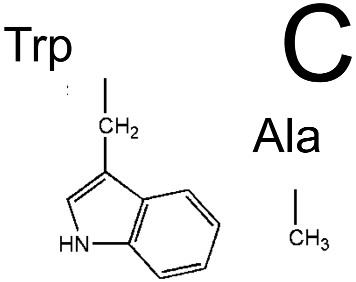


The increased amount of Ala was estimated as 8 residues, which is equivalent to the number of UGG codons in the MBP-WT mRNA.

The result strongly indicates that the tRNA variant introduced Ala to the UGG codon.

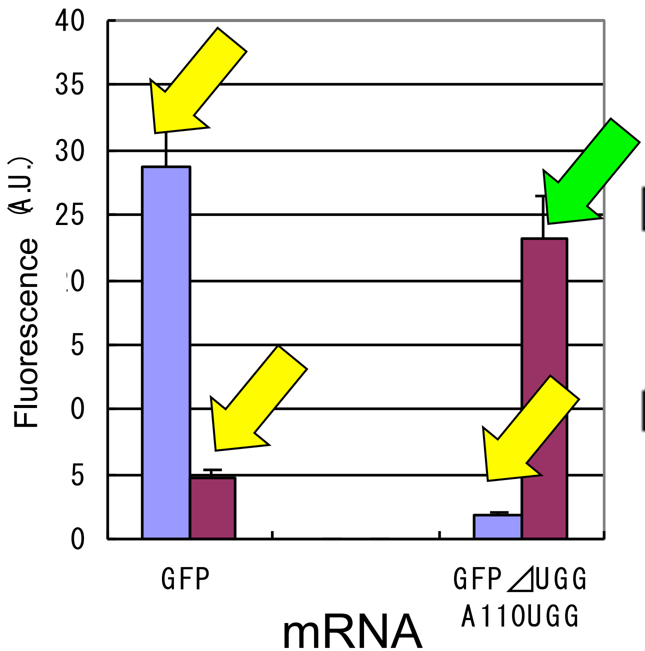
Results-I

Simplified genetic code Can produce active protein



Ala

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

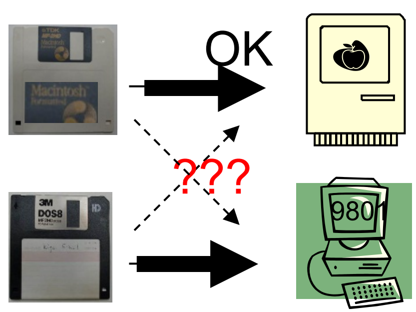


protein : GFP

Natural Universal code
(20 amino acid, UGG:Trp)

Engineered
Simplified code
(19 amino acid, UGG:Ala)

Gene specific for simplified genetic code



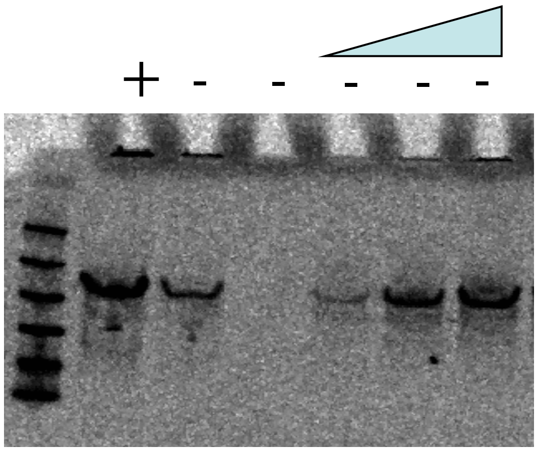
barrier against
horizontal transfer

Results-II

Generality of simplified genetic codes

Lys→Ser

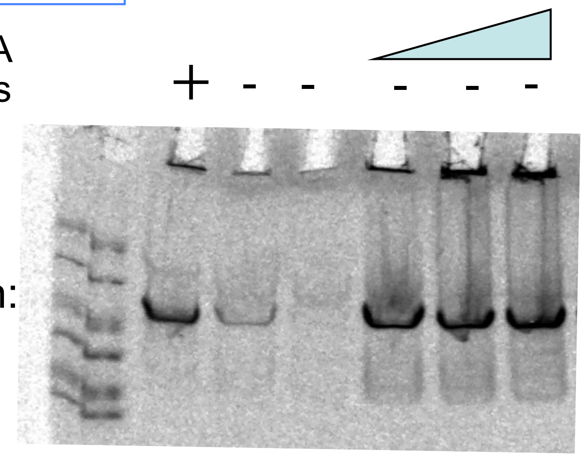
tRNA Lys + - - - - -



Protein :St.Av.

Cys→Ser

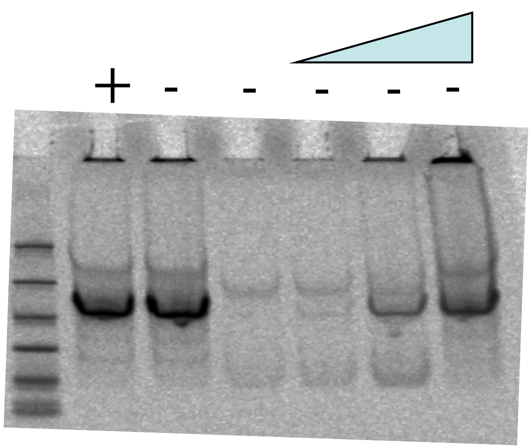
tRNA Cys + - - - - -



Protein: Ras

Asn→Ala

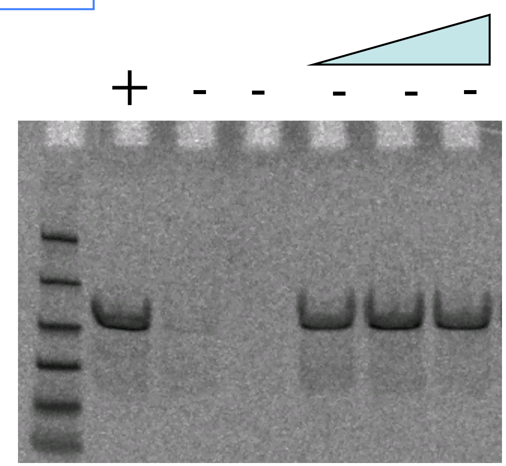
tRNA Asn + - - - - -



Protein: Ras

Trp→Ala

tRNA Trp + - - - - -



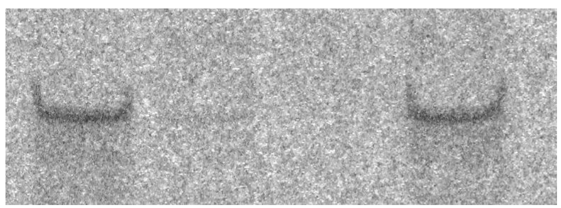
Protein :St.Av.

Cys→Ala, Asn→Ser, Trp→Ser Thr→Ser ….

The simplified genetic code comprising 16 amino acids

<u>Translation reaction with the tRNA^{Ser} variants in the *E.coli* cell extract lacking specific amino acids.</u>

tRNA variant	−	−	−	+
4amino acids	+	−	−	−
aa-SAs	−	−	+	+
Lane	1	2	3	4



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

tRNA variant : UGU/UGC:Ser, UGG:Ser, UAU/UAC:Ser, AAU/AAC:Ser

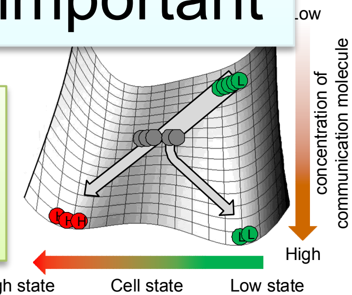
4amino acids : Cys, Trp, Tyr, Asn

aa-SAs : Cys-SA, Trp-SA, Tyr-SA, Asn-SA

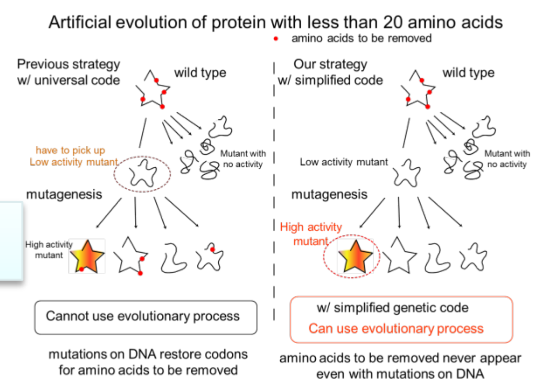
For resurrection of proteins before commonotes, engineered genetic code less than 20 amino acids is important

contents

Process for robust system

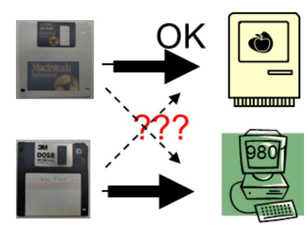


- Synthetic/Constructive approach in life science
  - proof by synthesis
  - construction of another life



- Simplified genetic code artificial evolution for reconstitution of ancient protein
  - code and protein with less than 20 amino acids

- How to engineer genetic code
  - Gene specific for simplified genetic code
  - 16 aa genetic code, and so on



- Q1: Find the secret message in the sentences.
  - Why is DNA necessary? Life is very special and cells are great things. Significantly advanced life science is easily mistaken as magic.
- Q2: Can I delete the “w” letter in these sentences?

