CARCINOGENESIS

Arno Helmberg

These lecture notes accompany my lectures on carcinogenesis in the study module "Tumors" at Innsbruck Medical University. The English version serves two purposes: as a learning aid for international students and to encourage German-speaking students to familiarize themselves with medical English; the lectures are delivered in German. The translation from the original German version is my own; I am afraid it will occasionally sound appalling to native English speakers, but it should at least be intelligible.

Our body is a society of cells. As in any functioning society, members have responsibilities and have to adhere to rules and regulations. In cells, the ability to perform tasks and to comply with regulations is implemented by gene expression. Cancer is caused by a somatic cell clone's "loss of rules", caused by genetic and epigenetic changes. A typical malignant cell develops over years in a step by step "microevolution" process, accumulating five to ten critical mutations in genes important for cell division, apoptosis, DNA repair and other aspects of cell behavior. Each of these mutations results in an additional small selective advantage vis-à-vis fellow cells, over time producing a rogue cell clone able to bring down the entire organism.

1. NORMAL GROWTH REGULATION: PROTO-ONCOGENES

When do cells proliferate?

Let's start by considering a simple, unicellular organism. Imagine a yeast cell, put into an optimal environment with lots of nutrients, oxygen and everything else the cell requires. What is going to happen? The cell will replicate as long as the environment allows it to do so.

As soon as multicellular organisms developed, a dilemma had to be solved. On the one hand, it was necessary to maintain favorable conditions for all cells, e. g. to constantly provide glucose and oxygen. On the other hand, a way had to be found to prevent cells from proliferating uncontrollably, blowing up the organism. While before, cell replication was mainly regulated by the availability of limiting nutrients, multicellularity required completely novel mechanisms to control cell proliferation.

To maintain a long-living multicellular organism, like the human one, is a complex task. Many of its cells have a limited life span. Lost cells have to be replaced, but that is only part of the problem. In addition, tissues need plasticity, enabling them to adapt to changing physiological demands. As an example, let's consider the total of red blood cells in our body. Over time, this total should be largely constant as long as personal circumstances --e. g., studying medicine in Innsbruck, Austria-- don't change. Yet, if the same person decides to go on a vacation at high

altitudes, e. g., for a trekking tour in Nepal, she requires a lot more erythrocytes. How to solve such a problem?

While for most tissues we don't understand the homeostatic mechanisms, we do for red blood cells. In an elegant feedback mechanism, a reduction in kidney partial oxygen pressure (pO_2) stimulates production of new erythrocytes in the bone marrow. Like most cells, kidney cells constantly produce a transcription factor, hypoxia inducible factor (HIF-1), which at normal pO_2 is broken down fast and efficiently. Breakdown via the ubiquitin-proteasome pathway is initiated by an oxygen-dependent enzyme, a prolyl-hydroxylase. This enzyme just barely works as long as there is enough oxygen in the cell. As soon as pO2 in the highly perfused kidney sinks below a certain threshold, the enzyme ceases to function. HIF-1 stops to be hydroxylated and starts to accumulate. In the nucleus, accumulating HIF-1 activates its target genes, including the gene encoding erythropoietin (EPO). In summary, this sensing mechanism translates a fall in kidney pO_2 into a secretion of EPO.

EPO increases the rate of proliferation in erythroid progenitors in the bone marrow (colony forming units, CFU_E and burst forming units, BFU_E). For the extracellular growth factor EPO to accelerate cell proliferation, a series of characteristic steps is induced, involving signal transduction and gene regulation, ultimately leading to changes in proteins regulating the cell cycle. The general pattern is typical for the response of proliferation-competent cells to growth factors; genes and proteins involved in these steps are critical to carcinogenesis.

In erythroid progenitor cells, EPO crosslinks two transmembrane erythropoietin receptor molecules, each of which is associated with an intracellular JAK2 kinase. The two kinases first phosphorylate each other, then several tyrosines of the receptor, creating docking sites for further proteins. One of these proteins is a member of the STAT (signal transducer and activator of transcription) family, STAT5. In turn, phosphorylated STAT5 molecules dimerize and translocate to the nucleus. They directly bind to DNA, activating genes promoting proliferation, e. g., genes encoding the transcription factor c-Myc and a protein called Cyclin D.

