Theoretical Molecular Biology

What is life?

Koichi Itoh M.D.,Ph.D.

The Institute for Theoretical Molecular Biology

21-13, Rokurokuso-cho, Ashiya, Hyogo, JAPAN 659-0011

http://www.i-tmb.com/

e-mail: itoh@i-tmb.com

Table of contents

Preface

This book is an introductory textbook and philosophy book. I think that academic fields need philosophy. Around 1900, physicists discussed about the quantum theory and physics divided into experimental physics and theoretical physics. In the case of biology, especially molecular biology, each scientist has been played a role as an experimental biologist and theoretical biologist for about 60 years. One of the reasons, I think, is that fundamental principles have not been discovered. However, scientists who have the techniques and the knowledge of molecular biology have been produced enormous data and those data are stored in official databases. I think that the time is coming to build up the new academic field to analyze those enormous data. The storage of data will continue. That is why I have been waiting the timing to build up the new academic filed as theoretical molecular biology. There exists theoretical biology, but the new academic field of theoretical molecular biology has not been existed. I venture to build up theoretical molecular biology, because materials are made of molecules and living bodies are also made of molecules. In addition, I am a molecular biologist and medical doctor. Hence, my theoretical molecular

biology is for health of human beings and medicine. Firstly, what theoretical molecular biologists must do is to decide propositions or hypotheses. If these propositions or hypotheses are not correct, it must be impossible to get logical theories. Hence, theoretical molecular biologists must pay careful attention to decide the propositions or the hypotheses which really have the solutions. If propositions or hypotheses do not have the solutions, theoretical molecular biologists will not be able to get the solutions from the survey of enormous databases. Therefore, theoretical molecular biologists rely on what they have been experienced and meditated. In other words, theoretical molecular biologist must have their own perspectives on nature. Theoretical scientists should concentrate and think logically to give the solutions to the propositions which they judge to be able to solve, in terms of their own perspectives on nature. What theoretical molecular biologists should do is to search for the fundamental principles which control life phenomena, and logically to prove them. There are two great books which are ‘Molecular Biology of the Cell’ and ‘Molecular Biology of the Gene’ as the bibles for molecular biologists. However, textbooks are constructed from figures and tables and

explanation of those. Hence, the most part of the textbooks explain results which were investigated by the past experiments, and short comments or discussions are written about unknown things. The theoretical molecular biologists must meditate and think about already known things, things which are thought to be matters of course and unknown things, and once again try to elucidate the fundamental principles of life phenomena. The theoretical molecular biologists must prove the fundamental principles which control life phenomena to the utmost. To do these things, it is necessary to be familiar with other academic fields such as mathematics, physics, and chemistry in addition to molecular biology. This book is constructed from chapter 1. What is molecular biology? chapter 2. What is theoretical molecular biology? chapter 3. Is Genome the blueprint of life? chapter 4. The principle of ‘fluctuations’ fundamentally control life phenomena, chapter 5. Images of future medicine. Chapter 6. What is life? I wrote this book compactly to be able to read it from cover to cover. I was strongly impressed by Watson-Click’s double helix. I came to have a dream to determine human genome sequences and to cure every disease to correct incorrect sequence to correct sequence and to get the Nobel Prize in

Physiology or Medicine. I was born in 30th September 1964. I graduated Osaka University Medical School and Osaka University Graduate School of Medicine. After graduation, I studied in Howard Hughes Medical Institute, Harvard Medical School. After coming back to Japan, I have been thinking about founding a new academic field as theoretical molecular biology. The concept is 1) to meditate and determine propositions or hypotheses, 2) to prove logically those propositions or hypotheses without doing experiments, 3) to get the solutions.

Chapter 1. What is Molecular Biology?

Before 1953, researches about molecules in living bodies had been doing, but Watson-Click's double helix hypothesis in 1953 was the beginning of molecular biology. Molecular biology was different from mathematics or physics which were described by mathematical formulas. This must be profoundly recognized as special characteristics in molecular biology. There was the fundamental thought that molecular biologists elucidated life phenomena to investigate action and function of molecules. Therefore, materials of researches were mainly bacteria, yeasts, worms and mice as models for researches. However, final aim of we life scientists and medical scientists must be elucidations of the fundamental principles of life phenomena in human beings. The missions for life scientists and medical scientists must contribute for the health of human beings. Richard Feynman who is the Nobel Prize winner in Physics said that if, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, he believe it is the atomic hypothesis that all things are made of atoms. Molecular biology is also the academic field to investigate the action and the function of molecules in

living bodies. The aim of molecular biology is to elucidate the fundamental principles in life phenomena. Molecular biology has been enormously contributing life science. Taking a long look at the indexes of ‘Molecular Biology of the Cell’ and ‘Molecular Biology of the Gene’, the progress of molecular biology is easily recognized in these 60 years. The mechanism of duplication of DNA, the sequences of genomes in many living bodies, the mechanism of gene expression, apoptosis, polymerase chain reaction, three dimensional structures of proteins, RNA interference, expression profiles from microarrays. These discoveries and inventions have been enormously contributing to life science and medicine. I think that the time is coming to elucidate fundamental principles in life science and medicine utilizing enormous data which were stored by experiments without doing experiments. Molecular biology must be divided into experimental molecular biology and theoretical molecular biology.

Chapter 2. What is theoretical molecular biology?

I define theoretical molecular biology as the molecular biology getting solutions to prove logically propositions which are significant in medicine without doing experiments. Firstly, the most important step is to meditate and determine appropriate propositions. If appropriate propositions were not determined, it is impossible to get solutions. Therefore, the most important step is to meditate and determine propositions which really have solutions. If the propositions were fundamental and significant in life science and medicine, it is possible to contribute to health of human beings. Theoretical molecular biologists must get solutions to prove them, meditating and thinking logically without doing any experiments. The data are already stored in National Center for Biotechnology Information (NCBI: http://www.ncbi.nlm.nih.gov/) and other official databases. These data are the precious property which many life scientists and medical scientists have been making efforts to investigate life phenomena. The databases in NCBI and other official databases have been improved to easy to be used and surveyed. From now on, user interface may be improved and data may be piled up. After determine propositions, theoretical molecular

biologists must meditate, think and analyze logically and profoundly data to get solutions using and surveying any materials such as databases, past experimental results, textbooks without doing experiments. Any materials are fine. This is theoretical molecular biology. There are many exceptions in biology different from mathematics and physics. Is it possible to describe life phenomena with mathematical formulas? In my opinion, it is impossible to describe life phenomena with mathematical formulas even if the analyzing ability in computers were improved and even if a lot of data were stored in the future. That is why living bodies are not machines. If machines go out of order, is it possible to repair them by themselves? Is it possible to describe human minds and emotions as mathematical formulas? Is it possible that machines accomplish evolution automatically? For these propositions, solutions are NO right now. Hence, the present aims of theoretical molecular biology are get solutions to propositions logically, not to try to describe life phenomena by mathematical formulas. Because it is unknown which life phenomena will be able to be described by mathematical formulas? Therefore, for the proposition which life phenomena are described by mathematical formulas, the solution is NO

right now. Scientists must not have illusions. Scientists must meditate, think and analyze propositions realistically. Hence, the aims of theoretical molecular biology are to give the solutions to propositions by logical thinking for the fundamental principles of life phenomena, not to make mathematical formulas.

Chapter 3. Theoretical analysis indicates human genome is not a blueprint and human oocytes have the instructions.

Abstract

Is Human Genome really a blueprint? If it is not a blueprint, how are human bodies constructed? This paper solves this hypothetical proposition. Firstly, I indicate 8 examples of important biological pathways and factors among house-keeping genes and proved that human genome is not a blueprint. Human Genome is storage of genes. Secondly, I proved that human oocytes have the instructions for development and differentiation. In this case, I used opened public database for expression profile of human oocytes. I selected 12700 genes which expressed in human oocytes. Among 12700 genes, more than 800 genes which are related to development and differentiation are expressed. Here I show that human genome is not a blueprint and human oocytes have the instructions.