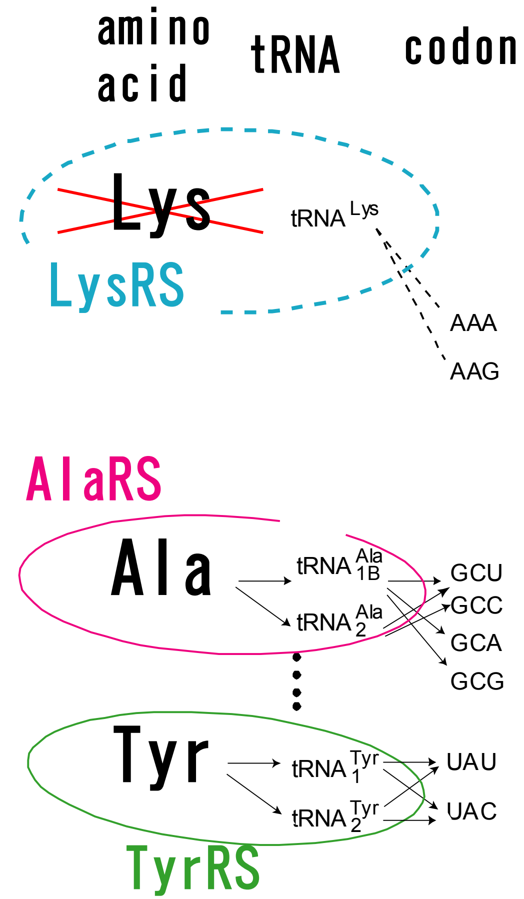
The above sentences use 20 of 26 letters in English alphabet

How about expression in German, Spanish, Japanese and so on?

Simplification of the code:

Construction of genetic codes with 19 amino acids

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

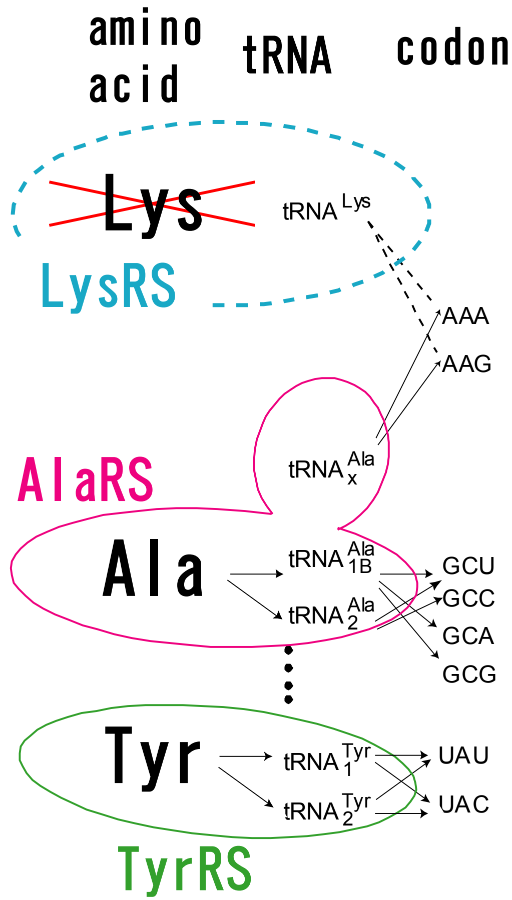


Removal of an amino acids ➡ Unassigned codon

Simplification of the code:

Construction of genetic codes with 19 amino acids

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



Removal of an amino acids → Unassigned codon

Addition of tRNA$^{Ala/Ser}$ mutant → Codon Reassignment

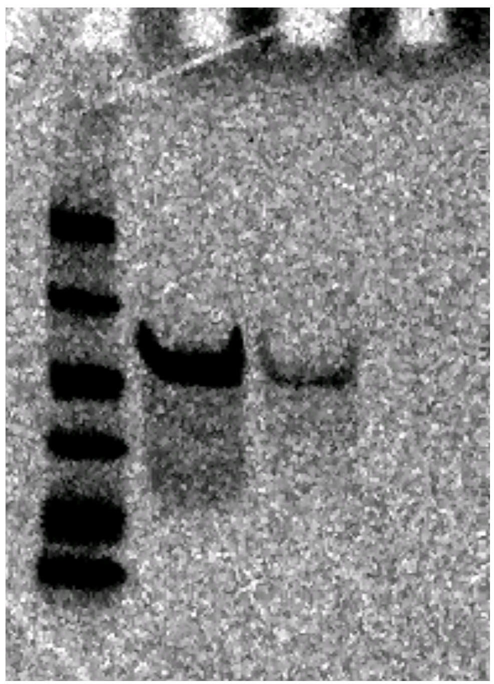
Simplification of the code:

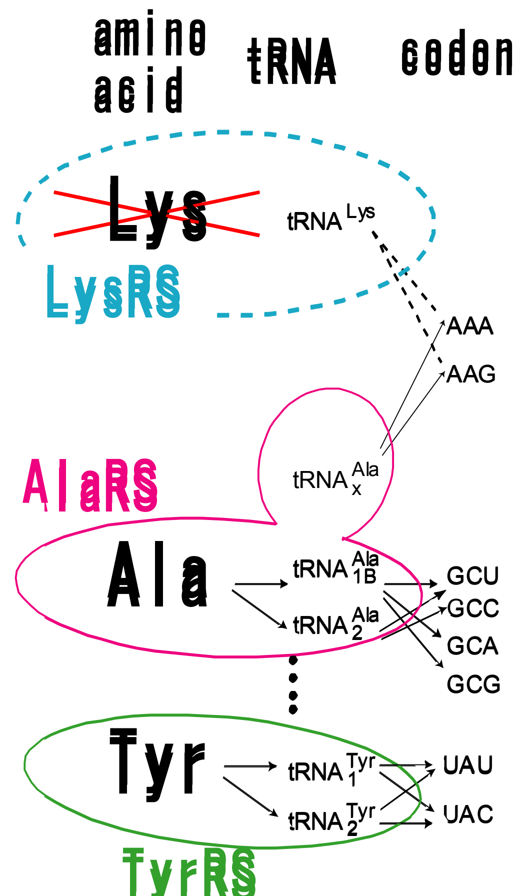
Codon reassignment by tRNA^{Ala} variant

