Slide 1

I would like to start my lecture on Chapter 1 Genomic Abnormalities and Cancer in this course on the Basics of Genomic Medicine and Cancer Genomes.

Slide 2

In this chapter, I will give teach you the basics of “What is cancer?” and “What is a genome” and I will explain about Genomic Abnormalities and Cancer.

Slide 3

To start off with, “What is cancer?”

Slide 4

In Japan today, the average lifespan for men and women is over 80 years. One out of every 2 people is diagnosed with “cancer” and 1 out of every 3 people dies from “cancer.” Therefore, “cancer” is a national affliction affecting modern Japan.

Slide 5

So why is the number of cancer patients continuing to increase? This graph is from a study on the incidence of colon cancer per 100,000 women in the UK, and it shows the incidence of colon cancer by age group. As you can see, the incidence of colon cancer rises sharply over the age of 60 と the incidence of colon cancer. The accumulation of genetic abnormalities or mutations with age is passed on at the cellular level. As a result, the risk of developing cancer increases.

Slide 6

Cancer is a malignancy that haphazardly grows, invades, and metastasizes inside the body, leading to the death of the individual. Cancer invades and metastasizes to surrounding tissue. This is the predominant characteristic of cancer.

Slide 7

Why does cancer grow chaotically and haphazardly? In normal cells, signaling relays are induced pursuant to an order to proliferate issued by other cells, and cells proliferate. Expression of a gene and its translation into a protein is strictly controlled by the number of receptors on the surface of a cell receiving an order to proliferate. The number or size of cells in tissue is strictly managed via this control.

Slide 8

Cancer cells, however, issue their own order to proliferate and then proliferate when that order is received, or the number of the number of receptors receiving an order to proliferate increases, or cells proliferate as a result of the production of a mutated receptor that is constantly activated independent of any order to proliferate. In addition, a genetic mutation can cause a signaling molecule downstream of a receptor to remain activated independent of any order to proliferate. This is because abnormalities in signaling pathways occur in cancer cells, and those cells proliferate continuously and faster than normal cells. Thus, cancer cells proliferate autonomously, chaotically, and excessively.

Slide 9

Unlike a benign tumor, the predominant characteristic of cancer is that it invades surrounding tissue and often metastasizes to other organs. If invasion by or metastasis of cancer could be inhibited, then outcomes of cancer treatment should improve.

Slide 10

In addition, cancer cells have various characteristics unlike those of normal cells. These include the acquisition of autonomous growth signaling, avoidance of growth-inhibiting signaling, avoidance of apoptosis (programmed cell death), acquisition of the capacity for unrestricted growth, induction of angiogenesis, acquisition of the ability to invade cells and metastasize, induction of genomic instability, mutations, and tumor-promoting inflammation, avoidance of destruction by the immune system, and reconfiguration of energy metabolism.

Slide 11

You've recently heard the term "genome," but what does it mean?

Slide 12

A genome is a combination of the German word for gene, which is also gene, and the Greek suffix -ome, meaning body, to refer to a body (or set) of genes. A genome is all of the genetic information for an organism included in its chromosomal DNA.

Slide 13

Each of the 37 trillion cells that make up our bodies contains the human genome, which is made up of 3.2 billion letters (base pairs) of DNA.

Slide 14

And what exactly is deoxyribonucleic acid, or DNA?

Slide 15

DNA is genetic information found in a cell nucleus. DNA comprises all of the genetic information that makes up humans and other organisms. It has a double helix structure as a result of complementary base pairing. The double helix structure of DNA is known to facilitate the accurate replication of DNA. The base adenine forms a hydrogen bond with the base thymine, and the base cytosine forms a hydrogen bond with the base guanine. These complementary base pairs create a double helix structure.

Slide 16

The nucleus of human cells contains DNA that is about 2 meters in length. DNA is tightly wound around histone proteins and tightly folded into chromosomes.

Slide 17

Typically, human diploid cells have 22 pairs of homologous chromosomes and 1 pair of sex chromosomes, for 46 chromosomes in total. Genomic information is found in those chromosomes.

Slide 18

DNA is information to produce proteins that serve various function and facilitate biological activity, and that information is written in genes. There are around 24,000 genes in the human genome.

Slide 19

When genetic information from DNA in the cell nucleus is transcribed into messenger RNA, the double helix structure of DNA unwinds. Genetic information from DNA is transcribed into messenger RNA by RNA synthase, using a single strand of DNA in the nucleus of cells as a template. The letters (base pairs) used to do this are not A, G, C, and T but A, G, C, and U; U stands for the base uracil.

Slide 20

Messenger RNA that is transcribed in the cell nucleus migrates to the cytoplasm via nuclear pores, and it is translated into proteins, which are one or more chains of amino acids, by ribosomes.

Slide 21

In ribosomes, each 3-base sequence of messenger RNA, known as a codon, is translated into an amino acid.

Slide 22

The genetic information written in DNA is transcribed into messenger RNA. Information from messenger is translated into proteins, which are one or more chains of amino acids. This is known as the central dogma of gene expression, and it a basic rule for organisms, though there are exceptions like retroviruses.

Slide 23

Results of analyzing the human genome were reported in 2004. There are 22,000 human genes, which is less than the anticipated 100,000. Humans have about the same number of genes as sea urchins but fewer genes than rice. Accordingly, the expectation that humans surely had significantly more genes than other organisms was clearly wrong. Over 90% of the human genome consists of non-coding DNA, and its biochemical activity is actively being analyzed. Mapping the human genome took 13 years and cost a massive 3 billion dollars. Now, however, it can be analyzed in about 1 week at a cost of 500,000 yen.