Introduction

Human genome has been thought to be a blueprint, but what type of the

blueprint has been a mystery. Human genome project was over in 2003, and seven years are already passed, but the number of human genes still unknown. Analysis of human genomes has been continuously done, but the discussion which a human genome is a blueprint has not been done. Far from that, any traces of a blueprint are not found in human genomes. This may be evidence that a human genome is not a blueprint. The Watson-Click's DNA double helix is very beautiful. Hence, we life-scientists have been imprinted that a human genome is a blueprint. If we hypothesize that a human genome is a blueprint, what types of absurdity do emerge? And if a human genome is not a blueprint, what must be needed to construct human bodies? To solve these hypothetical propositions are the aim of this document. In the case of unicellular organisms such as *E.coli*, their genomes may play a role for blueprints. However, biological mechanisms of multicellular organisms such as *Homo Sapiens*, are much complex and it is difficult to contain all information as a blueprint in their genomes. Therefore, a human genome plays a role for storage of genes, and I think that human oocytes have the instructions and a fertilized egg selects necessary genes from that storage, and expresses

genes for development and differentiation.

Materials and Methods

Table I was made from NCBI database (http://www.ncbi.nlm.nih.gov/) and KEGG (http://www.kegg.jp/ja/). One hundred ninety six key words in Supplemental Table I were selected from reference3-7. Supplemental Table II was made from Supplementary Data 1, 2, 3 which were originally located in http://www.canr.msu.edu/dept/ans/community/people/cibelli_jose.html (Kocabas 2006). I re-locate Supplementary Data 1, 2, 3, in http://www.i-tmb.com/text.html. Supplementary Data 1 contains up-regulated genes in human oocytes, Supplementary Data 2 contains down-regulated genes in human oocytes, and Supplementary Data 3 contains uniquely expressed genes in human oocytes. I combined Supplementary Data 1, 2, 3, and eliminated duplicated genes. Finally, I got 12764 genes which expressed in human oocytes (Supplemental Table II). I surveyed 12764 genes with 196 key words and I selected 823 genes which are thought to be important in development and differentiation in GenBank release 175.0 (Supplemental Table III). Table II shows the number of

important genes for development and differentiation. Supplemental Table I, II and III are located in http://www.i-tmb.com/.

Results and Discussion

Human genome is not a blueprint. At first the definition of a blueprint must be determined. According to a dictionary, a blueprint for something is a plan or set of proposals that shows how it is expected to work. I scrutinized loci of genes for 8 important biological pathways and factors, and their loci are scattered all over the human genome at random (Table I). I think that a blueprint must have regularity, periodism, harmony, some types of patterns, consistency or beauty which a blueprint itself has. But there were not existed such things. On the contrary, more than half of human genome sequence consists of Lines, Sines, retroviral-like elements, DNA-only transposon fossils, *Alu* sequences and pseudogenes (Alberts 2008). The loci of genes for 8 pathways and factors are scattered all over the human genome, and there do not exist any operons such as in bacterial genomes. Some reports exist that genes that make a cluster in one-dimensional, construct a cluster in three-dimensional, but there are no report that

scattered genes in one-dimensional construct a cluster in three-dimensional (Schneider 2007). In mathematics, one opposite example is enough for proof. But biology has some exceptions. However, genes in Table I are biologically important genes, and if a human genome is a blueprint, 8 exceptions must not be permitted. Here, I logically show that a human genome is not a blueprint. Hence, how are human bodies constructed from a human genome which is storage of genes?

Human oocytes have the instructions. Before fertilization, human oocytes express genes. If a human genome is storage of genes, mRNAs which are important for development and differentiation must be expressed in human oocytes and translated into proteins as soon as fertilization begins. Therefore, I surveyed public databases and I found an expression profile in human oocytes. In that profile, there are 12700 genes, and among 12700 genes, I found more than 800 genes which are related to development and differentiation. In general, many sample data must be necessary for comparison of gene expression levels for statistical analysis. But in my case, I do not need statistical analysis. Because the importance is only in which certain types of genes are expressed in human oocytes. I think that

human oocytes play a major role because of the amount of genes related to development and differentiation. Essential genes for human development and differentiation such as *Oct3, Oct4* are not existed in Table II. But I do not think that it is critical. I just think that mRNAs of *Oct3, Oct4* did not hybridize on the microarray chips. Because the genes which must be expressed must be expressed in human oocytes. And because of RNA interference, some mRNA might be broken. However, the amount of genes in human oocytes related in development and differentiation indicates that human oocytes have the instructions. Definition of instruction must be done. Instructions are clear and detailed information on how to do something. In this point, I think that human oocytes have the simple instructions. If human oocytes do not have the simple instructions, where is the blueprint or the instructions? I already indicate that a human genome is not a blueprint. Hence, it is logical that human oocytes have the simple instructions because a human body begins to be built from only one cell, a fertilized egg. If other cells except for human oocytes give proteins or mRNAs from outside of human oocytes, nurse cells or stromal cells might be candidates for the simple instructions. But it is not realistic that those

cells give most of biologically important proteins or mRNAs into fertilized eggs. Therefore, I logically proved that human oocytes have the simple instructions.

Important genes for the instruction in human oocytes (Gilbert 2006, Moody 2007, Schoenwolf 2009, Slack 2006, Wolpert 2007). The homeodomain is an approximately 60 amino acid sequence containing many basic residues, and forms a helix-turn-helix structure that binds specific sites in DNA. The homeodomain sequence itself is coded by a corresponding homeobox (HOX) in the gene. The homeobox was given its name because it was initially discovered in homeotic genes. However, there are many transcription factors that contain a homeodomain as their DNA-binding domain and although they are often involved in development, possession of a homeodomain does not guarantee a role in development, nor are mutants of homeobox genes necessarily homeotic. A very large number of homeodomain proteins have important functions, e.g. Engrailed in *Drosophila* segmentation, Goosecoid in the vertebrate organizer, Cdx proteins in anteroposterior patterning. An important subset are the HOX proteins which have a special role in the control of anteroposterior pattern

in animals. Homeobox genes are found in animals, plants, and fungi, but the Hox subset are only found in animals. The LIM domain is a cysteine-rich zinc-binding region responsible for protein-protein interactions, but is not itself a DNA-binding domain. LIM-homeoproteins possess two LIM domains together with the DNA-binding homeodomain. Examples are Lim-1 in the organizer, Islet-1 in motorneurons, Lhx factors in the limb bud, and Apterous in the *Drosophila* wing. PAXs are characterized by a DNA-binding region called a paired domain with 6 alpha-helical segments. The name is derived from the paired protein in Drosophila. Many of pax proteins also contain a homeodomain. Examples are Pax6 in the eye and Pax3 in the developing somite. Zinc-finger protein is a large and diverse group of proteins in which the DNA-binding region contains projections (“fingers”) with Cys and/or His residues folding around a zinc atom. Some examples are the GATA factors important of the blood and the gut, Krupple in the early *Drosophila* embryo, WT-1 in the kidney. Basic helix-loop-helix (bHLH) protein transcription factors are active as heterodimers. They contain a basic DNA-binding region and a hydrophobic helix-loop-helix region responsible for protein dimerization.