Lys codon AAA/G →Ala

tRNA^{Ala} mutant

Lys	+	-	-
Lys-SA	-	-	+
laneNo	1	2	3



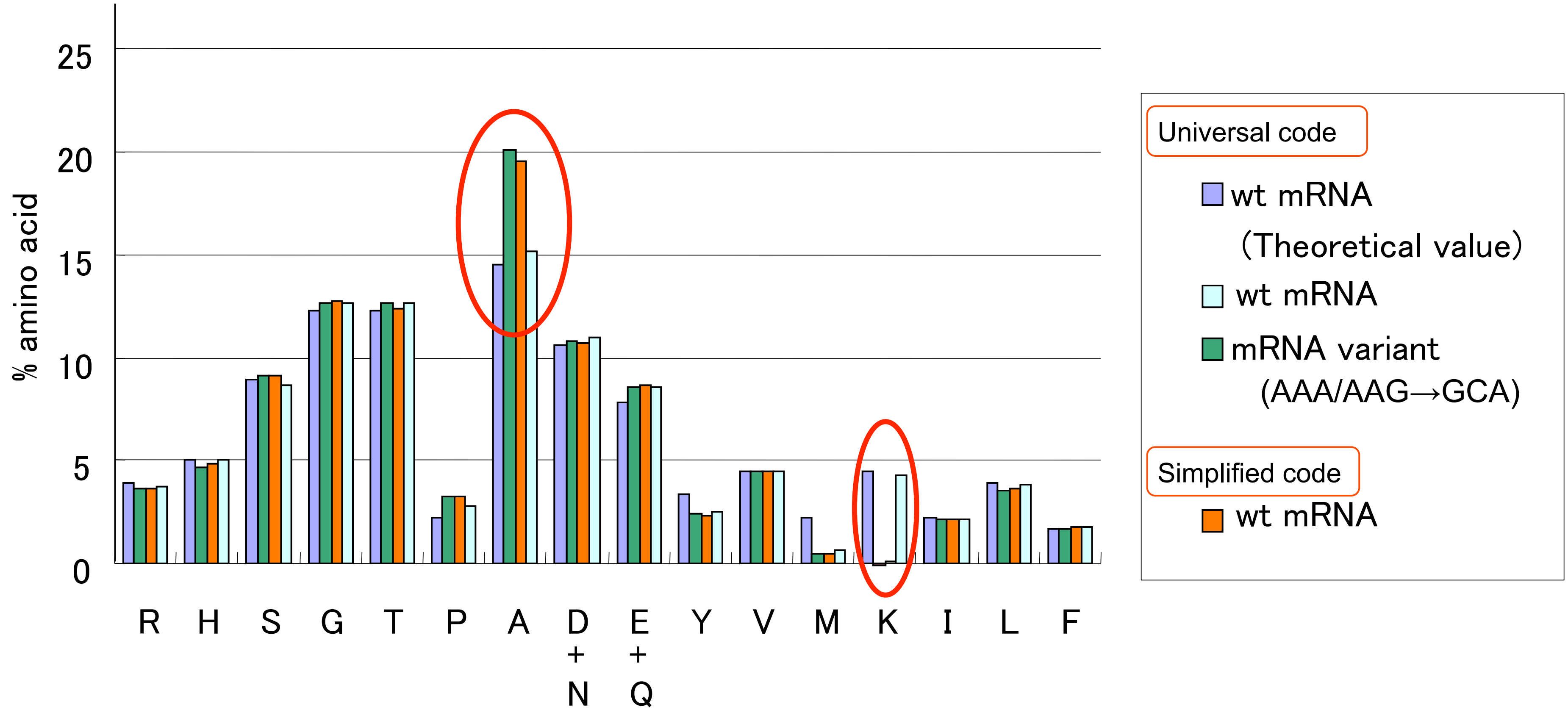


Simplification of the code:

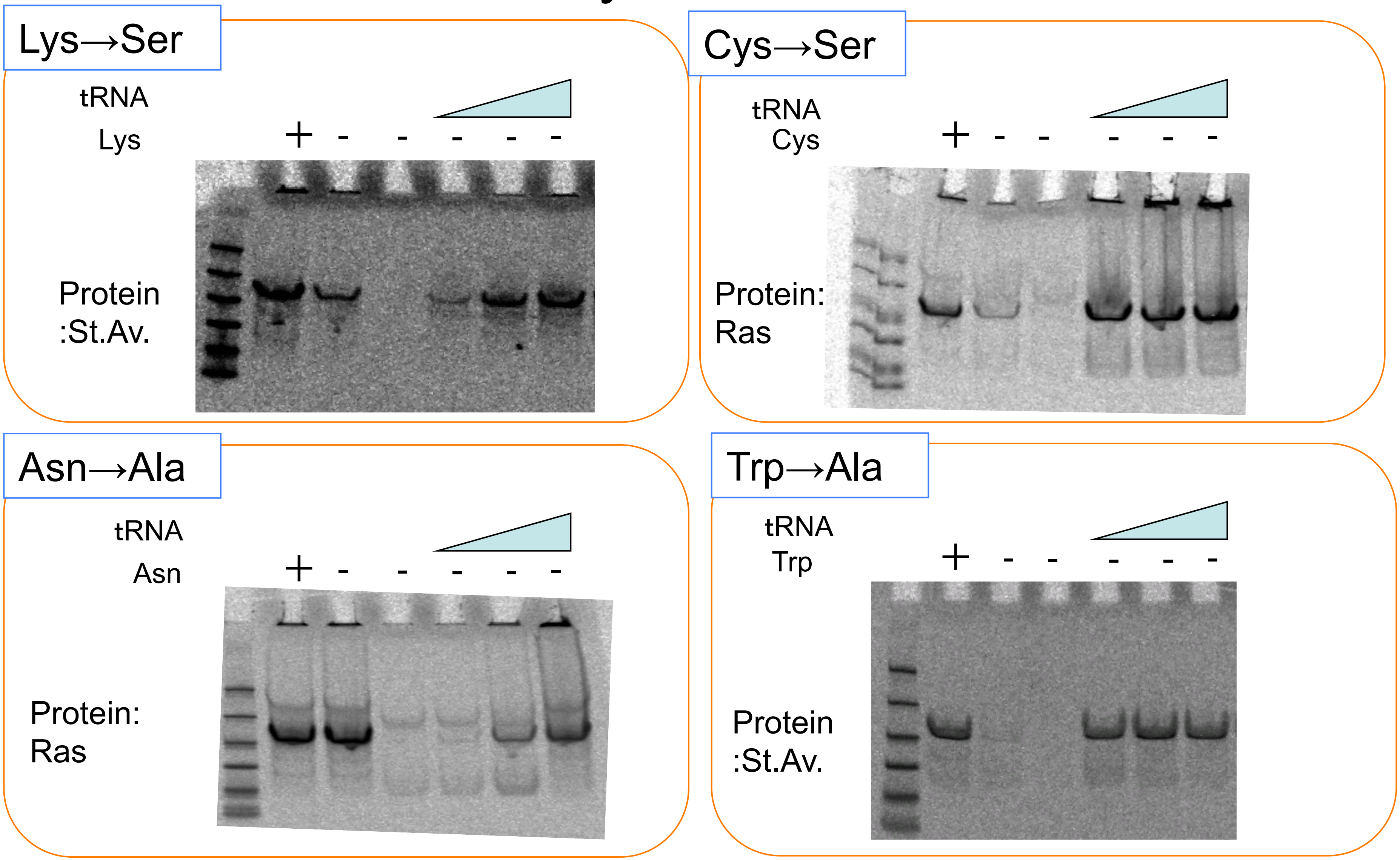
Amino acid composition analysis

w/ RIKEN Dohmae group

AAA/AAG codons are reassigned from Lysine(K) to Alanine(A)



Simplification of the code:
Generality of the method



Cys→Ala, Asn→Ser, Trp→Ser Thr→Ser ….

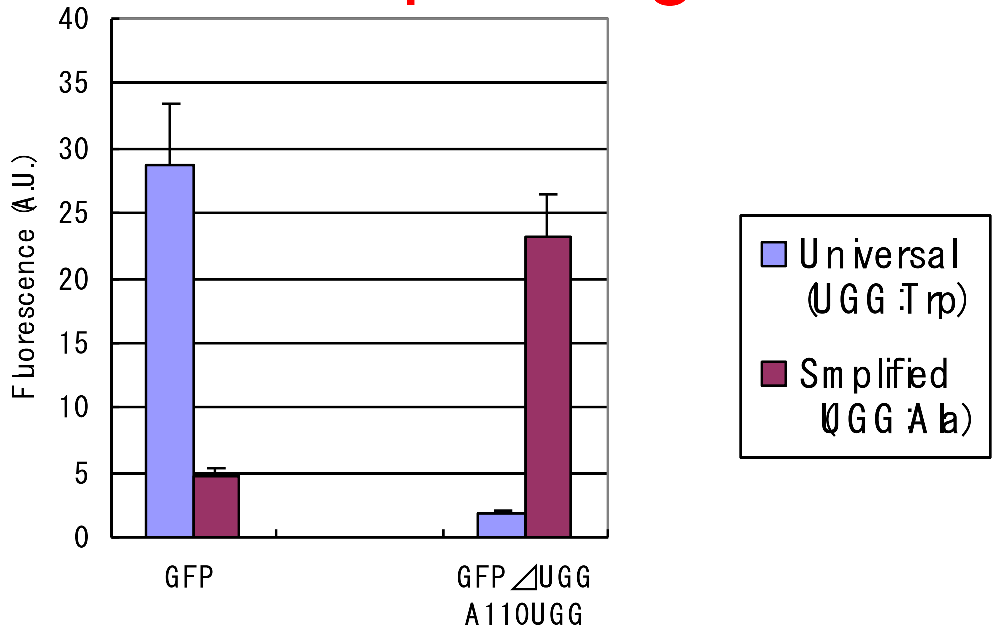
Simplification of the code:

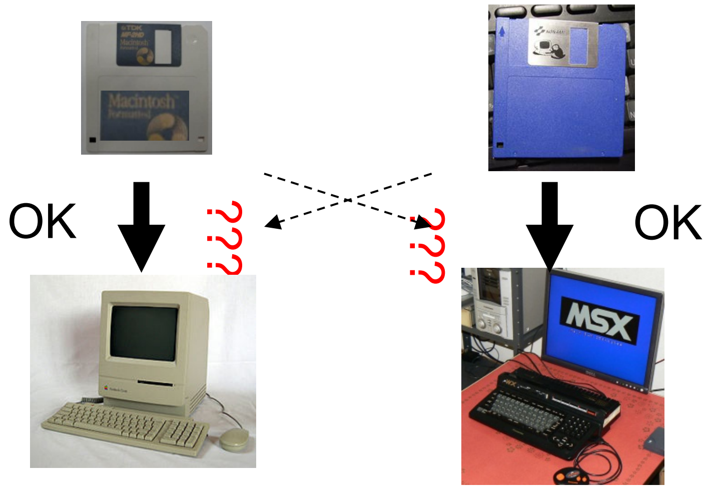
Simplified genetic code can produce active protein

Gene specific for simplified genetic code

Ala

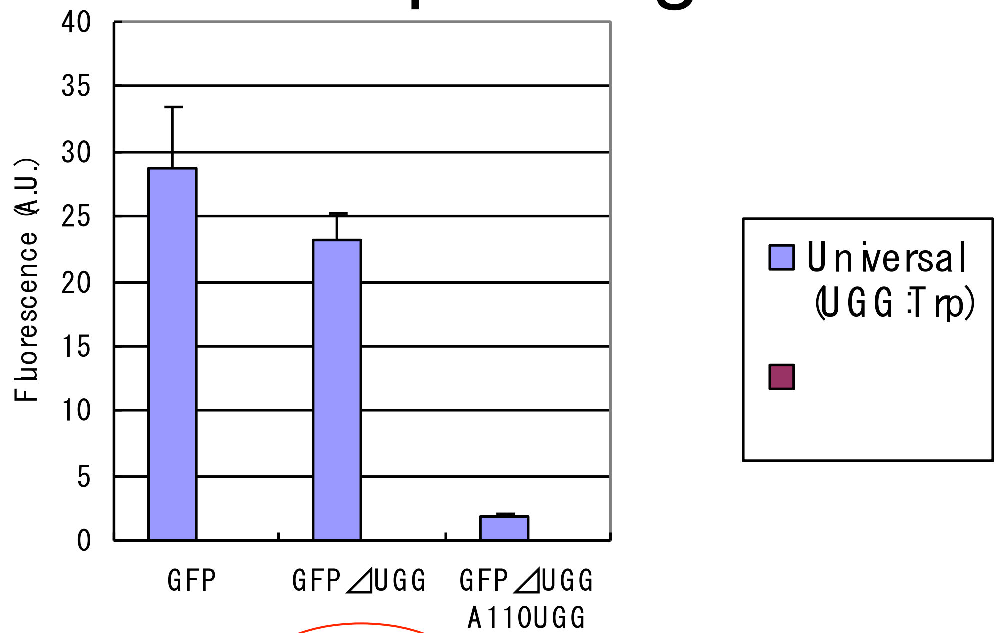
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
UUG	Leu	UCG	Ser	UAG	Stop	UGG	~~Trp~~ (Ala)
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser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





Simplification of the code:

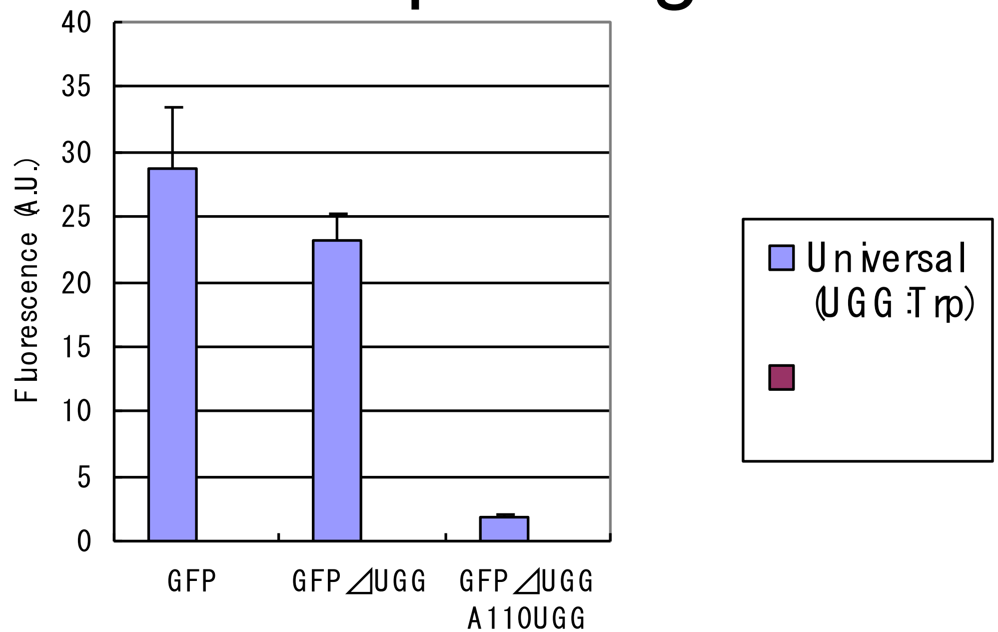
Gene specific for simplified genetic code



codon 57	UGG	UUU	UUU
amino acid required at 57	Trp	Phe	Phe
amino acid programmed at 57	Trp	Phe	Phe
codon 110	GCU	GCU	UGG
amino acid required at 110	Ala	Ala	Ala
amino acid programmed at 110	Ala	Ala	Trp

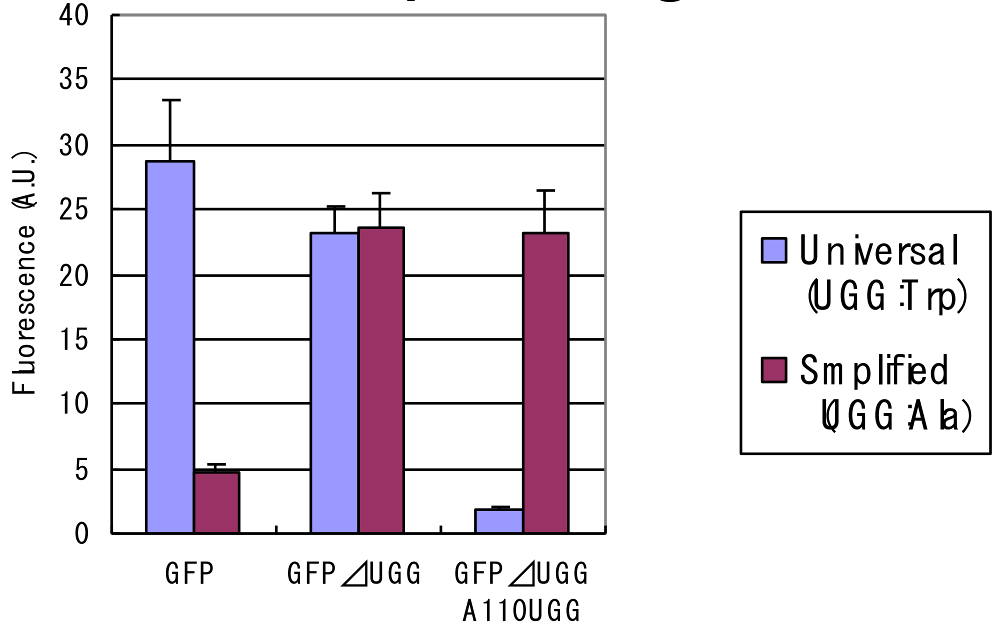
Simplification of the code:

Gene specific for simplified genetic code



codon 57	UGG	UUU	UUU
amino acid required at 57	Trp	Phe	Phe
amino acid programmed at 57	Trp	Phe	Phe
codon 110	GCU	GCU	UGG
amino acid required at 110	Ala	Ala	Ala
amino acid programmed at 110	Ala	Ala	Trp