Slide 24

What made this possible was scientific and technological advancement brought about by a dramatic improvement in the speed of DNA analysis. Technology for DNA analysis is advancing every day.

Slide 25

Next, I will explain the association between Genomic Abnormalities and Cancer.

Slide 26

Biological activity depends on cell activity. Cell activity depends on protein functioning. Protein functioning depends on amino acid sequences. And amino acid sequences depend on base sequences in genetic DNA. DNA is a blueprint. Thus, if the code of DNA changes, then cell activity can also change.

Slide 27

DNA sequences, gene locations, and the number of chromosomes are stringently controlled to maintain the stability of genomic information. Over 100 genes help to maintain genomic stability by controlling mechanisms of DNA repair, the mechanism of chromosome segregation, and the functioning of cell cycle checkpoints.

Slide 28

So, does one’s genome never change during one’s life?

Slide 29

A large amount of DNA located in chromosomes is accurately replicated when cells follow precise steps and processes known as the cell cycle. A cell divides into 2 daughter cells that are genetically the same. The cell cycle always proceeds in 4 stages — G1, S, G2, and M, and there are cell cycle checkpoints after each stage.

Slide 30

However, gene mutations can occur as a result of an error in DNA replication during cell division, damage to DNA or an error in replication as a result of exposure to radioactivity or a chemical substance, and chromosomal damage. Genetic mutations constantly occur in various cells of the body, but normally those mutations are corrected before they are passed on to future generations of cells. However, the cellular mechanisms for repairing damaged DNA can be disrupted and the capacity for repair diminishes with age. As a result, mutations accumulate.

Slide 31

A mutation in DNA can cause a base to change, thus causing an amino acid to change, translation to stop, or causing an abnormality in an amino acid sequence. If, for example, cytosine changes to thymine and the codon still codes for glycine, then the amino acid sequence will not be affected, resulting in what is known as a silent mutation. If, however, guanine changes to adenine, then the encoded amino acid will change from glycine to serine. Such a mutation is referred to as a missense mutation. In addition, adenine can change to thymine, creating a stop codon. Such a mutation is known as a nonsense mutation. Deletion or insertion of a base can cause a missense mutation or a nonsense mutation.

Slide 32

Thus, a mutation in DNA can cause a base to change, thus producing an abnormality in an amino acid sequence. In such an event, a change in an amino acid sequence will cause the structure or function of a protein to change, which is certain to affect cell activity and the individual.

Slide 33

So how are gene mutations linked to the development of cancer?
Does this apply to all 24,000 genes?

Slide 34

We now know that abnormal cell proliferation is induced by the activation of cancer genes known as

oncogenes that serve to accelerate cell growth and by the inactivation of tumor suppressor genes that slow down cell growth.

Slide 35

Moreover, we know that tumor suppressor genes responsible for apoptosis, mechanisms of DNA repair, the mechanism of chromosome segregation, and the functioning of cell cycle checkpoints are inactivated, inducing genomic instability. This is closely related to the malignant transformation and progression of cancer.

Slide 36

The instability of a cancer genome as a result of the accumulation of gene abnormalities leads to chromosomal instability.

Normal cells, like those in the photo on the left, have 46 chromosomes. However, colon cancer cells, like those in the photo on the right, have 70 chromosomes and various abnormalities in parts of those chromosomes, such as deletions and translocations. We now know that genomic abnormalities cause the development and malignant transformation of cancer.

True or false?

・Genetic information in every cell is processed in the following order: DNA→RNA→protein. This basic principle is known as the central dogma. (Answer: True)

・Genes consist of a regulatory region and coding region. Mutations in the former after the gene's products and mutations in the latter alter the gene's expression. (Answer: False)

・Cancer is a disease of the genome at the cellular level. (Answer: True)

・Patients can be expected to develop increased sensitivity to chemotherapy targeting DNA repair pathways. (Answer: True)

・Most cancer cells have variants of normal genes. (Answer: True)

・Abnormal regulation of the cell cycle is unrelated to the transformation of normal cells into cancerous cells. (Answer: False)

・The human genome does not change over one's life. (Answer: False)

The Basics of Genomic Medicine and Cancer Genomes

Chiaki Takahashi

Slide 1

The previous chapter featured an explanation of the manner in and mechanism by which somatic mutations occur in cancer genomes.

In this chapter, I will specifically discuss the genes found in somatic mutations in cancer genomes. In addition, I will talk about what we have learned from an analysis of the human genome as a result of advances in DNA sequencers. Last, I will talk about the relationship between variations in the human genome and cancer.

Slide 2

This figure shows the process of cancer progression from normal epithelial cells, where colon cancer originates, to the ultimate formation of a metastatic focus. This takes about 10-30 years. A series of genetic abnormalities accumulates, as shown in the figure, and pathological forms develop. In the case of colon cancer, groups of genes cause dysfunction due to somatic mutations. These groups include tumor suppressor genes such as APC and p53 and cancer genes such as KRAS, BRAF, and PIK3CA.