One member of the dimer is found in all tissues of the organism and the other member is tissue specific. There are also proteins containing the HLH but not the basic part of the sequence. These form inactive dimmers with other bHLH proteins and so inhibit their activity. Examples of bHLH proteins include E12, E47 which are ubiquitous in vertebrates, the myogenic factor MyoD, and *Drosophila* pair-rule protein hairy. An inhibitor with no basic region is Id, which is an inhibitor of myogenesis. FOX have a 100 amino acid winged helix domain which forms another type of DNA-binding region and known as “FOX” proteins. Examples are Forkhead in *Drosophila* embryonic termini and Fox2A in the vertebrate main axis and gut. T-box factors have a DNA-binding domain similar to the prototype gene product known as “T” in the mouse and as brachyury in other animals. They include the endodermal VegT and the limb identity factors Tbx4 and Tbx5. High mobility group (HGM)-box factors differ from most others because they do not have a specific activation or repression domain. Instead they work by bending the DNA to bring other regulatory sites into contact with the transcription complex. Examples are SRY, the testis-determining factor, Sox9, a “master switch” for cartilage

differentiation, and the TCF and LEF factors whose activity is regulated by the Wnt pathway. Trnasforming growth factor (TGF) beta was originally discovered as a mitogen secreted by “transformed” (cancer-like) cells. It has turned out to be the prototype for a large and diverse superfamily of signaling molecules, all of which share a number of basic structural characteristics. The mature factors are disulfide-bonded dimmers of approximately 25 kDa. They are synthesized as longer pro-forms which need to be protrolytically cleaved to the mature form in order for biological activity to be shown. The TGF-beta themselves are in fact often inhibitory to cell division and promote the secretion of extracellular matrix materials. They are involved mainly in the organogenesis stages of development. The activin-like factors include the nodal-related family, which are all involved in induction and patterning of the mesoderm in vertebrate embryos. The bone morphogenetic proteins (BMPs) were discovered as factors promoting ectopic formation of cartilage and bone in rodents. They are involved in skeletal development, and also in the specification of the early body plan. There are a number of receptors for the TGF-beta superfamily. Their specificity for different factors is complex and overlapping, but in general

different subsets of receptors bind to the TGF-beta themselves, the activin-like factors, and the BMPs. In all cases the ligand binds first to a type II receptor and enables it form a complex with a type I receptor. The type I receptor is a Ser-Thr kinase and becomes activated in the ternary complex. Activation causes phosphorylation of smad proteins in the cytoplasm. Smads 1, 5, and 8 are targets for BMP receptors; smad 2 and 3 for activin receptors. Smad 4 is required by both pathways, and smad 6 is inhibitory to both by displacing the binding of smad 4. Phosphorylation causes the smads to migrate to the nucleus where they function as for transcription factors, regulating target genes. The hedgehogs were first identified because mutations of the gene in *Drosophila* disrupted the segmentation pattern and made the larvae look like hedgehogs. Sonic hedgehog is very important for the dorsoventral patterning of the neural tube and for anteroposterior patterning of the limbs. Indian hedgehog is important in skeletal development. The full-length hedgehog polypeptide is an autoprotease, cleaving itself into an active N-terminal and an inactive C-terminal part. The N-terminal fragment is normally modified by covalent addition of a fatty acyl chain and of

cholesterol, which are needed for full activity. The hedgehog receptor is called patched, again named after the phenotype of the gene mutation in *Drosophila*. This is of the G-protein-linked class. It is constitutively active and is repressed by ligand binding. When active it represses the activity of another cell membrane protein, smoothened, which in turn represses the proteolytic cleavage of Gli-type transcription factors. Full-length Gli factors are transcriptional activators that can move to the nucleus and turn on target genes, but the constitutive removal of the C-terminal region makes them into repressors. In the absence of hedgehog, patched is active, smoothened inactive, and Gli inactive. In the presence of hedgehog, patched is inhibited, smoothened is active, and Gli is active. Activation of protein kinase A also represses Gli and hence antagonizes hedgehog signaling. The founder member of the Wnt family was discovered through two routes, as an oncogene in mice and as the wingless mutation in *Drosophila*. Wnt factors are single-chain polypeptides containing a covalently linked fatty acyl group which is essential for activity and renders them insoluble in water. The Wnt receptors are called frizzled after another Drosophila mutation. There are several classes of

receptor for different ligand types and they do not necessarily cross-react. Wnt 1, 3A, or 8 will activate frizzleds that cause the repression of a kinase, glycogen synthase kinase 3 (gsk3) via multifunctional protein called dishevelled. When active, gsk3 phosphorylates beta-catenin, an important molecule involved both in cell adhesion and gene regulation. When gsk3 is repressed, beta-catenin remains unphosphorylated and in this state can combine with a transcription factor, Tcf-1, and convey it into the nucleus. This pathway is important in numerous developmental contexts, including early dorsoventral patterning in *Xenopus*, segmentation in a *Drosophila*, and kidney development. Other Wnts, including Wnts 4, 5, and 11, bind to a different subset of frizzled that activate two other signal transduction pathways. In the planar cell polarity pathway a domain of the dishevelled protein interacts with small GTPases and the cytoskeleton to bring about a polarization of the cell. In the Wnt-Ca pathway phospholipase C becomes activated by a frizzled. This then acts to generate diacylglycerol and

inositol 1,4,5 triphosphate, with consequent elevation of cytoplasmic calcium, as described above under G-protein-coupled receptors. For the Delta-Notch system both the ligand (Delta, Jagged) and receptor (Notch) are integral membrane proteins. Their interaction can therefore only take place if the cells making them are in contact, as for the ephrin-Eph system. Binding of ligand to Notch causes cleavage of the cytoplasmic portion of Notch by an intramembranous protease, gamma-secretase, and this causes release into the cytoplasm of transcription factor, CSL-kappa. This migrates to the nucleus and activates target genes. The gamma-secretase is the same protease that generates the peptide whose accumulation in the brain leads to Alzheimer's disease. Notch can carry O-linked tetrasaccharides and presence of this carbohydrate chain can affect its specificity, increasing sensitivity to Delta and reducing sensitivity to Jagged. Control is often exercised through the activity of the glycosyl transferase Fringe, which adds GlcNAc to the O-linked fucose. The Delta-Notch system is important in numerous developmental situations, including neurogenesis, somitogenesis, and imaginal disc development. Cadherins are families of single-pass transmembrane glycoproteins which can adhere tightly to

similar molecules on other cells in the presence of calcium. Cadherins are the main factors attaching embryonic cells together, which is why embryonic tissues can often be caused to disaggregate simply by removal of calcium. The cytoplasmic tail of cadherins is anchored to actin bundles in the cytoskeleton by a complex including proteins called catenins. One of these, beta-catenin, is also a component of the Wnt signalimg pathway, providing a potential link named for the tissues in which they were originally found, so E-cadherin occurs mainly in epithelia and N-cadherin occurs mainly in neural tissue. The integrins are cell-surface glycoproteins that interact mainly with components of the extracellular matrix. They are heterodimers of alpha- and beta- subunits, and require either magnesium or calcium for binding. There are numerous different alpha and beta chain types and so there is a very large number of potential heterodimers. Integrins are attached by cytoplasmic domains to microfilament bundles, so, like cadherins, they provide a link between the outside world and the cytoskeleton. They are also thought on occasion to be responsible for the activation of signal transduction pathways and new gene transcription following exposure to particular extra cellular components.

After the birth of molecular biology, we life-scientists proved only two things, in my opinion. Firstly, there is high possibility that genes or proteins which have similar nucleic acid or amino acid sequences have similar 3-demensional structures and functions. Secondly, Genes or proteins have many functions because of the timing of working, permutation and combination. The number of human genes might be 40000 at most. In the first place, only 40000 genes cannot control complex biological mechanisms. Therefore, I think that limited number of genes and proteins change the timing of working, permutation and combination, and control the diverse biological mechanisms in human bodies. Genomes of viruses or bacteria might have the possibility that those genomes play a role for blueprints. But it will become impossible that human genome play a role for a blueprint. Hence, I think that human genome begins to exist as storage of genes. And human oocytes express essential genes for development and differentiation as the simple instructions. After fertilization, a fertilized egg differentiates according to micro-environment surround the fertilized egg. Therefore, human oocytes expresses genes for adhesion molecules such as integrins, cadherins and so on. From now on, a lot of evidence will be piled

up to support my hypothesis. Finally, I foresee that once organogenesis begins, tissue differentiation proceeds autonomously and human bodies are built. This is, I think, theoretical molecular biology and ‘Itoh hypothesis’.