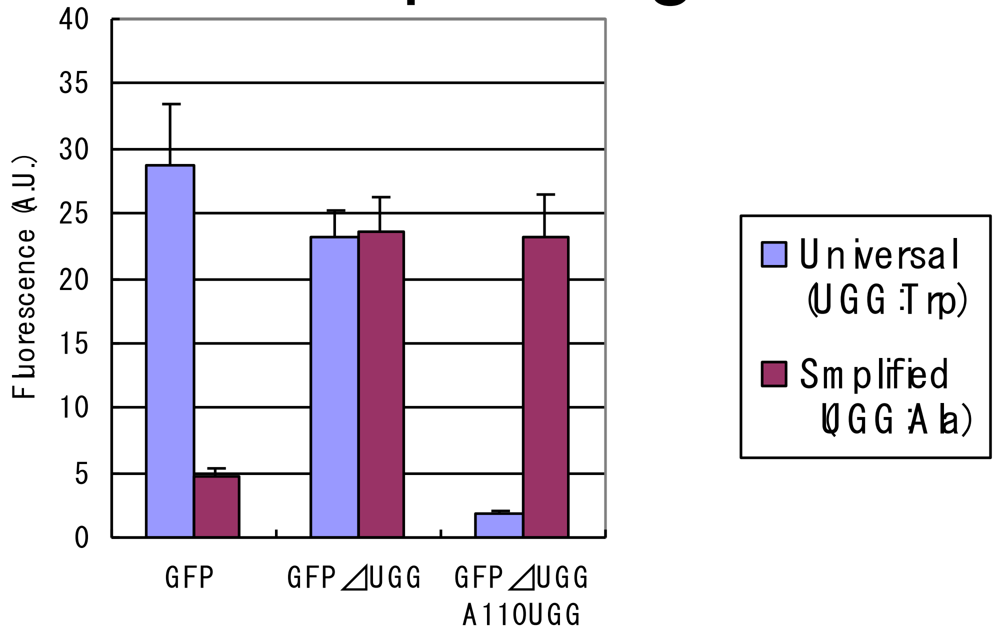
Gene specific for simplified genetic code



codon 57	UGG	UUU	UUU
amino acid required at 57	Trp	Phe	Phe
amino acid programmed at 57	Trp Ala	PhePhe	PhePhe
codon 110	GCU	GCU	UGG
amino acid required at 110	Ala	Ala	Ala
amino acid programmed at 110	AlaAla	AlaAla	TrpAla

Simplification of the code:

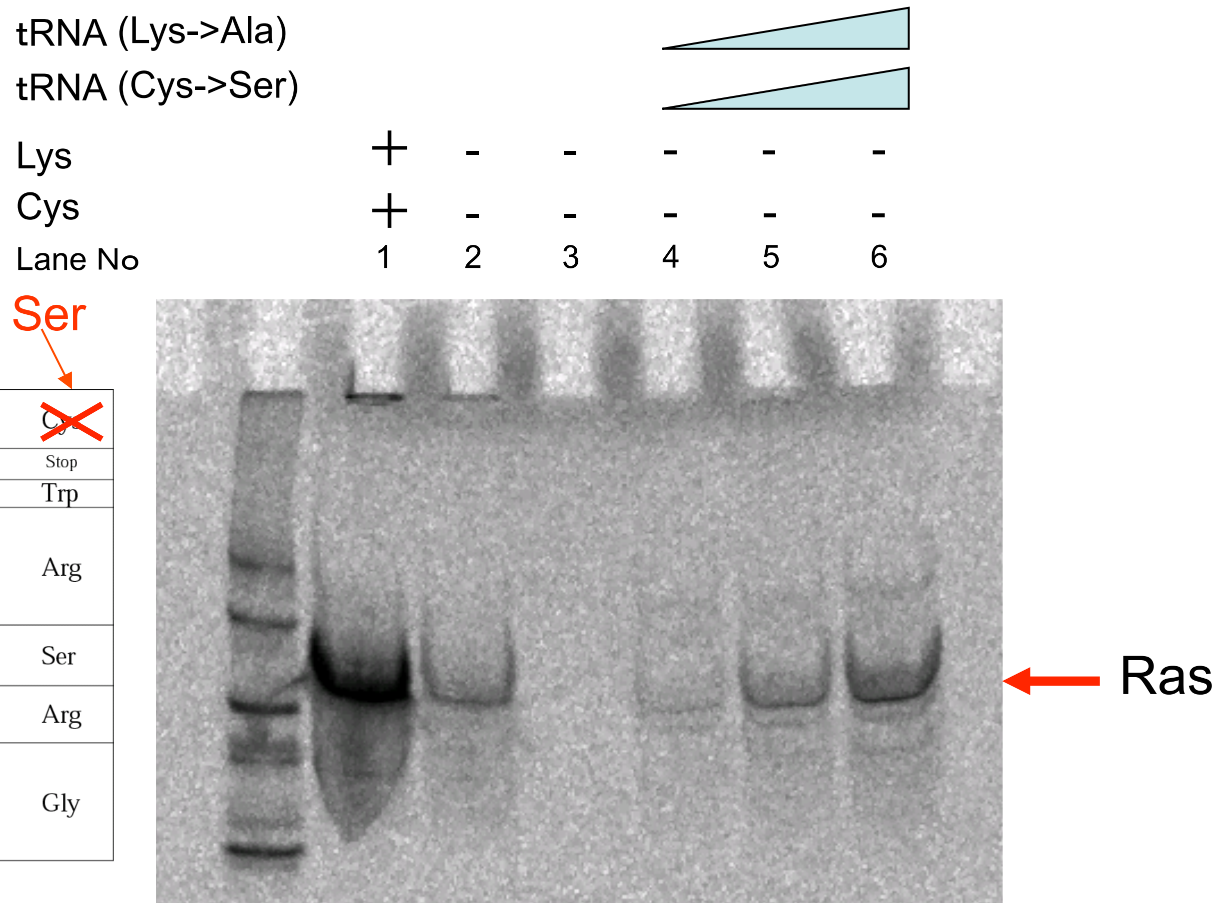
Gene specific for simplified genetic code



codon 57	UGG	UUU	UUU
amino acid required at 57	Trp	Phe	Phe
amino acid programmed at 57	Trp Ala	Phe Phe	Phe Phe
codon 110	GCU	GCU	UGG
amino acid required at 110	Ala	Ala	Ala
amino acid programmed at 110	Ala Ala	Ala Ala	Trp Ala

Simplification of the code:

Genetic code with 18 amino acids



Ser

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

Ala

Genetic Code: Conclusion

The "universal" number of amino acids
in the genetic code can be changed.

Future works

- Protein with less than 20 amino acids
  - evolution under different genetic codes
    - Screening System

- Cell with less than 20 amino acids ?
  - have to construct artificial cell from scratch

Number of amino acids and aminoacyl tRNA synthetases

Canonical case : 20 aaRS :20 amino acids

19 aaRS : 20(=19+1) AA

Organisms without Glutaminyl tRNA synthetase
→ transfer of amino group to Glu on tRNA

20 aaRS : 20 AA + additional AA

Selenocysteine incorporation
→ transfer of selenium to Ser on tRNA

21 aaRS : 20 AA + additional AA

Pyrrolysine incorporation (in nature)
Iodotyrosine incorporation (engineered)

19 aaRS : 19 amino acids

→ In a past form of life,
less than 20 amino acids would be assigned in a genetic code

Objectives of Synthetic biology

- Engineering
  - protein engineering
  - metabolic engineering
  - tissue engineering

- Science
  - Reconstitution: proof-by-synthesis
    - ATPase, RNA Pol, Ribosome
    - PURE Translation system
    - Regulatory circuit
  - Construction of "another life"

We still cannot construct "cell" w/ 20 aa

Japanese Society for Cell Synthesis Research

President 2010 Shoji Takeuchi
President 2012 Daisuke Kiga

Japanese Society for Cell Synthesis Research

Umeno H.R.Ueda Itaya

Kiga

Gene Network

Genome Engineering

Yomo

Shimizu

T. Ueda

Protein synthesis

Cell synthesis

Observation

Suga

Toyoda

Noji

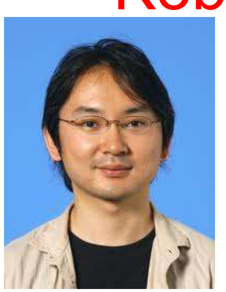
Information Processing

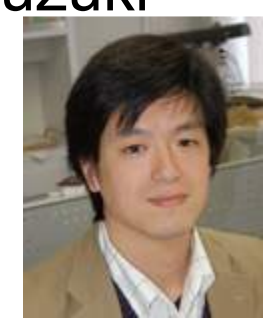
Compartment.