Cancer research has experienced various changes as a result of identification of the association between abnormalities in cancer genomes and pathologies. To start with, the classification of cancer is transitioning from a pathological or morphological classification in the past to a classification based on genomic information. Next, somatic mutations in cancer were found to be appropriate therapeutic targets. Then, the use of cancer genomic information allowed the optimal treatment to be designed for each patient. Now, cancer genomics has progressed dramatically as a result of the widespread use of next-generation sequencing, paving the way for accurate diagnosis and follow-up of cancer and the development of new treatments.

Slide 3

The discovery of groups of cancer genes originated with the discovery of viruses that cause tumors in chickens and rodents. Several of the oncogenes found in those viruses have been found to acquire somatic mutations and become active in human cancers as well. The Ras family of cancer genes is one example, and they are referred to as cellular proto-oncogenes. Today, we know of various cellular proto-oncogenes. In human cancers, those proto-oncogenes are activated by various mechanisms, such as a point mutation or abnormal expression, and they confer a growth advantage to tumor cells. As shown in the figure, many cellular proto-oncogenes act on various levels of

proliferative signaling pathways. In other words, these proto-oncogenes encode growth factors, receptors, signaling molecules, transcription factors, and various regulatory proteins. Proto-oncogenes also include genes that inhibit cell death.

Slide 4

The Ras family of proto-oncogenes has several members, and each has a somatic mutation found in various cancers. Mutations most often occur in the K-ras gene, and those mutations occur in virtually all bile duct cancers and pancreatic cancers. Ras family mutations are directly linked to a poor prognosis for various cancers.

Slide 5

I would now like to explain in simple terms the mechanism by which cancer develops as a result of a point mutation in a Ras proto-oncogene. Ras proteins are molecules involved in GTP-mediated signaling, so they are switched on when they bind to GTP. When Ras proteins hydrolyze GTP to GDP, they are switched off. However, a mutation in codon 11 or codon 61 of Ras causes Ras proteins to lose their ability to hydrolyze GTP, and the proteins remain switched off. Presumably this is how cell growth signaling is constantly turned ON and normal cells become cancerous cells.

Ras mutations are the most common genomic abnormalities in cancers, but they have yet to be targeted effectively.

Slide 6

I will now explain about tumor suppressor genes. Tumor suppressor genes can be divided into gatekeeper genes that slow down cell growth and that promote differentiation and caretaker genes that maintain chromosomal stability. Tumor suppressor genes originally acted to inhibit cancer but lost their original action due to a somatic mutation or deletion, thus allowing oncogenesis. A germline mutation in most of these genes is linked to familial cancer. Let me give you a typical example. RB1 is the first tumor suppressor gene to be identified, and a deletion of or mutation in RB1 causes retinoblastoma. A germline mutation causes retinoblastoma in both eyes while a somatic mutation causes retinoblastoma in one eye, so RB1 is a well-known example of the “two-hit theory” that predicts the presence of an oncogene. The product of this gene is a master regulator of the cell cycle and is known to be involved in controlling the terminal differentiation of cells.

Slide 7

Shown here is the relationship between RB1 and various oncogenes and tumor suppressor genes. Orange indicates oncogenes and blue indicates tumor suppressor genes. You can see how various forms of oncogenic signaling ultimately lead to control of RB1 functioning and how signaling includes

various oncogenes and tumor suppressor genes. Moreover, many oncogenes act in opposition to tumor suppressor genes. A mutation in the RB1 gene is somewhat uncommon, but RB1 is highly phosphorylated as a result of a large volume of oncogenic signaling and it loses its function, ultimately leading to activation of the cyclin D-CDK4 complex. Such a phenomenon is common to every cancer.

Slide 8

I will now explain about p53, which is another typical tumor suppressor gene. A mutation in p53 and inactivation of RB pathways are noted in various cancers. Most of the mutations in P53 are missense mutations, which stand in marked contrast to most of the mutations in tumor suppressor genes, which are frameshift mutations. A mutation in P53 is known to have a dominant-negative effect in inhibiting its formation of a tetramer, and mutated p53 is known to acquire new oncogenic activity. Normal p53 plays an important role as a guardian of the genome.

When cells experience events such as the depletion of nucleic acids, accumulating damage to DNA due to its exposure to ultraviolet radiation or ionizing radiation, enhanced oncogenic signaling (which causes overreplication of DNA that leads to damaged DNA), hypoxia (which induces the production of harmful reactive oxygen species or gene expression), or inhibition of transcription (which hampers cell homeostasis), p53 is activated. It arrests the cell cycle, it promotes the repair of DNA, or, if DNA is too extensively damaged to be repaired, it permanently arrests cell division and induces cells to die ("cellular senescence").

Slide 9

The oncogenes and tumor suppressor genes mentioned thus far are genes that were found to be related to cancer prior to advances in cancer genomics. However, long exome sequences were analyzed using the Sanger method, and then whole genomes and exome sequences were exhaustively analyzed using next-generation sequencing. This progress in cancer genomics has yielded various new findings.

One such finding with regard to cancer genomes is that a vast number of somatic mutations — from several dozen to more than several hundred per specimen – are found in coding regions. This will be shown in the next slide. The frequency of those mutations differs widely depending on the type of cancer. Even if two cancers were pathologically identical, sequencing has revealed that the type of mutation differed completely. That is, a tumor cannot be correctly diagnosed based on its pathological classification alone. Moreover, each type of cancer has a characteristic pattern of genomic changes. Analysis of cancer genomes has led to the concept of driver mutations and passenger mutations. This is described in detail in the coming slides. Of the several thousand or so somatic mutations identified in cancer genomes, only a few driver mutations contribute to oncogenesis. Genes that are believed to harbor driver mutations include genes that were previously believed to be

unrelated to cancer.