References

Alberts B, Johnson A, Walter P, Lewis J, Raff M, Roberts K (2008). Molecular Biology of the Cell, 5th edition, Garland Science, Mortimer Street, London. 1-1601

Gilbert SF (2006). Developmental Biology, 8th edition, Sinauer Association Inc. Sunderland, MA. 3-751

Kocabas AM, Crosby J, Ross PJ et al, (2006). The transcriptome of human oocytes, Proc. Natl Acad. Sci. USA. 103:14027-14032.

Moody SA (2007). Principles of Developmental Genetics. Academic Press, New York. 2-1022

Schneider R, Grosschedl R (2007). Dynamics and interplay of nuclear architecture, genome organization, and gene expression. Genes Dev. 21:3027-3043.

Schoenwolf GC, Bleyl SB, Brauer PR, et,al, (2009). Larsen's Human Embryology, 4th edition, Churchill Livingstone, New York. 1-644

Slack JMW (2006). Essential Developmental Biology 2nd edition, Blackwell Publishing, West Sussex, UK. 3-336

Wolpert L (2007). Principles of Development, 3rd edition. Oxford University Press. 1-522

Table I. Loci of genes for major biological pathway

I. Glycolysis	
Gene Name	**Locus**
Glucokinase (Hekisokinase 4)	7p15-p13
Phosphoglucose isomerase	19q13.1
Phosphofructokinase, Liver Type	21q22.3
Phosphofructokinase, Muscle Type	12q13.3
Phosphofructokinase, Platelet Type	10p15.3-p15.2

VI. Purine biosynthesis	
Gene Name	**Locus**
amidophosphoribosyltransferase	4q12
phosphoribosylamine glycine ligase	21q22.1; 21q22.11
phosphoribosylglycinamide formyltransferase	21q22.1; 21q22.11
phosphoribosylformylglycinamidine synthase	17p13.1
phosphoribosylformylglycinamidine cyclo-ligase	21q22.1; 21q22.11

Aldolase A	16p11.2	phosphoribosylaminoimidazole carboxylase	4q12
Aldolase B	9q22.3	phosphoribosylaminoimidazole-succinocarboxamide synthase	4q12
Aldolase C	17cen-q12	adenylosuccinate lyase	22q13.1; 22q13.2
Glyceraldehyde 3-phosphate dehydrogenase	12p13.31-p13.1	phosphoribosyl aminoimidazole carboxamide formyltransferase	2q35
Phosphoglycerate kinase 1	Xq13	IMP cyclohydrolase	2q35
Phosphoglycerate mutase 2 (muscle)	7p13-p12	adenylosuccinate synthase	14q32.33
Phosphoglycerate	10q25.3	IMP	7q31.3-q32

mutase 1 (brain)	
Enolase 1, (alpha)	1p36.3-p36.2
Enolase 2 (gamma, neuronal)	12p13
Enolase 3 (beta, muscle)	17pter-p11
Pyruvate kinase, muscle	15q22
Pyruvate kinase, liver and RBC	1q21
II. TCA cycle	
Gene Name	**Locus**
Aconitase	22q11.21-q13.31

dehydrogenase	
GMP synthase	3q24
VII. Primidine biosynthesis	
Gene Name	**Locus**
carbamoyl-phosphate synthase	2p22-p21
aspartate carbamoyltransferase	2p22-p21
dihydroorotase	2p22-p21
dihydroorotate dehydrogenase	16q22
orotate phosphoribosyltransferase	3q13

Isocitrate dehydrogenase	15q26.1	orotidine-5'-phosphate decarboxylase	3q13
2-oxoglutarate dehydrogenase E1 component	7p14-p13	CTP synthase	1p34.1
2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase)	14q24.3	thymidylate synthase	18p11.32
succinyl-CoA synthetase alpha subunit	2p11.2	**VIII. Basal transcription factors**	
Succinate dehydrogenase	5p15	**Gene Name**	**Locus**
Fumarase	1q42.1	TATA-box-binding protein	14q22.3

Malate dehydrogenase	7cen-q22	transcription initiation factor TFIID subunit D1	9p21.1
Citrate synthase	12q13.2-q13.3	transcription initiation factor TFIID subunit D2	8q24.12
III. Pentose phosphate pathway		transcription initiation factor TFIID subunit D3	20q13.33
Gene Name	**Locus**	transcription initiation factor TFIID subunit D4	1q42.13
Glucose-6-phosphate dehydrogenase	Xq28	transcription initiation factor TFIID subunit D5	11q12.3
6-phosphogluconolactonase	19p13.2	transcription initiation factor TFIID subunit D6	Xq22.1

6-phasphogluconate dehydrogenese	1p36.3-p36.13	transcription initiation factor TFIID subunit D7	Xq13.1-q21.1
Ribrose 5-phosphate ketoisomerase	2p11.2	transcription initiation factor TFIID subunit D8	11p15.3
transketolase	3p14.3	transcription initiation factor TFIID subunit D9	5p15.1
transaldolase	11p15.5-p15.4	transcription initiation factor TFIID subunit D10	1p35.3
IV. Urea cycle		transcription initiation factor TFIID subunit D11	1p13.3
Gene Name	**Locus**	transcription initiation factor TFIIB	1p22-p21

Carbamoyl phoshpate synthase I	2q35	transcription initiation factor TFIIA large subunit	2p16.3
Ornithine transcarbamylase	Xp21.1	transcription initiation factor TFIIA small subunit	15q22.2
Argininosuccinic acid synthase	9q34.1	transcription initiation factor TFII-I	7q11.23
Argininosuccinase	7cen-q11.2	transcription initiation factor TFIIF alpha subunit	19p13.3
Arginase	6q23	transcription initiation factor TFIIF beta subunit	13q14

V. Fatty acid metabolism		transcription initiation factor TFIIE alpha subunit	3q21-q24
Gene Name	**Locus**	transcription initiation factor TFIIE beta subunit	8p21-p12
long-chain acyl-CoA synthetase	4q34-q35	transcription initiation factor TFIIH subunit H1	11p15.1-p14
acyl-CoA dehydrogenase	1p31	transcription initiation factor TFIIH subunit H2	5q12.2-q13.3
acyl-CoA oxidase	17q24-q25.1	transcription initiation factor TFIIH subunit H3	12q24.31

enoyl-CoA hydratase	10q26.2-q26.3
3-hydroxyacyl-CoA dehydrogenase	3q26.3-q28
long-chain 3-hydroxyacyl-CoA dehydrogenase	2p23
acetyl-CoA acyltransferase	18q21.1

transcription initiation factor TFIIH subunit H4	6p21.3

TableII. Genes for development and differentiation in human oocytes

Gene Group	Number of Genes	Gene Group	Number of Genes
Activin	6	lim	28
AKT	3	lin	4
armadillo	10	MAP	36
ATM	1	meltrin	1
BCL	25	mindbomb	1
BDNF	1	mix	1
beta-catenin	1	Myf	1
BMP	12	nanos	1
Cadherin	4	NCAM1	1
caspase	15	NENF	1
catenin	4	netrin	1
caudal	1	neuregulin	2
ced	7	neuropilin	3
chordin	4	NF-kappa-B	3

CNTF	1	nodal	2
dachshund	2	NOTCH	4
deformed	1	Numb	1
delta	2	odd-skipped	1
dickkopf	2	Orthodenticle	2
dishevelled	2	paired	1
distal-less	2	par	4
E-cadherin	1	PAX	4
EGF	1	plexin	7
ephrin	7	polycomb	8
Even-skipped	1	pumilio	2
F-box	3	Ras	13
FGF	10	Rhomboid	4
follistatin	3	robo	4
FOX	17	runt	4
frizzled	8	semaphorin	9
GATA	7	sex comb	6
GDF	2	SMAD	10

geminin	1	snail	1
GFAP	1	SOX	10
giant	1	STAT	1
hairy	6	T-box	5
hedgehog	2	TCF3	1
helix-loop-helix	9	TGF	8
HGF	1	Trk	1
hmg	20	twist	3
HOX	38	VEGF	1
I-kappa-B	3	vimentin	1
insulin	6	WNT	6
integrin	15	WT1	1
JAK	3	XIST	1
Kruppel	14	zinc finger	324

Chapter 4. Theoretical analysis indicates ‘the principle of fluctuations’ fundamentally control life phenomena.