Suzuki

Kobayashi

Tabata





Iwasaki

Interactions w/ Society

Takeuchi

Yoshizawa Kato Hashimoto Hibino

Objectives of Synthetic biology

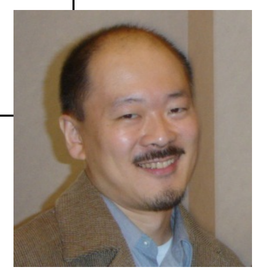
- Engineering
  - Protein engineering
  - metabolic engineering
  - tissue engineering

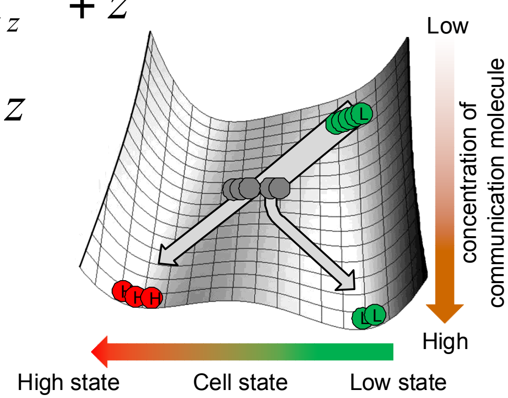
$$\frac{dx_i}{dt} = \frac{\alpha_x}{1 + \left(y_i / K_y\right)^{ny}} - d_x \cdot x_i$$

$$\frac{dy_i}{dt} = \frac{\alpha_y}{1 + \left(x_i / K_x\right)^{nx}} \cdot \frac{z^{nz}}{K_z^{\ nz} + z^{nz}} - d_y \cdot y_i$$

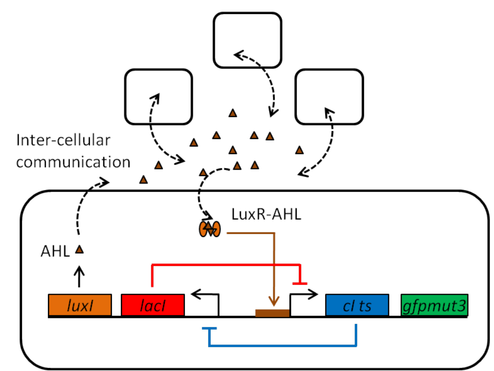
$$\frac{dz}{dt} = \frac{N_{tot}}{N} \sum \lambda \cdot x_i - d_z \cdot z$$

Kiga Lab
PNAS (2011)
vol108, 17969-

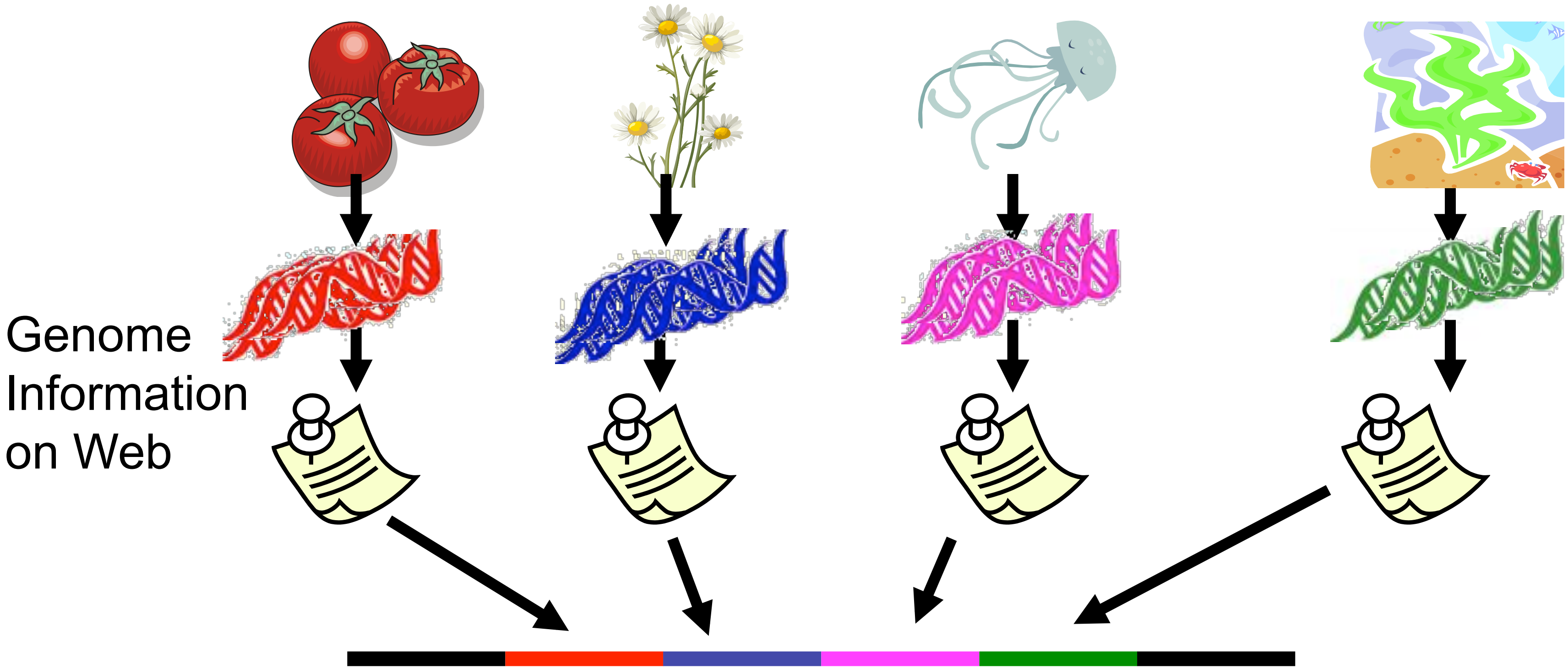




- Science
  - Reconstitution:
    proof-by-synthesis
    - ATPase, RNA Pol, Ribosome
    - PURE Translation system
  - Regulatory circuit



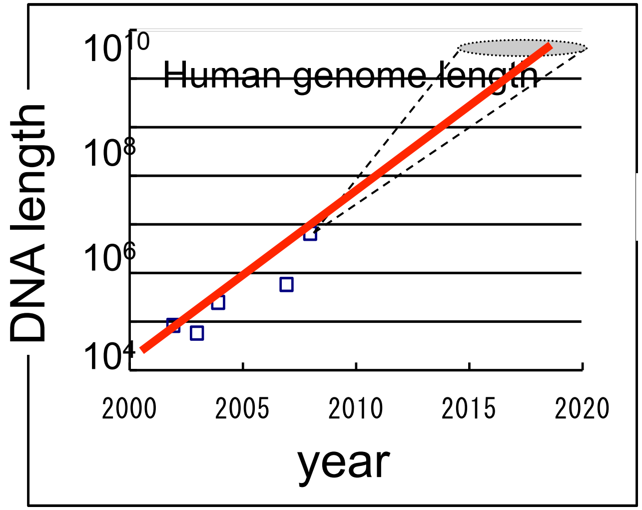
Construction of Artificial Genetic Circuit by Copy-and-Paste from DNA sequence of Organisms