The fact that whole-genome analysis is possible has indicated that mutations in non-coding regions may be associated with the development of cancer. Slight chromosomal changes or copy number variations that could not be detected with conventional methods of chromosomal analysis can now be identified.

What is extremely important is that, based on various aspects of tumors in the same patient or based on sequencing of sites of recurrence or metastases in a patient, tumors are extremely heterogeneous in terms of the somatic mutations that led to them.

Slide 10

Each point in the figure represents a cancer specimen, and the red lines indicate the median value for somatic mutations found in each type of cancer. The types of cancer are arranged in order of those median values. It is hard to see, but melanoma is on the far right. Next in order are squamous cell carcinoma of the lung, adenocarcinoma of the lung, bladder cancer, and small cell lung cancer. In contrast, pilocytic astrocytoma, ALL, medulloblastoma, and AML are on the far left. Many genes that are known to harbor potent driver mutations are located on the right. Genomic instability increases, and somatic mutations accumulate. In addition, mutations may also be caused by exposure to a mutagen like tobacco.

Slide 11

Now that the concept of driver mutations and passenger mutations has come up, I would like to briefly talk about what those terms mean.

A driver mutation is thought to be harbored by a few of the genes found in somatic mutations in cancer genomes. This mutation confers a growth advantage to tumor cells. Detection of a duplication mutation in the same amino acid in multiple patients is crucial to a driver gene. These genes include many genes that are known to be closely related to oncogenesis, such as Ras and p53, but genes not thought to be related to oncogenesis also acquire driver mutations.

A passenger mutation is harbored by most of the genes detected in somatic mutations in cancer genomes. These mutations do not confer a growth advantage to tumor cells. Absence of a duplication mutation in the same cancers is crucial to distinguishing a passenger gene. There are examples of passenger genes that are driver genes, though they seldom appear. Distinguishing between driver genes and passenger genes will be a major issue in the future.

Slide 12

Shown here are 6 genes that are most often mutated in different types of cancer. These genes were identified via an analysis of various cancer genomes or exomes. Frequent mutations of PIK3CA and

BRAF were not anticipated prior to advances in genomics. These are good examples of genomics closing in on the etiology of cancer and its ability to identify promising therapeutic targets.

Slide 13

Cancer genomics has also led to the identification of new cancer-related genes. A group of genes involved in metabolism harbors somatic mutations in cancer. Mutations in IDH1 or IDH2 have been found in neuroblastoma, glioblastoma, cartilaginous tumors, and acute myelogenous leukemia. Mutations in the SDH gene have been found in paraganglioma and pheochromocytoma, and mutations in the FH gene have been found in rhabdomyosarcoma and renal cell carcinoma. IDH mutations are missense mutations. These mutations preclude IDH from producing 2-oxoglutarate as it normally would; instead, IDH produces 2-hydroxyglutarate. As a result, enzymes such as TET2 and JmjC that use 2-oxoglutarate as a cofactor are inhibited. This causes changes in the epigenetic profile and leads to oncogenesis. This was a hallmark discovery, i.e. that mutations in genes involved in regulating metabolism are associated with cancer. This was all thanks to cancer genomics.

Slide 14

One concept that has been definitively proven by cancer genomics is intra-tumor heterogeneity. Intra-tumor heterogeneity is a concept where genetic mutations or epigenetic changes occur in a monoclonal tumor, resulting in groups of cells following different paths of clonal evolution in that tumor. This concept was succinctly demonstrated in a paper published in 2012. The study described in that paper obtained multiple specimens of the primary focus and metastatic foci from a single patient and analyzed their genomes. Of the 128 genetic mutations that were identified, 1/3 were found at all of the sites. The only driver mutation was VHL, so groups of cells had presumably evolved from monoclonally initially tumors as a result of the VHL mutation. Mutations in SETD2, KDM5C, and mTOR were found in the primary focus, increasing its heterogeneity. Further heterogeneity was noted at site R4. Some of the cells had acquired a mutation in SETD2 or KDM5C, producing groups of cells with metastatic potential.

Slide 15

Results of cancer genomics and microarray analysis are disclosed to everyone and not just certain specialists. The Cancer Genome Project started in the UK in 2004, the TCGA started in 2006, and the ICGC started as a global project in 2007. Data from these efforts have, for example, helped to search for gene expression signatures that can be used in clinical diagnosis, to elucidate the process of oncogenesis, and understand inter-tumor and intra-tumor heterogeneity, as was mentioned earlier. I encourage you all to visit these sites and uncover some important information.

Slide 16

In this chapter, you learned about abnormalities found in cancer genomes. You learned that various mutations in oncogenes and tumor suppressor genes cause oncogenesis. You learned about how variations in cancer genomes acquired after birth cause cancer. At the end of this chapter, I will touch on variations found in the human genome, and I would like to discuss whether or not they are related to a risk of developing cancer. Listed here are variations found in the human genome. There are several types of polymorphisms and variations in the human genome. A polymorphism is a genomic structural change appearing in the general population at a frequency higher than 1%. A variation is a genomic structural change appearing in the general population, regardless of frequency. A representative polymorphism is a single nucleotide polymorphism, or SNP. An SNP is a change in a base occurring in 1% or more of the population. Ten million SNPs have been found in the human genome. SNPs are known to be involved in the expression of complex traits such as skin and eye color and build and to be associated with the incidence of lifestyle-related cancers and other non-communicable diseases. Please examine this chart carefully for the items to come.