Abstract

The proposition of the existence of fundamental systems which control or manage life phenomena has not given the solution. The profiles of gene expression or the pathways for the protein interactions have been elucidated. However, those are the results of the gene expression patterns and the pathways only under steady states and have not been elucidated the fundamental systems or principles of complex life phenomena. Hence, do really the systems or principles exist which fundamentally control or manage the complex life phenomena? I logically proved that ‘the principle of fluctuations’ control or manage the fundamental life phenomena. In other words, life phenomena exist on the basis of ‘the principle of fluctuations’. Hence, living bodies can cope with the change of diverse conditions. Replication of DNA, DNA mismatch repair, gene expression, translation into amino acids, production of proteins, the process of energy productions and the process of signal transductions are not be firmly operated in 100%. Notwithstanding, living bodies operate life

phenomena without hindrance. This means the existence of ‘fluctuations’ fundamentally. Life phenomena are operated harmoniously. Since living bodies are constructed by molecules, living bodies must be accepted the uncertainty principle in the field of physics. It is impossible to make mathematical formulas, because life phenomena are too complex and too flexible to make such formulas. Living bodies are not machines. Therefore, I suppose that life is the states of operation of life phenomena on the basis of ‘fluctuations’, because the boundary line between living conditions and dead conditions is not be able to be defined.

Introduction

Since physicist Dr.Schrödinger published the book ‘What is life?’ in 1944, the proposition “What is life?” has been one of the most important propositions in the field of life science[1]. But still now, the solution of the proposition has not been elucidated. The fields of systems biology and bioinformatics emerged to solve to the proposition which the systems control or manage life phenomena. However, these fields have not been given the solution to the existence of the fundamental systems which control or manage life phenomena so far. I think that the way of trials to

elucidate life phenomena in terms of systems biology or bioinformatics are correct. However, even if gene expression profiles by microarrays and protein interaction pathways were elucidated, or analysis of biological information were performed, those trials have not been given the solution to the proposition of the existence of the systems which control or manage fundamental life phenomena. Living bodies maintain homeostasis under the steady state, but if once those conditions are damaged by some kinds of stresses, the homeostasis brake and other life phenomena set in motion[2, 3]. Do really replication of DNA, gene expression, translation into amino acids, protein production, pathways of energy production of glycolysis or TCA cycle and pathways of signal transductions which are essential for life, support life phenomena all together harmoniously? The solution to this proposition is NO! Many systems and pathways have been elucidated, but even one of them has not been the fundamental systems which control or manage life phenomena. In the fields of systems biology and bioinformatics, the concept of robustness advocated and those scientists emphasize that systems cope with diverse life phenomena by the existence of robustness[4, 5]. And in the fields of chemical biology, biophysics, physical

biology, those scientists emphasized the existence of the system on the assumption of mechanism or physicalism[6-20]. Trials to elucidate life phenomena have also been performed by complex system and self-organization, these trials have not been successful so far [21-25]. Are life phenomena systematic such as machines which can be designed by mathematical formulas? When living bodies fall into danger and certain systems do not operate fully, living bodies operate the other systems to compensate for the danger to survive. It is redundancy. Hence, how does make an interpretation of the existence of redundancy? How is the uncertain principle of physics adapted for life phenomena[26, 27]? It is very significant to elucidate life phenomena. Hence, the propositions which theoretical biologists try to elucidate, are 'what is the fundamental principle to control and manage life phenomena?' and "What is life?" I emphasize that the system does not exist to control or manage life phenomena fundamentally, but 'fluctuations' exist on the basis of life phenomena. The uncertainty principle of physics is the basis of the existence of 'fluctuations'. Because living bodies are constructed by molecules, life phenomena are operated by the uncertainty principle of physics. Finally,

the solution of the proposition “What is life?” is supposed to be the condition which life phenomena are controlled or managed by ‘fluctuations’.

The definition of words and phrases.

Before the discussion, the definition of words or phrases is significantly important. Because scientists must use appropriate words or phrases. In the fields of systems biology and bioinformatics, the word “robustness” is used to express flexible strength of the systems of life phenomena. But the real meaning of robustness is to withstand or overcome adverse conditions by dictionaries[28-38]. If life phenomena are not based on the systems, flexibilities, randomness and vagueness, ‘fluctuations’ is thought to be appropriate to express the principle of control the system of life phenomena.

Systematic or nonsystematic?

Are life phenomena systematic or nonsystematic? There must be only two choices. In the fields of systems biology and bioinformatics, scientists emphasize that life phenomena are the aggregation of individual systems, and life phenomena are smoothly controlled or managed by robustness on

the basis of the aggregation of those systems. Forthermore, some scientists in those fields try to make mathematical formulas on the basis of mechanism or physicalism. On the contrary, I emphasize that systems are controlled or managed by ‘the principle of fluctuations’ which are constructed which is existed on the basis of unstable life phenomena. Because of the existence of ‘the principle of fluctuations’, the values of blood examinations from one healthy human have the unevenness (data not shown). However, human beings can act life phenomena harmoniously. This means that life phenomena are controlled or managed fundamentally on the basis of ‘fluctuations’. In other words, life phenomena are controlled or managed on the basis of ‘the principle of fluctuations’ fundamentally. Endosymbioses have dramatically altered eukaryotic life, but were thought to have negligibly affected prokaryotic evolution. By analyzing the flows of protein families, the evidence that the double-membrane, Gram-negative prokaryotes were formed as the result of a symbiosis between an ancient actinobacterium and an ancient clostridium. The resulting taxon had been extraordinarily successful, and had profoundly altered the evolution of life by providing endosymbionts necessary for the emergence of eukaryotes

and by generating Earth's oxygen atmosphere. Their double-membrane architecture and the observed genome flows into them suggest a common evolutionary mechanism for their origin: an endosymbiosis between a clostridium and actinobacterium[39]. Why sex evolved and persists is a problem for evolutionary biology, because sex disrupts favorable gene combinations and requires an expenditure of time and energy. Further, in organisms with unequal-sized gametes, the female transmits her genes at only half the rate of an asexual equivalent. Many modern theories that provide an explanation for the advantage of sex incorporate an idea originally proposed by Weismann more than 100 years ago: sex allows natural selection to proceed more effectively because it increases genetic variation. Dr. Goggard and colleagues tested this hypothesis, which still lacked robust empirical support, with the use of experiments on yeast populations. Capitalizing on recent advances in the molecular biology of recombination in yeast, they produced by genetic manipulation strains that differed only in their capacity for sexual reproduction. They show that, as predicted by the theory, sex increases the rate of adaptation to a new harsh environment but has no measurable effect on fitness in a new benign