Design of DNA sequence

Developments in preparation of biomolecules

Moor's Low in DNA synthesis: exponential growth



Kiga and Yamamura, New Generation Computing, 26 (2008), p347-364

biological synthesis：copy template
chemical synthesis：any sequence

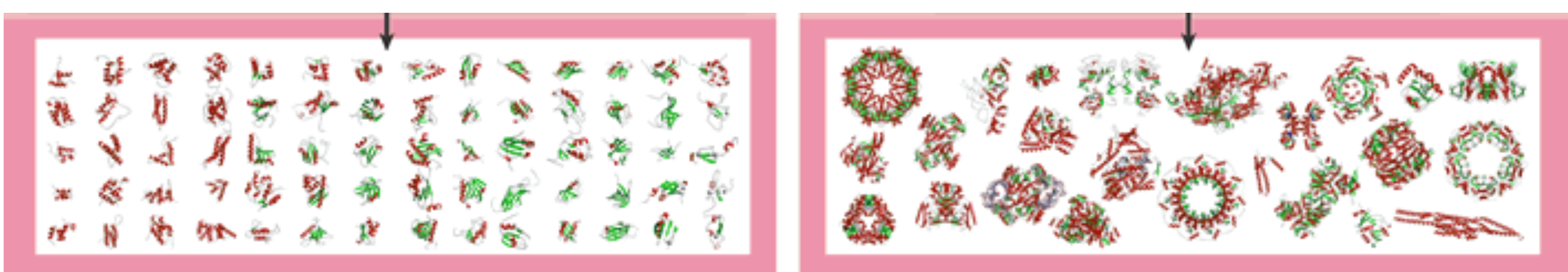
2 JULY 2010 VOL 329 SCIENCE 52

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

J. Craig Venter[1,2*]

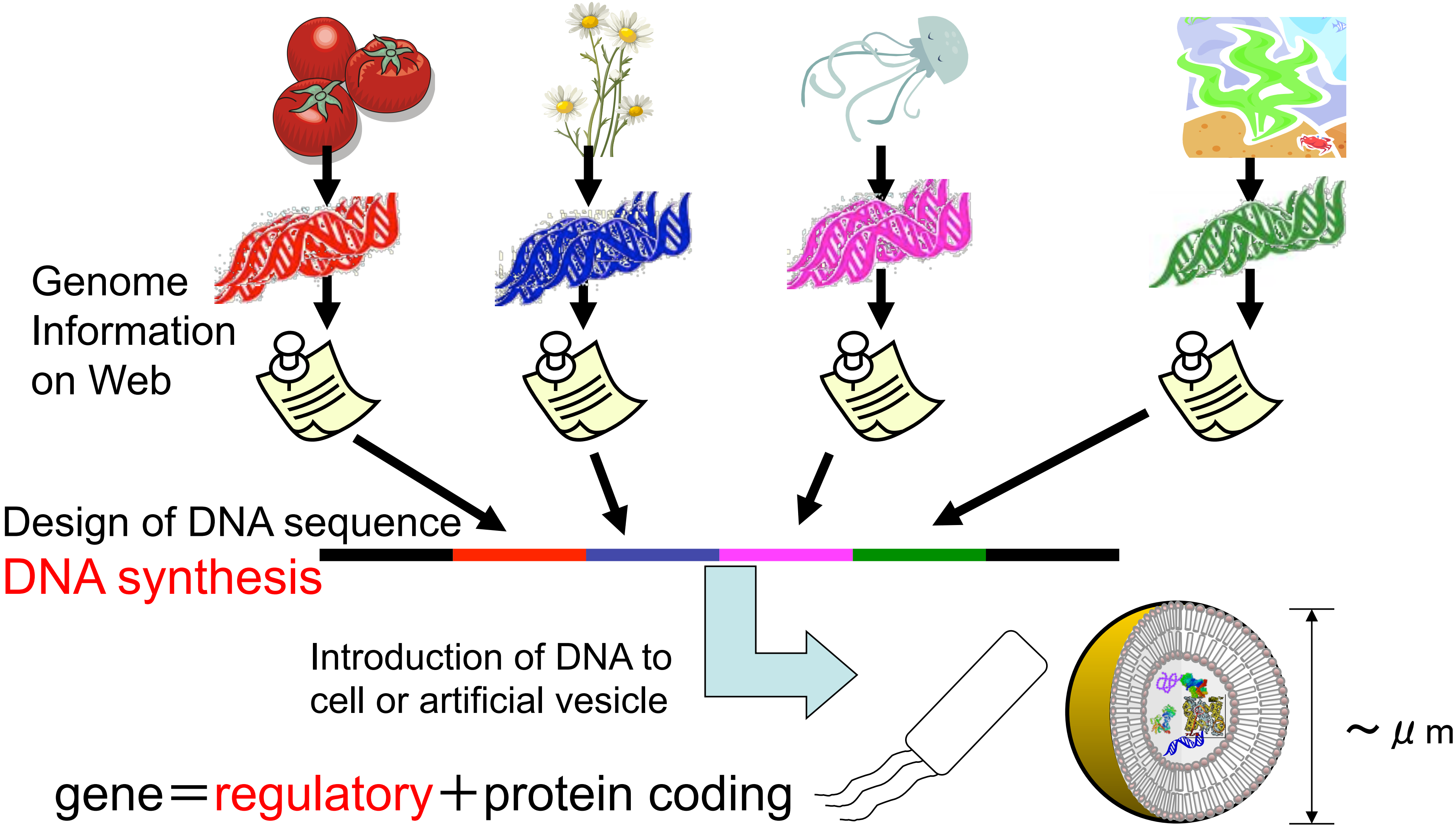
2015-2020 artificial DNA
with human genome length!?

Preparation of proteins



Construction of Artificial Genetic Circuit by Copy-and-Paste from DNA sequence of Organisms

have to design regulatory program



Programming in computer and life

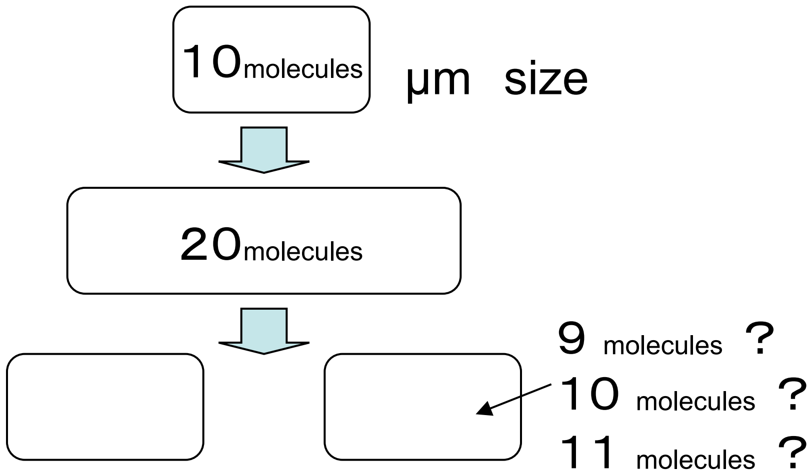
- computer

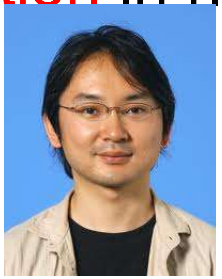
- Fix 0-1 state in a bit
- High energy consumption

- life

- Low energy consumption
- unable to fix its state

- Reaction in tiny space containing a ***small number*** of molecules
  →fluctuation in number



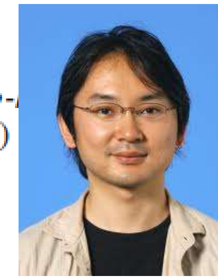


PRL **104,** 228104 (2010) PHYSICAL REVIEW LETTERS week ending 4 JUNE 2010

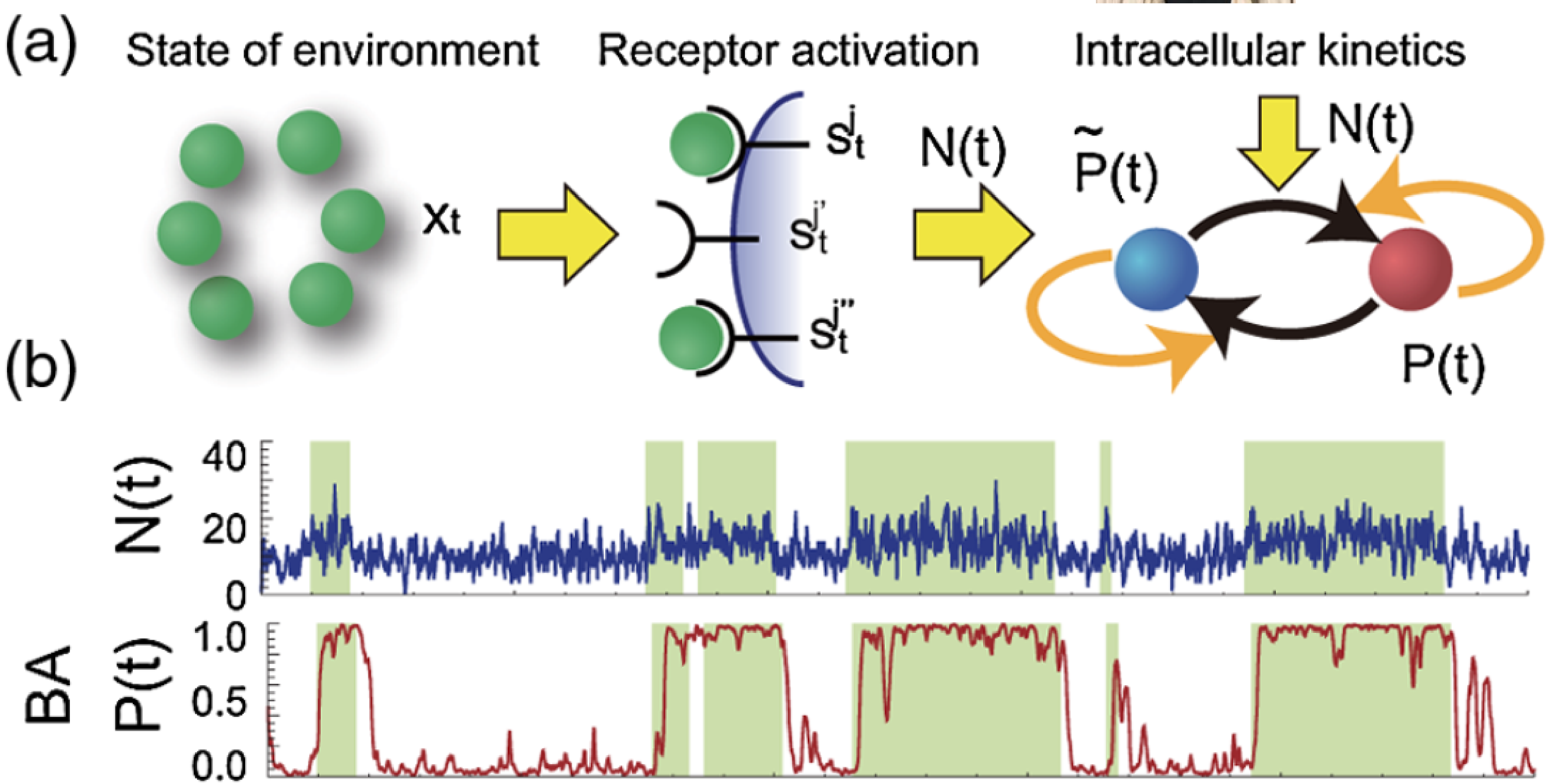
Implementation of Dynamic Bayesian Decision Making by Intracellular Kinetics