Slide 17

So are variations in the human genome actively related to the development of cancer or not? An extremely close correlation has yet to be found, but certain SNPs are reported to significantly increase the risk of breast cancer, colon cancer, prostate cancer, and gastric cancer. In addition, deletion of the UGT2B17 gene on chromosome 4q13 has been found to increase the risk of prostate cancer. This region will yield new information as analysis of the human genome progresses in the future.

The Basics of Genomic Medicine and Cancer Genomes

Chiaki Takahashi

Slide 1

In this chapter, I will explain methods of detecting genomic abnormalities in cancer based on a structural analysis of chromosomes and chromosomal abnormalities found in specific types of cancer.

Slide 2

To start off, exactly what are chromosomes? A chromosome is a complex consisting of DNA located in the cell nucleus and proteins (primarily histones). A chromosome condenses DNA to 1/10,000 of its original length. There are 2 types of cell division, somatic mitosis and meiosis. Chromosomes are condensed in the metaphase of mitosis and have a morphology like that shown in the figure. During the metaphase, chromosomes align along the equatorial plane of a cell nucleus. If differential staining is performed at this point, then numerous bands are evident. Thus, chromosomes in the metaphase of mitosis are typically used to observe chromosomal morphology. Chromosomes can be stained using differential staining such as Giemsa staining to distinguish up to 2,000 patterns of bands, fluorescence in situ hybridization (FISH) to detect a hybridization probe (DNA) as a fluorescent signal, a combination of chromosome painting and multicolor fluorescence, spectral karyotyping (SKY) that stains 24 chromosomes with different colors, and array CGH that identifies a small abnormality in the copy number of several kb using a DNA chip with oligo DNA probes. Array CGH can detect amplification or deletion of various cancer-related genes at one go, and it has a high diagnostic capacity. Based on an analysis using such methods, chromosomal morphology can be categorized, in order of size, into gametes, autosomal chromosomes, sex chromosomes, short arms, long arms, telomeres, centromeres, and heterochromatin. The chromosomal makeup of an individual or cells is expressed as the karyotype.

Slide 3

What I want to show you here is chronic myeloid leukemia painted using SKY. As I will explain later, the Philadelphia chromosome – a reciprocal translocation of chromosome 9 and chromosome 22 — is a chromosomal abnormality noted in over 95% of chronic myeloid leukemias. SKY is a useful method of detecting such a translocation.

Slide 4

The entire human genome is around 3,000 Mb in size. Differential staining normally distinguishes around 400 bands and can distinguish up to 2,000 bands. Differential staining has a resolution of 5-

10 Mb. FISH and array CGH have a resolution of around 100 kb. Of course, these methods are far inferior to next-generation sequencing, which covers the whole genome at a resolution of singe bases. Chromosomal analysis is a method of testing that is extremely effective at detecting a disease-specific translocation such as that found in a hematopoietic malignancy or osteocartilaginous tumor, a chromosome deletion or duplication at the arm level, or aneuploidy. Compared to next-generation sequencing, chromosomal analysis is relatively simple and easy, so it can be used to detect and follow-up on residual lesions after chemotherapy. Thus, chromosomal analysis remains an essential test for the diagnosis and treatment of cancer.

Slide 5

Next, I will explain about various abnormalities in chromosomes. Aneuploidy is an abnormality associated with an increase or decrease in 1-2 chromosomes instead of the normal number of 46. A well-known example is trisomy 21 noted in Down's syndrome. The condition is caused by nondisjunction during cell division. In polyploidy, somatic cells are 3n or 4n instead of 2n. Most triploidy occurs when 2 sperm fertilize 1 egg. These abnormalities occur in the germline, and most are structural chromosomal abnormalities found in an individual as a whole. In contrast, temporary chromosomal abnormalities are temporarily noted in only cancer cells or some cells as a result of irradiation, viral infection, or exposure to a chemical substance. A break occurs in a chromosome, and an abnormality occurs when a fragment reattaches. Reciprocal translocation occurs when 2 or more chromosomes break and fragments are switched and then reattached to the other chromosome. An insertion is why a fragment is attached to another chromosome. A deletion is when a fragment is lost. An inversion is when a fragment is rotated 180 degrees and reattached. A ring chromosome is where the distant ends of arms beyond the cleavage site are lost and the remaining short arm and long arm are attached to one another. An isochromosome is where one arm is a mirror image of the other. Although it is not shown in the figure, a duplication is where part of a chromosome is "duplicated" lengthwise as a result of the unequal crossover of homologous chromosomes.

Slide 6

Shown here are typical temporary chromosomal abnormalities that are noted in tumors. A reciprocal translocation of chromosome 9 and chromosome 22 causes fusion of the BCR gene of chromosome 22 and the ABL gene of chromosome 9. This produces a fusion protein, tyrosine kinase, that is constantly activated, and it is the direct cause of Philadelphia chromosome-positive chronic myeloid leukemia. The drug imatinib has been developed to inhibit tyrosine kinase activity, and it is therapeutically effective. A reciprocal translocation of chromosome 8 and chromosome 14 places an enhancer of the immunoglobulin heavy chain region of chromosome 14 upstream of MYC, which is an oncogene on chromosome 8. This causes overexpression of the MYC protein, resulting in

Burkitt's B cell lymphoma.