environment where there is little selection[40]. If the systems are robust in human cells, tissues and organs, life phenomena may not cope with flexibly the dangerous conditions which menace the homeostasis. Further, if systems exist on the basis of life phenomena, living bodies could not acquire these flexibilities, in other word, ‘the principle of fluctuations’. Hence, any living bodies such as bacteria, yeast, human beings may be disturbed evolution. In that case, the systems must not have the space to acquire other systems, because the systems must be constructed completely. I deductively and logically proved that life phenomena do not exist on the basis of systems. It is very difficult to prove logically that life phenomena are not fundamentally controlled or managed by the systems. Even if only 5000 molecules control or manage all biological activities in a certain living body, the systems maintain homeostasis. If the systems are damaged, the living body copes with redundancy. But if the systems were not able to maintain homeostasis, the living body will die. Is it possible to predict which and how the pathways or the systems cope with those crises? It depends on the size and type of crises. Therefore, as a result, it is impossible to predict how to cope with those crises. Because, life

phenomena are controlled or managed by the principle of uncertainty in the field of physics. In other words, I deductively and logically proved life phenomena are unstably fluctuated under those crises. If ‘fluctuations’ do not exist under the crises for life, living bodies may be accepted the crises and stop biological activities. Notwithstanding, living bodies manage to survive. This is for the sake of existence of ‘fluctuations’. But it is impossible to predict how to manage to survive. I proved by abduction as stated an above-mentioned. If the systems exist fundamentally control or manage life phenomena, life phenomena may be controlled or managed by the gene products of house-keeping genes. However, these genes products must be classified into the several essential pathways such as DNA replication, DNA mismatch repair, gene expression, translation into amino acids, production of proteins, the process of energy productions, and the process of signal transductions and so on. And the existence of upper systems or the link to totally control or manage to these pathways is not identified still now. In addition, it is undeniable to predict how much amount or genes and proteins must be different from individual living bodies, and how to respond.

Mechanism, physicalism, probability theory and the uncertainty principle

Systems biology, bioinformatics, chemical biology, biophysics and physiological biology ultimately exist on the basis of mechanism or physicalism. But because of the uncertainty principle in the field of physics, life phenomena are not able to be predictable. In case of DNA replication or DNA mismatch repair, there exist mistakes in certain probabilities. And the timing of gene expression and gene expression pattern are also considered by probability theory. The timing of working, permutation, combination and the efficiency of working of proteins are also considered by probability theory. Hence, how much amount of proteins is secreted? How fast are the proteins degraded? Do pathways of energy productions usually produce the same amount of energy? How are those pathways exact and fast under stress? How fast does the concentration in blood of antibiotics increase, in case of giving antibiotics? It is impossible to solute these propositions. Because life phenomena are exceedingly complex and unpredictable. It is further more impossible to design mathematical formulas. Because all of life phenomena must be considered by probability

theory. There manifestly exist the differences of biological activities among individual living bodies from the results of research and treatment. This is the way Heisenberg stated the uncertainty principle originally: If the measurements on any objects are made, and the *x*-component of its momentum with an uncertainty Δp can be determined at the same time, it is impossible to know its *x*-position more accurately than $\Delta x = h/\Delta p$, where *h* is a definite fixed number given by nature. It is called “Plank’s constant”. Hence, it means that life phenomena have the uncertainty and are not predictable even in an instant future. This means that the positions and momentums of molecules are not predictable.

Living bodies are not machines.

The academic discipline which I advocate theoretical molecular biology, is a science to elucidate life phenomena logically and theoretically. Life phenomena must be considered by probability theory, and exist on the basis of ‘the principle of fluctuations’. According to ‘fluctuations’, life phenomena which are not machinelike, flexibly cope with the changes of environments and crises of homeostasis. I inductively proved as a stated above. I will elucidate the proposition which living bodies are machines.

Firstly, if living bodies were machines, living bodies could not accomplish evolution. Furthermore, DNA replication, DNA mismatch repair, gene expression, translation into amino acids and productions of proteins might have mistakes. If living bodies were machines, living bodies must not accomplish evolution and not make mistakes in case of DNA replication, DNA mismatch repair, gene expression, translation into amino acids, productions of proteins and so on, because living bodies must be created in 100% machinelike. Hence, the solution is that living bodies are not machines. Firstly, if there do not exist ‘fluctuations’, individual cells cope with crises of homeostasis in 100% uniformly. And tissues or organs which are the aggregate of cells also cope with in 100% uniformly. But living bodies operate biological activities harmoniously without hindrance and cope with crises of homeostasis. To sum up, life is on the basis of ‘fluctuations’. Blood examinations were performed from only one male (Data not shown). The results show that there were certain different measured values of two blood samples which were took at interval of only one hour. Even at the same time, the measured values of two blood samples have difference. These were measurement errors. However, instead of the

existence of unevenness of blood examinations, human beings can perform life phenomena without any obstructions. This means that life phenomena fundamentally have unevenness. Hence, life phenomena are based on the uncertainty principle in the field of physics, and the measured values of blood examinations must not be able to predictable only in one hour. Because, the systems which control or manage life phenomena are based on ‘fluctuations’. I proved the existence of ‘fluctuations’ inductively. Secondly, I will prove life phenomena are controlled or managed by the uncertainty principle. Can we predict our life phenomena or body conditions in one hour, one week or one year? This is impossible. We will be able to interpret the events such as life phenomena or body conditions by the analysis of gene expression profiles or the pathways of protein interactions. Hence, life phenomena must not be predictable according to the uncertainty principle in the field of physics. This means that there exist uncertainties of life phenomena on the basis of ‘fluctuations’. I proved an above-mentioned deductively. Thirdly, why living bodies can perform evolution? If the systems which control or manage are robust, evolution might not be performed. Hence, the systems which control or manage life

phenomena must have flexibilities to acquire new characters or traits. This means that there do not exist the robust systems, but must namely exist flexible ‘fluctuations’. I proved an above-mentioned deductively.

It is impossible to make mathematical formulas.

It is also impossible to make mathematical formulas. That is not why analytical capabilities of the present time computers are not sufficient to analyze more than billions of interactions of molecules in living bodies. If it will be possible to analyze more than billions of interactions of molecules in living bodies, will it be possible to make mathematical formulas in the future? And if all systems of life phenomena were elucidated in the future, will it be declared to elucidate life phenomena completely? The solution of these prepositions is NO! It is impossible to make mathematical formulas to elucidate the systems or the principles of life phenomena, because life phenomena are too complex, and biology is different from mathematics or physics. And many systems operate together and are connected with other systems on the same time in life phenomena. Hence, it is impossible to make mathematical formulas and elucidate the systems or the principles of life phenomena fundamentally. I inductively proved that the systems do not

exist on the basis of life phenomena. Some theoretical biologists try to make mathematical formulas, but living bodies do not live and cope with crises of homeostasis in 100% uniformly. That is why that it is impossible to make mathematical formulas.

What is life?

Can the boundary line between living conditions and dead conditions be defined in terms of biological and philosophical point of views? Is it possible to define when living bodies die? The solutions for these propositions may be that living conditions and dead conditions are continuous sequentially. Because living bodies are not machines, it will not be impossible to define the boundary line between living conditions and dead conditions. The important fact is that life phenomena are not predictable, not be able to make mathematical formulas to elucidate the systems, and exist on the basis of ‘fluctuations’. That is why living bodies are able to operate diverse biological activities and cope with crises of homeostasis harmoniously.

Conclusion

Here, I logically and theoretically proved that the solutions for the propositions of the systems or principles which control and manage life phenomena fundamentally are ‘the principle of fluctuations’. I name this thought as Itoh’s ‘the principle of fluctuations’. And the proposition of “What is life?” may be supposed to operate or perform biological activities on the basis of ‘fluctuations’.

Methods summary

Blood examinations were performed from only one male at interval of only one hour. And blood examinations were performed from the same person on the same time as a negative control.