The net effect is an increase in erythrocyte output, eventually increasing hematocrit or, in other words, the blood's oxygen-transport capacity. In time, this leads to normalization of pO_2 in the kidney, reactivating the oxygen-dependent hydroxylase. As soon as HIF-1 is degraded again, the negative feedback loop is closed.

This example illustrates a general principle: proliferation is not something that just "happens". It is regulated stringently and occurs only in situations where new cells are required. This requirement is measured by other cells, which communicate it to the proliferation-competent cells via extracellular signaling molecules generally called growth factors.

Cell division versus differentiation

In our example, the division of labor is obvious. On the one hand, we have cells that do the work (transporting oxygen and CO_2), but can't proliferate: red blood cells. On the other hand, we have cells that are able to proliferate, but lack the ability to do specialized work: erythropoietic progenitors. For other tissues, this division is equally true, though less obvious. Usually, only a small part of the cells of a tissue are able to proliferate: stem cells and a few additional generations of cells. In the intestinal epithelium, stem cells are located at the bottom

of crypts; fast-proliferating *transit amplifying cells* fill the wall of the crypt. All proliferation occurs in response to growth signals. Stem cells divide only occasionally, as rarely as possible. In mature tissues, they divide asymmetrically. Only one of the daughter cells remains a stem cell, replacing the original one to maintain the stem cell pool. The second daughter cell gives rise to a clone of rapidly proliferating transit amplifying cells.

In the bone marrow, the prevailing pattern of growth factors nudges transit amplifying cells in the direction of specific colony forming units, which give rise to only one of the hematopoietic lineages. At this stage, the number of divisions per unit of time is strongly influenced by growth factors. Erythropoietin, for example, exerts its effect at this level. Following a phase of clonal expansion encompassing a limited (10-15) number of generations, cells shut down their ability to proliferate and enter a phase of terminal differentiation. From then on, cells only mature into their final state.

In summary, the first stem cell-generated transit amplifying cell gives rise to a relatively short-lived clone leading to a large number of differentiated worker cells. Mutations arising in cells of this clone are of little importance, as these cells have limited proliferative potential and limited survival time. Mature erythrocytes live for about three months, intestinal epithelial cells for a few days only.

Stem cells

For cancer to develop, a number of mutations has to accumulate in a cell lineage. This process takes many generations of cells, and is therefore much more likely to occur in the stem cell lineage. It is therefore assumed that for many types of cancer, the initial events occur at the stem cell level. Interestingly, in many cases, cancerous tissue maintains the division into (cancer) stem cells, transit amplifying cells and more or less differentiated cells. As we will see, certain properties of stem cells make them especially hard to attack by therapeutic means.

Several safeguards are in place to minimize accumulation of mutations in stem cells:

1. Stem cells divide as rarely as possible to minimize the number of rounds of DNA replication, which is inherently error-prone. Many stem cells seem to spend most of their lives in a state called dormancy or quiescence. They are activated only in case of a requirement for additional cells. This property makes cancer stem cells hard to eradicate, as many of our therapeutic options work best in cells that are actively proliferating.
2. Stem cells reside in niches offering anatomical protection: hematopoietic stem cells are located in bone, protected from ionizing radiation. Stem cells of the epidermis have to be protected primarily from UV radiation, so they are tucked away deeply at the bulge of hair follicles. Stem cells from the intestinal epithelium sit at the bottom of crypts, as far away as possible from the noxious contents of the gut and protected by a mucus escalator moving away from them.
3. Stem cells express high levels of P-glycoprotein, a transmembrane protein encoded by the MDR1 (multi-drug resistance) gene able to pump questionable alkaloids out of the cells. While this mechanism is highly beneficial in the normal situation, it makes cancer stem cells particularly hard to attack: many chemotherapeutic agents do not reach adequate intracellular concentrations.
4. Some stem cells, like those of the intestinal epithelium, have a lowered threshold for apoptosis. DNA damage in an intestinal stem cells causes the cell to enter apoptosis. The

logic seems to be: better to lose a stem cell than to risk replication of a damaged genome. This causes problems in radiotherapy. If an intestinal loop is irradiated inadvertently, all of its stem cells are forced into apoptosis, causing loop necrosis with peritonitis a few days later.