Slide 7

Various reciprocal translocations that produce oncogenic fusion proteins, as exemplified by BCR-ABL, are known to be mainly associated with hematopoietic malignancies.

Slide 8

Oncogene overexpression due to enhancer translocation, as exemplified by overexpression of MYC in Burkitt's B cell lymphoma, is primarily noted in hematopoietic malignancies, as would be expected. However, such oncogenic chromosomal abnormalities are seldom found in solid epithelial tumors.

Slide 9

That said, Dr. Mano, now at the National Cancer Center of Japan, and his colleagues introduced a cDNA library from lung cancer into NIH-3T3 cells and they used classic methods to search for transformable genes. Dr. Mano and his colleagues identified a new type of EML4-ALK gene. This was the first oncogenic chromosomal abnormality found in lung cancer.

Slide 10

An inversion occurs in chromosome 2, fusing the EML4 gene with the ALK gene to produce an EML4-ALK fusion protein. As a result, ALK tyrosine kinase activity is constantly elevated. This chromosomal abnormality accounts for around 6% of non-small cell lung cancers, and it is effectively treated with an inhibitor of ALK tyrosine kinase. Chromosomal abnormalities were thought to be largely absent in cancers other than hematopoietic malignancies and sarcomas, but the discovery of that fusion protein showed that chromosomal abnormalities cause epithelial carcinoma.

Slide 11

Following the EML4-ALK fusion gene, fusion genes such as LRIG3-ROS1 and KIF5A-RET have been identified in adenocarcinoma of the lung. KIF5A-RET accounts for around 2% of adenocarcinomas of the lung. Hopes are high that it can serve as a new therapeutic molecular target. And now this chapter has concluded.

Hokushin e-learning for cancer specialists
Genomic Medicine
The Basics of Genomic Medicine and Cancer Genomes
Cancer Epigenomes

Takeshi Suzuki (The Cancer Research Institute of Kanazawa University)

Slide 1
This e-learning lecture on the Basics of Genomic Medicine and Cancer Genomes, Chapter 4, Cancer Epigenomes will now begin.
I am Takeshi Suzuki of the Cancer Research Institute of Kanazawa University and I will be teaching this chapter.

Slide 2
Based on what you have learned thus far, you now understand that cancer is a genetic illness. Cancer develops as a result of gene mutations and abnormal gene expression, so genomic changes, i.e. changes in genetic information, are crucial. Genomic changes are irreversible and are consistently passed on to the next generation of cells via cell division. In the past, cancer was thought to develop primarily due to genomic changes. However, genomic changes are typically infrequent, and those changes alone are unlikely to cause cancer to directly develop. What explains this are epigenomic changes. With epigenomic changes, the pattern of expression changes but genetic information itself does not change. Epigenomic changes are reversible. Thus, one could envision a treatment strategy to return the epigenome to its normal state, so these changes have garnered attention as a target for the next generation of cancer therapies. Just what is an epigenome? An epigenome is a concept referring to epigenetic information as a whole. I will start by explaining what epigenetics is.

Slide 3
Epigenetics is the mechanism causing phenotypic changes without causing changes in gene sequences. Epigenetics is mechanism to decipher the same genomic information in various ways. The human genome codes for around 22,000 genes. However, there are around 200 types of cells, such as neurons and muscle cells, in the human body. Every cell has the same genome and the same genetic information, but why are multiple cells with completely different phenotypes produced? The answer is the particular manner in which genes are expressed in each of those

cells. The mechanism that determines ow genes are expressed is epigenetics.

Slide 4

So what sort of mechanism is epigenetics? What exactly is an epigenome? Do cells contain epigenomic information? If so, in what state? Those questions must be answered. There are 2 simple answers, posttranslational modification of histones and methylation (modification) of genomic DNA. Histones are basic proteins that comprise nucleosomes, which are the basic units of chromatin. Histones bind to DNA and store DNA in the nucleus of a cell. These histone proteins are chemically modified, e.g. they are methylated or acetylated. Modification of histones and methylation of DNA can be reversed or counteracted, but those changes can be consistently passed on to new cells after cell division.

Slide 5

I will now explain about histone proteins in detail. The human genome contains around 6.4 billion base pairs. DNA has a total length of about 2 meters. DNA is stored in the nucleus, which has a diameter of 5 to 10 micrometers. What makes that possible is a structure known as chromatin. Chromatin is primarily a complex of DNA and histone proteins. The basic units of chromatin are nucleosomes. The 4 histones H2A, H2B, H3, and H4 are known as core histones, and an octamer or histone octamer contains 2 copies of each of the 4 core histones. A nucleosome is DNA wrapped around an octamer. Chromatin produces histones, and the structure of chromatin is vital to every event involving DNA, including control of transcription and DNA replication and repair. Epigenetics can be referred to as chromatin biology. As an example, chromatin exists in a less condensed form as euchromatin in a genomic region with abundant gene expression while it exists in a condensed form as heterochromatin in a region where gene expression is quiescent.