References

1. Schrödinger, E. *What Is Life?: with "Mind and Matter" and "Autobiographical Sketches"* (Cambridge University Press, 1992).
2. Alberts, B. *et al. Molecular Biology of the Cell: Reference Edition 5th edition* (Garland Science, 2007)
3. Fauci, S. A. *et al. Harrison's Principles of Internal Medicine, 17th Edition* (McGraw-Hill, 2008)
4. Kitano. H, Systems biology: a brief overview. *Science* **295**, 1662-4 (2002).
5. Kitano. H, Computational systems biology. *Nature* **420,** 206-10 (2002).
6. Creighton. E. T, *The Physical and Chemical Basis of Molecular Biology* (Helvetian Press, 2010).
7. Meyer B. Jackson. B, M, *Molecular and Cellular Biophysics* (Cambridge University Press, 2006).
8. Beard. A, B, & Qian. H, *Chemical Biophysics: Quantitative Analysis of Cellular Systems* (Cambridge University Press, 2010).
9. Sener. B, *Innovations in Chemical Biology* (Springer Netherlands, 2010).

10. Waldmann. H, & Petra. J, *Chemical Biology* (Wiley-VCH, 2009).
11. Mayer. G, *The Chemical Biology of Nucleic Acids* (Wiley, 2010).
12. Miller. D, A, & Tanner. J, *Essentials Of Chemical Biology: Structure and Dynamics of Biological Macromolecules* (Wiley, 2008).
13. Al-Rubeai. M, & Fussenegger. M, *Systems Biology* (Springer Netherlands, 2010).
14. Nagasaki. M, Saito, A, Doi. A, Matsuno. H, & Miyano. S, *Foundations of Systems Biology: Using Cell Illustrator and Pathway Databases* (Springer-Verlag New York Inc, 2009).
15. Iglesias. A, P, & Ingalls. P, B, *Control Theory and Systems Biology* (The MIT Press, 2009).
16. Choi. S, *Introduction to Systems Biology* (Springer, 2010).
17. Glaser. R, *Biophysics: An Introduction* (Springer, 2010).
18. Claycomb. J, & Tran. J, *Introductory Biophysics: Perspectives on the Living State* (Jones & Bartlett Publishers, 2010).
19. Nelson. P, *Biological Physics* (W. H. Freeman, 2007).
20. Scherer. O. J, P, & Fischer. F, S, *Theoretical Molecular Biophysics* (Springer, 2010).

21. John H. Miller. H, J, & Scott E. Page. E, S, *Complex Adaptive Systems* (Princeton University Press, 2007).

22. Page. E, S, *Diversity and Complexity* (Princeton University Press, 2010).

23. Kauffman. S, *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity* (Oxford University Press, 1996).

24. Kauffman. S, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, 1993).

25. Camazine. S, *et al. Self-Organization in Biological Systems* (Princeton University Press, 2003).

26. Feynman. R, Leighton. R, B, & Sands. M, L, *The Feynman Lectures on Physics* (Addison-Wesley Publishing Co, Inc., 1963).

27. Dirac. M, A, P, The principles of quantum mechanics 4th edition (Oxford University Press, USA, 1982).

28. Trumble. R, W, *Shorter Oxford English Dictionary 6th edition* (Oxford University Press, USA, 2007).

29. Hornby. S, A, Turnbull. J, Lea. D, & Parkinson. D, *Oxford Advanced Learner's Dictionary 8th edition* (Oxford University Press, USA; (April

26, 2010).

30. Title. B, *Cambridge Advanced Learner's Dictionary 3th edition* (Cambridge University Press, 2008).

31. *Collins English Dictionary 30th anniversary edition of 10th revised edition* (Collins, 2010).

32. House. R, *Random House Webster's Unabridged Dictionary* (Random House Reference, 2006).

33. Soanes. C, & Stevenson. A, *Concise Oxford English Dictionary* (Oxford University Press, 2008).

34. *Cambridge Academic Content Dictionary* (Cambridge University Press, 2008).

35. *Oxford Dictionary of English 3rd Revised edition* (Oxford University Press, 2010).

36. *Webster's 3rd New International Dictionary* (Merriam Webster, 2000).

37. *Longman Dictionary of Contemporary English* (Pearson Longman, 2009).

38. Anderson. S, Katharine Coates. K, & Hands. P, *Collins Cobuild Advanced Dictionary* (Heinle, 2008).

39. Lake. A, J, Evidence for an early prokaryotic endosymbiosis. *Nature* **460**, 967-971 (2009).

40. Goddard. M, R, Godfray. H, C, & Burt. A, Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**, 636-40 (2005).

Chapter 5. The future images of medicine.

1) Gene therapy

Recently, we do not hear about the phrase ‘gene therapy’. Because it is impossible to do gene therapy right now. Original definition of gene therapy is that to correct incorrect genome sequences into correct sequences. Before profound discussion about the possibility of accomplishment of the technique of genome alteration, development of the technique of gene transfer began. Hence, viral vectors were developed because of high efficiency of gene transfer into cells. Adenoviral vectors transfer genes the most efficiently. However, adenoviral vectors are not inserted in to genomes and were not used gradually. Using retroviral vectors, DNA fragments are able to be inserted in genomes. However, because of carcinogenesis, retroviral vectors are not able to be used in gene therapy. The helpes vectors are used to transfer DNA fragment into brain cells. Howsoever, it is impossible to do gene therapy because the techniques to correct incorrect genome sequences into correct sequences were not accomplished. Human cells have many barriers to protect

mutations. That is why it is impossible to correct incorrect genome sequences into correct sequences. If it is easily DNA fragments inserted in human genomes, it means that carcinogenesis occur very easily. And human genome in mucosal cells in intestines must be inserted DNA fragment easily. Hence, the concept of gene therapy is going to disappear. There are a lot of unknown things to get solutions without doing experiments. But before doing experiments, life scientists and medical scientists must meditate and think profoundly if propositions really have solutions. Compared with human somatic cells, human germ cells have the ability to DNA recombination. However, is that techniques really efficient to cure diseases? In case of human somatic cells, it is almost impossible to correct incorrect genome sequences into correct sequences. If it becomes possible to correct incorrect genome sequences into correct sequences, germ cells must be used. However, there are big obstacles ethically and technically. If it were permitted ethically, do really correct only incorrect sequences in 100%? If mutations occur other portions of human genomes, what will happen? It is impossible to predict what happen if these problems were not solved. In my opinion, if germ cells were not used, there will not

be any chance to accomplish gene therapy in the future. However, if germ cells were used, it is almost impossible to find out new techniques to correct incorrect genome sequences into correct sequences. Therefore, doing gene therapy is impossible right now.

2) Genomic diagnostics and gene diagnostics

Medical treatment in 5-10 years from now will be totally different from present medical treatment. Because China government officially announces that they will sequence of a whole human genome of one person by about $100 in 2-3 years. Actually, Stanford University sequences a whole human genome of one person by about $200. Hence, we will be able to know our own genome sequences in about $100 in 2-3 years. If human genomes among 10 million people were compared, analyzed those sequences statistically and found the positions of insertion, deletions, mutations and small nucleotide polymorphisms (SNPs), it will be possible to predict when patients become diseases, which types of diseases patients suffer from and which types of drug combination are efficient for patients. In other words, in 5-10 years,

hospital will become the place to go to check if patients do not suffer from predicted disease yet, but will not become the place to go after suffering from diseases. There actually exist the pedigrees of cancers. If we predict about when they suffer from cancers, preventive medicine will be important. And it will be much easier to find cancers in early stage and to cure cancers early. However, it is unsafe to expect too much expectation. Because the causal genomic insertions, deletions, mutations and SNPs or causal genes for hypertension, hyperlipidemia, diabetes mellitus, autoimmune disorder, connective tissue diseases and other diseases will be found, but those findings will not be able directory to contribute to treatment of those diseases. However, if patients know when and which types of diseases they will suffer from, it will be possible to delay the timing of suffering from those diseases. Therefore, as the results, it will be possible to find the new treatment to cure those diseases.