Tetsuya J. Kobayashi*

Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba Meguro- *8505, Japan*



(Received 20 January 2010; published 3 June 2010)



Programming in computer and life

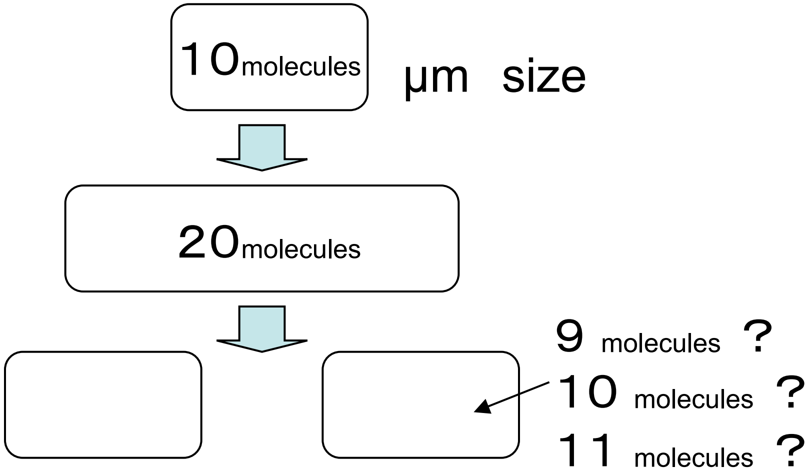
- computer

- Fix 0-1 state in a bit
- High energy consumption

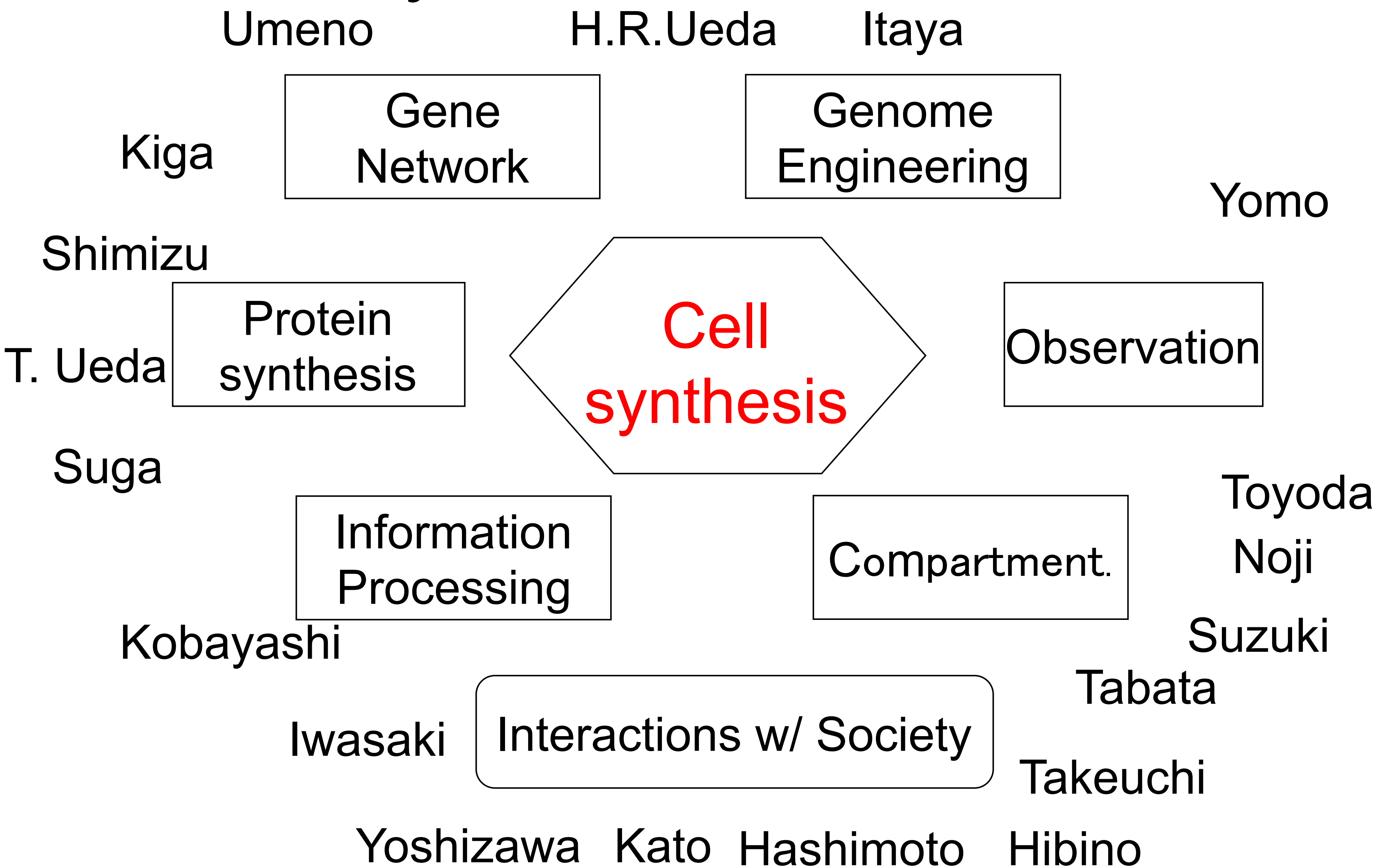
- life

- Low energy consumption
- unable to fix its state

- Reaction in tiny space containing a ***small number*** of molecules
  →fluctuation in number



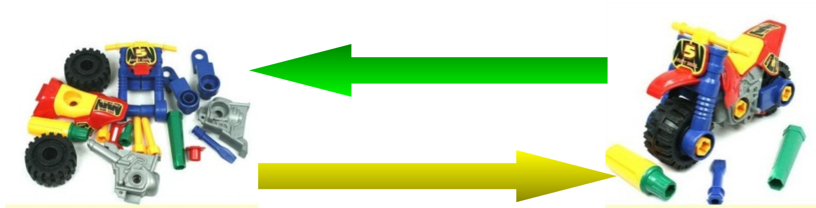
Japanese Society for Cell Synthesis Research



Objectives of Synthetic biology

- Engineering

- Science
  - Reconstitution:

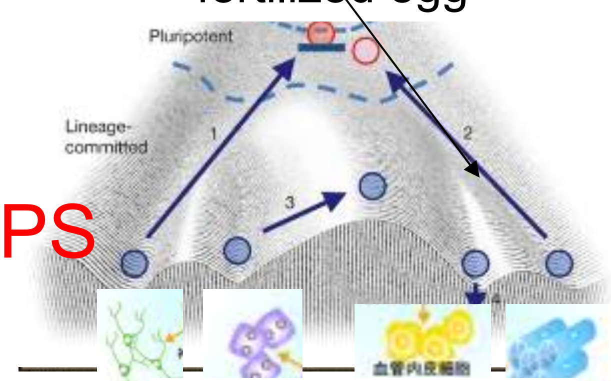


Epigenetic landscape

fertilized egg

one phenotype
↓
multi phenotypes

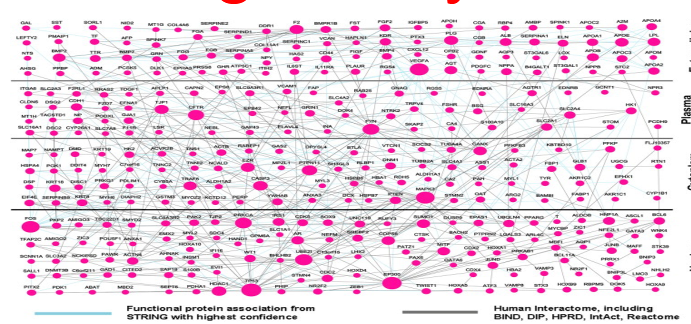
iPS



differentiated cells

S Yamanaka *Nature* (2009)

- Regulatory circuit



Molecular Interaction Network in Human Organogenesis
Fang et al , Dev Cell 2010