5. Finally, observations suggest that in some types of stem cells, the original DNA single strand may be kept within the stem cell lineage like a family heirloom. Semiconservative replication implies that misincorporations preferentially affect the new strand. If the new strand is systematically handed to the transit amplifying daughter cell, and the original strand systematically given to the stem daughter cell, this mechanism helps to minimize mutations in the stem cell line.

Cell division is a program involving activation and silencing of multiple genes

Once a cell "decides" to divide, its inner workings are turned upside down. Many genes active in G0/G1 have to be silenced, while others need to be switched on, e. g. those with specialized functions in DNA replication and mitosis.

From extracellular signal to changes in gene expression: proto-oncogene classes

Information reaching the cell in the form of growth factors needs to be processed at several levels to result in coordinated changes in gene expression and in a structured cell division. Proteins involved in these processes are vulnerable to be activated by mutations. In their activated, oncogenic form, they generate false proliferation-stimulating signals. In their normal, physiological form these genes are called proto-oncogenes ("pre-oncogenes"). Typically, encoded proteins function in one of the following capacities:

Class I: growth factors
Class II: growth factor receptors
Class III: signal transductors
Class IV: regulators of transcription
Class V: components of the machinery controlling the cell division cycle

The essence of proto/oncogene activation: a mutation feigning a growth signal

Let us look at signal processing in response to an extracellular growth factor. Physiologically, involved proteins can be switched between "off" and "on" states, depending on the presence of the growth factor.

Specific mutations may "freeze" these proteins in the "on" state. Switching off the signal becomes physically impossible. For the cell, this results in the "impression" that there is a lot of growth factor around. Hence, the cell reacts in the way it is programmed to: by proliferating.

For erythropoietin signal processing, there is a easy-to-grasp real-world example. The mutation Val617Phe in JAK2 results in continuous activation of this kinase, feigning continuous presence of EPO or other growth factors. This mutation is common in a group of myeloproliferative diseases including *Polycythemia vera*. In *Polycythemia vera*,

overproduction of red blood cells has a blood-thickening effect, greatly increasing the risk of thrombosis.

It is important to note that only a tiny fraction of mutations in proto-oncogenes have an activating effect. Usually, the opposite is true: mutations result in proteins that don't work anymore (loss-of-function). In that case, the only result is that the respective cell loses its ability to react to a growth factor. Therefore, loss-of-function mutations in proto-oncogenes are irrelevant for carcinogenesis. Only the rare gain-of-function mutations are able to transform a proto-oncogene into an oncogene.

Think of an automobile parallel: most defects will stop a car from functioning. This is a nuisance, but not immediately dangerous. Compared with these common problems, activating defects in cars are rare, but not unheard of, like a jammed accelerator: this is much more critical.

2. MUTATIONS

Carcinogens

At the beginning of the 20th century, a link was recognized between certain substances and cancer. In one example, men who were engaged in distilling the bicyclic aromatic substance 2-Naphthylamine later developed cancer of the bladder. Obviously, this substance somehow caused cancer. Such substances were termed "carcinogens".