Slide 6

Once histones are synthesized, they undergo various types of chemical modifications. These modifications include acetylation, methylation, phosphorylation, and ubiquitination. Certain amino acid residues are chemically modified, and the main residues are listed here for each histone. Lysine residues and arginine residues undergo acetylation and methylation, serine residues and threonine residues undergo phosphorylation, and lysine residues undergo ubiquitination. Research on the relationship between posttranslational modification of the histone H3 and its function has proceeded. This has revealed that the amino terminus referred to as the tail undergoes various modifications, resulting in a variation of functions.

Slide 7

Epigenetic marks such as histone modification are consistently passed on to future generations of cells, but they change dynamically in accordance with the stage of development and environment. まず, Writers are induced by guides called initiators, which are molecules known as long-chain non-coding RNAs, to leave a mark at a specified site. Readers are molecules that interpret those marks, which erasers are molecules that erase those marks. Remodelers are molecules that alter the position of nucleosomes and the structure of chromatin, and insulators are molecules that form boundaries limiting the extent of epigenetic control. Dynamic epigenetic control and control of chromatin function are achieved these molecules fulfilling their various roles.

Slide 8

The most well-known type of control is the histone code. The histone code refers to the hypothesis that a combination of chemical modifications to histones induces chromatin-specific functioning. As an example, there are lysine residues at 4 locations at the tail end of the amino terminus of histone H3 that mainly undergo methylation (modification). There are enzymes that methylate individual lysine residues and demethylating enzymes that remove methyl groups. Lysine 4 is referred to as K4 and lysine 36 is referred to as K36. Methylation of K4 and L36 is related to the activation of transcription, while methylation of K9 and K27 is linked to the suppression of gene expression.

Slide 9

These enzymes and proteins are the writers, erasers, and readers involved in the methylation of the histone H3K4. Writers are histone lysine methyltransferases (KMTs); these enzymes add a methyl group to K4. Erasers are histone lysine demethylases (KDMs). Readers include proteins with a PHD domain that recognize methylation sites, and readers serve as an effector of methylation (modification) signaling. Methylation of K4 is related to the activation of transcription.

Slide 10

Next are proteins involved in the methylation of K27. A well-known writer is the enzyme EZH2. EZH2 is an important enzyme that is closely associated with the development and malignant transformation of cancer. Enzymes that typically control histone modification form a protein complex consisting of multiple proteins. EZH2 is a subunit of polycomb repressive complex 2 (PRC2) that catalyzes the methylation of K27. K27 erasers include demethylating enzymes such

as UTX. A well-known reader is the CBX7 protein, which is a component of polycomb repressive complex 1 (PRC1). The CBX7 protein recognizes and binds to K27 methylated by PRC2 and it monoubiquitinates the core histone H2A, thus inducing the suppression of gene expression.

Slide 11

The concept of the histone code was accepted about 10 years ago. Since then, a succession of studies has reported abnormalities in the methylation (modification) of histones and abnormalities in the enzymes regulating methylation in cancer. To summarize the major finding of those studies, abnormalities in most of the enzymes that methylate and demethylate the histone H3 have been detected in cancer. Most abnormalities related to gain of function (GOF) involve high levels of expression of those enzymes in cancer while abnormalities related to loss of function (LOF) involve a mutation in or suppressed expression of those enzymes. An interesting fact is that the same enzyme, e.g. the enzyme EZH2, can have GOF or LOF depending on the type of cancer. That is, a mutation linked to the development of cancer can act in the exact opposite manner as a result of differences in the cell context. Enzymes control which genes to target, but they can act differently depending on the cell type and conditions.

Slide 12

Next, we will look at abnormalities in the regulation of H3K27 methylation in detail. EZH2 is an enzyme that methylates K27, and it is a catalytic subunit of the complex PRC2. EZH2 is an oncogene. This is because amplification of the EZH2 gene reveals high levels of its expression in many solid tumors and because mir101, a microRNA that suppresses EZH2 expression, is detected in cancer. Moreover, a GOF variant of EZH2 that has increased enzymatic activity as a result of a point mutation is detected in leukemia, and an LOG variant of UTX, an enzyme that demethylates K27, is detected in numerous cancers. This is evidence that increased methylation of K27 is associated with cancer. In contrast, an LOF variant of EZH2 has been identified in some leukemias and lymphoid tumors, suggesting that EZH2 also acts as a tumor suppressor gene. That is, EZH2 can act as an oncogene or as a tumor suppressor gene depending on differences in the cell context. Thus, abnormal K27 methylation in cancer is not a simple issue. Instead, caution is required when formulating a treatment strategy to target that abnormality.

Slide 13

Next, I will explain about the acetylation (modification) of histones and their functions. These

proteins are involved in the acetylation of histones. Writers are histone acetyltransferases (HATs), and the protein CBP is a well-known writer. Erasers are histone deacetylases (HDACs), and there are numerous types. Readers include proteins with a bromodomain that recognize acetylation sites, and readers facilitate acetylation signaling. One characteristic that differentiates methylation from acetylation is that both groups of enzymes involved in the acetylation of histones do not specifically act on amino acid residues at certain sites. Using the histone H3 as an example, lysines such as K9 and K27 are acetylated, but certain enzymes are not restricted to acting only on certain sites. Another characteristic is that acetylation is basically related to the activation of transcription, regardless of the site that is acetylated. This differs completely from control via methylation.