3) Drug delivery system (DDS)

The concept of drug delivery system (DDS) emerges instead of gene therapy. DDS is the method to deliver chemical materials to target cells or tissues. The best examples are the molecular target drugs. In other words, DDS means cell-targeting or tissue-targeting. These methods are already used clinically and the efficiency is quite good. That is why DDS will be developed more and more. However, it is impossible to accomplish DDS without using antibodies right now. Will it be possible to find chemical materials instead of antibodies in the future? I think there are possibilities to find such chemical materials. However, it is very difficult to find such chemical materials because antibodies bind very efficiently the target molecules and affinity of between antibodies and target molecules are quite high. Therefore, antibodies are the best DDS tools right now.

4) Molecular target drugs

In case of cancer therapy, molecular target drugs are attracted a great deal of attention. This is a matter of course. I have been put forward to this method is highly efficient to cure cancers. In case of treating cancers, there are only two approaches which utilize the difference of inside of cancer cells or outside of cancer cells. Many oncologists have been doing researches about epigenetics such as signal transduction by phosphorylations or methylation of genomes and so on, mainly inside of cancer cells. However, even if we know epigenetics in detail inside of cancer cells, it is useless to be utilized efficient DDS. In my opinion, if the proposition is to cure cancers, the solution is to utilize the difference of outside of cancer cells. In case of treatment of cancers, it is not significant what happen inside of cancer cells. If the aim is to cure cancers, DDS is the most important tool. Even if what happens inside of cancer cells is a black box, it does not matter to cure cancers. When it comes to cure cancers by DDS, antibodies are thought to be the best tool. The target of those antibodies must be the portion of membrane proteins which are outside of cancer cells. In human genomes, there do not exist

cancer specific genes. Therefore, certain amount of normal cells must be destroyed. However, for advanced and end stages of cancer patients, such kind of side effect will be permitted, because the aim is to save patient's life. Even if we know all epigenetics inside of cancer cells, will it contribute to cure cancers? I do not think that doing researches about epigenetics inside cancer cells are not significant. However, in case of treatment of cancers, doing researches about inside of cancer cells are not significant to cure cancers. Compared with cancer cells and normal cells, we will be able to find highly expressed membrane genes in cancer cells using results of microarray in NCBI or other official databases. In this case, the products of those highly expressed membrane genes are suitable targets of molecular target therapy using antibodies. It is possible to produce hybridomas against the portion of outside of those membrane proteins in cancers. It is possible to produce single chain antibodies (ScFvs) from hybridomas which express those portions. I determined the amino acid sequences in variable region of ScFvs against Prostate Specific Membrane Antigen (PSMA). ScFv is monovalent. However, making divalent human type antibodies, these

antibodies must be used as molecular target drugs. The important points are to select the appropriate portions of membrane proteins in cancer cells. Furthermore, several types of hybridomas must be produced. In addition, it is possible to predict three dimensional structures of proteins. Therefore, preparing several types or peptide antibodies on the same time is safe. I experienced that more than 30 types of cDNA were emerged to determine cDNA sequences from 4 types of hybridomas by using RT-PCR. I translated those cDNA sequences into amino acid sequences, but still there were more than 20 types of amino acid sequences in ScFvs. I made my mind to select 5 amino acid sequences and tried to make ScFvs. It was very difficult to prepare 5 types of ScFvs on the same time, because ScFvs making from bacteria were very fragile. I spent one and half year to produce 5 ScFvs on the same time because I needed to determine the best condition to produce ScFvs. To my surprise, 5 types of ScFvs had the same affinities against target proteins. In my opinion, if there exist several types of amino acid sequences in variable regions of ScFvs, all ScFvs will bind efficiently the target portions of proteins. I experienced any difficult steps to

produce ScFvs and I have know-how to produce ScFvs. I try to collaborate with many pharmaceutical companies, but I have not been gotten good responses from them. However, it is easily possible to select suitable membrane proteins as targets of molecular therapy drugs, surveying NCBI and other official databases. The molecular target drugs using my thought will be efficient to early stage to end stage cancer patients. Therefore, for those patients, molecular target drugs using this method will be the good news. In addition, I think that these molecular therapy drugs will be also efficient to auto-immune diseases.

5) Regenerative medicine

The most significant problem in regenerative medicine is carcinogenesis. If regenerative medicine were enforced, results of carcinogenesis will be known in 20-30 years from now. If patients want to accept regenerative medicine to improve their quality of life (QOL), it will be difficult to stop them regardless of carcinogenesis. In my opinion, if regenerative medicine were enforced, I do not think that stem cells go to only target tissues. As a result of this condition, a whole body will be

suffered from carcinogenesis. However, if patients want to accept regenerative medicine to improve their QOLs regardless of carcinogenesis and other clinical risks, regenerative medicine will be performed with thorough informed consent. In my opinion, the most suitable cells for regenerative medicine are germ cells. Because germ cells keep intact genome sequences. However, it will be difficult to produce a whole tissue. Even if intact cells were used, mutations are unavoidable in the step of DNA duplication. Therefore, it is logically difficult to produce a whole tissue without DNA mutations. That is why it is not realistic to produce a whole tissue. However, even if it were impossible to produce a complete whole tissue, it is useful to improve function of tissue. It will be possible to improve cardiac function to transplant stem cells into a heart which were suffered from myocardial infarction. It will be possible to improve brain function to transplant stem cells into brain which were suffered from cerebral infarction or cerebral palsy. Hence, according to diseases, treatment must be changed. For cancers, DDS will be mainly used and for dysfunction of tissues, regenerative medicine will be mainly used. In my opinion, in case of regenerative medicine, the goal is to improve

dysfunction of tissues. I have a great interest to improve brain function in Down syndrome patients. However, it is unclear if regenerative medicine were really useful to improve dysfunction of brain in those patients.

Chapter 6. What is life?

(Proposition 1) What is life?

(Proposition 2) Is it possible to make mathematical formulas to describe life phenomena?

Do there exist solutions for these two propositions? Systems biology, nonlinear science, phase transition, the quantum theory, complex systems, self-assembly, chaos, bioinformatics, biophysics, biological physics, chemical biology. Among these academic fields, the point in common is to try to make mathematical formulas to describe life phenomena. Making mathematical formulas is the thought based on mechanism or physicalism. Let us assume that life phenomena were able to be described by mathematical formulas. I explain some counterexamples. 1) In that case, life phenomena in the future must be able to be predicted. However, it is impossible to predict life phenomena in the future. It is unknown when human beings become ill and what human beings think in the future. 2) If life phenomena were able to be described by mathematical formulas, life phenomena must be mechanical. Do machines repair their troubles by themselves? If machines break down, machines continue to be out of order.

However, if human beings catch cold, it is cured in few days by themselves. This is a proof that human beings are not machines. 3) If life phenomena were mechanical, evolution must not happen. However, human beings have been evolved. Hence, human beings are not machines. These counterexamples are enough to prove that human beings are not machines and it is impossible to describe life phenomena by mathematical formulas. Even if it were possible to make mathematical formulas, it is limited in only a small portion of life phenomena. It is impossible to describe the whole life phenomena by mathematical formulas. Because life is not a machine. Let’s go back proposition 1. What is life? I list up what are not alive. Machines, computers, books, air, water. These are not alive. On the contrary, I list up what are alive. Bacteria, yeasts, plants, insects, animals. These are alive. What are common factors? 1) These are made of cells. 2) These have the ability to survive and leave offspring. Hence, the solution of proposition 1 is things which are made of cells and have the ability to survive and leave offspring. The uninterrupted process to survive and leave offspring and die. This is life!