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.