Slide 14

So how is the acetylation of histones controlled in cancer? The mechanism is simpler than that for methylation. Mutations and fusion genes are detected in cancer and produce LOF variants of writers, i.e. acetylating enzymes. In contrast, there are few mutations in erasers. An abnormality occurs when these enzymes are overexpressed or they are mistakenly transported to a gene other than the original target gene due to a mutation in a transcription factor. Regardless of the change, the abnormality ultimately decreases acetylation, causing the suppression of gene expression. The relationship between that phenomenon and cancer is being discussed.

Slide 15

Next, I will talk about the methylation (modification) of DNA separate from histones. DNA methylation refers to substitution of a methyl group for the hydrogen bound to the fifth carbon of cytosine, which is one of the bases that make up DNA. The enzymes responsible for this reaction are known as DNMTs. A reaction to remove a methyl group is known as demethylation. Until recently, active DNA demethylation pathways as shown here were unclear. Previously, a reaction known as passive demethylation was considered to be the only pathway for DNA demethylation. In other words, DNA was replicated and cytosine was taken up by the newly synthesized DNA strand without being methylated. If replication was repeated while cytosine was still unmethylated, demethylation of DNA would occur spontaneously. As was recently revealed, what holds the key to the active demethylation of DNA is an enzyme known as TET that controls the initial stage of the reaction.

Slide 16

DNMTs are writers of DNA methylation and can be divided into 2 groups based on their action.

The first is DNMT1, which is a maintenance methylase. During DNA replication, unmethylated cytosine is taken up by the complementary strand and the cytosine site is partially methylated. The maintenance methylase known as DNMT1 methylates the complementary strand. The second group of DNMTs includes DNMT3A and DNMT3B, which are de novo methylases that attach a methyl group to unmethylated DNA. The TET family of enzymes, which are involved in active demethylation, are important erasers. Readers are well-known proteins, such as MBDs, that can recognize and bind to methylated cytosine and that are involved in the regulation of gene expression.

Slide 17

An important aspect of the regulation of gene expression via DNA methylation is the concept of CpG islands. A CpG island is a place where CG dinucleotides often appear in a genome. There are CpG islands in promoter regions and expression-regulating regions of about half of human genes. When a CpG island promoter is not normally methylated, a gene is constantly expressed. When, in contrast, a CpG island promoter is highly methylated, gene expression is consistently suppressed.

Slide 18

The abnormal methylation of genomic DNA has long been studied in cancer. It has 2 characteristics. One is that a decrease in the methylation of DNA occurs in a wide range of genomic regions. The other is that extensive methylation is noted in a local CpG island promoter. In normal cells, for example, heterochromatin regions near the centromere at the center of a chromosome are hypermethylated, and gene expression is constantly turned OFF. In cancer cells, hypomethylation of DNA occurs genome-wide. Repeats have an open chromatin structure, and transposons not expressed in normal cells are expressed. The frequency of recombination and genomic instability increases. In normal cells, genes are expressed while the CpG island promoter is hypomethylated, but local hypermethylation occurs in cancer cells, and gene expression is constantly turned OFF. When this occurs due to a tumor suppressor gene, the tumor suppressor gene is silenced, leading to oncogenesis.

Slide 19

Next, we will look at the relationship between the TET protein and cancer in relation to DNA demethylation. An LOF variant of the TET2 enzyme is often detected in leukemia and glioblastoma. In these cancer cells, hypermethylation of CpG islands is induced due to diminished demethylation. In the same types of cancer, mutations in the isocitrate

dehydrogenase enzymes (IDH1 and IDH2), which are involved in glycolysis and the TCA cycle, are also detected. These enzymes convert isocitrate to α-ketoglutarate, but the enzyme variant produces a metabolite known as 2-hydroxyglutarate (2HG). 2HG competitively inhibits the activity of the TET family of enzymes (which are α-ketoglutarate-dependent) and histone-demethylating enzymes with the Jumonji domain and subsequently induces CpG island hypermethylation. An LOF variant of TET2 and an IDH variant are mutually exclusive in cancer. In other words, hypermethylation of CpG islands is crucial to the development of cancer. Incidentally, a substance like 2HG, i.e. a metabolite associated with cancer, is referred to as an oncometabolite.

Slide 20

Well, I have explained about epigenomic changes in cancer. As I initially mentioned, epigenomic changes are reversible, so one could envision a treatment to return the epigenome to its normal state. Here, I will describe well-known examples of treatments targeting the epigenome. One treatment involves azacitidine and decitabine, which are DNA methylase inhibitors. These compounds inhibit every form of DNMT3. When used at a low concentration, they are efficacious in treating MDS myelodysplastic syndrome, so they have been approved by the FDA. Another treatment involves the histone deacetylase (HDAC) inhibitor vorinostat. This drug is approved for treatment of T cell lymphoma. These drugs halt the silencing of tumor suppressor genes by reducing the methylation of DNA or increasing the acetylation of histones. This action inhibits the proliferation of cancer cells. Epigenomic drug discovery has garnered attention as other compounds targeting the methylation of histones are being successively developed.

Slide 21

This slide summarizes the content of this lecture. In cancer, genomes and epigenomes influence one another. Epigenome regulators are affected by mutations in the genome. Subsequent epigenomic changes pave the way for genomic changes. Mutations or instability in the genome and epigenomic abnormalities act in a coordinated manner and are involved in determining various phenotypes associated with cancer and characteristics of cancer cells を involved in determining, as I hope you now understand.

With that, I'd like to conclude my lecture on cancer epigenomes.