Many people are afraid of "chemicals", which they suppose are the most relevant causes of cancer. Yet, this is a misconception. For example, **Aflatoxin B1**, one of the most relevant carcinogens, is a purely natural product. It is produced by the fungus *Aspergillus flavus*, which thrives on peanuts, corn, grain or pistachios, if these foods are stored in warm or humid conditions. In many tropical and subtropical parts of the world, this is inevitable for lack of better storage infrastructure. Aflatoxin by itself is not mutagenic. After its ingestion with food, it is taken up via the gut and reaches the liver via the portal vein. In hepatocytes, it is oxidized by the cytochrome P450 enzyme system to a highly reactive intermediate, Aflatoxin-epoxide, which proceeds to bind covalently to nitrogen or oxygen atoms in cellular macromolecules. A typical acceptor in DNA is the N7-atom of guanine. The combination forms a bulky complex termed a DNA adduct. We have efficient repair systems to deal with such adducts, but if the next DNA replication fork comes first, misincorporation may occur. The DNA-aflatoxin-adduct is able to form hydrogen bonds not only with the correct cytosine, but also with adenosine; yet the replication fork seems to get stuck with a C, while it is able to hobble past an A. Altogether, the adduct favors misincorporation of an A. Enter the repair system, which removes the adduct and resynthesizes the strand, placing a T opposite the A. In summary, the process replaces an original G with a T. Likely, this mechanism is responsible for the G→T mutation at the third position of codon 249 of the tumor suppressor p53. This mutation, which is frequently observed in regions burdened by aflatoxin-contaminated food, exchanges Arginine 249 (AGG), which is important for DNA-binding, for a Serine (AGT), inactivating the tumor suppressor. This process is thought to contribute to the high frequency of hepatocellular carcinoma observed in the tropics.

Heating protein-rich food causes numerous chemical reactions. When cooking meats or fish at high temperatures, **heterocyclic amines** (HCAs) such as PhiP (2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine) are formed. PhiP causes mutations much like aflatoxin. Oxidation of PhiP's exocyclic amino group by a cytochrome P450 oxidase generates highly reactive N-OH-PhiP, which again reacts with guanine to form a DNA adduct.

In today's society, potential "hidden" environmental carcinogens are the focus of a lot of anxiety. At the same time, people voluntarily expose themselves to established carcinogens all the time: take carcinogens in tobacco smoke or the mutagenic effect of UV irradiation.

Polycyclic aromates like benzo[a]pyrene are part of the "tar" component of cigarettes. Mainly, these mutagens cause lung cancer, as they reach their highest concentrations in a smoker's bronchial epithelia. The formal mechanism is analogous to that of aflatoxin: benzo[a]pyrene forms DNA adducts with guanine, leading to misincorporation of adenosine. In a heavy smoker, mutagenic effects are so pronounced that there is virtually no need to worry about "hidden" carcinogens. In addition to polycyclic aromates, numerous other carcinogens have been identified in tobacco smoke: **aldehydes**, **nitrosamines** and **heavy metals**, even the radionuclide Polonium 210, which is formed by Uranium-Radium decay (Uranium is a contaminant of phosphate used as fertilizer). Nicotine in cigarettes is primarily addictive; while mutagenic derivates like N'-nitrosonornicotine have been identified, their carcinogenic effect is small compared to the total mutagenic load of tobacco smoke.

Mutation-causing radiation

Normal exposure to the sun leads to a large amount of DNA damage in the skin. The energy of **UV rays** induces covalent bonds between adjacent pyrimidines (resulting in, e. g., thymine dimers). The intensity of this constant mutagenesis becomes obvious in patients suffering from *Xeroderma pigmentosum*, who cannot repair these DNA lesions. Even with maximum possible UV protection, patients develop multiple skin cancers, sometimes even at the tip of the tongue or in the anterior part of the oral cavity! A certain amount of UV is healthy and required for production of vitamin D, but most of us absorb a lot more than necessary. It would be smart to use sunscreen and to limit the amount of direct exposure to the sun. From this perspective, solariums do not seem like a good idea.

Ionizing radiation causes single and double strand breaks in DNA. Everybody is exposed to a certain amount of ionizing radiation, depending on medical issues, geological properties of her place of residence and building materials used in home construction. Occupational issues (medical professions, frequent intercontinental flying) may increase exposure.

Assaying mutagenic potential: the Ames test

To test individual substances for mutagenicity, American scientist Bruce Ames developed an assay to quantify induced mutations in a bacterial strain. The test strain fails to grow on agar lacking histidine, due to a defect in a gene necessary to produce this amino acid. However, a mutation hitting the right spot is able to correct the defect, allowing to quantify mutation rates by counting revertant bacterial colonies that are able to grow in the absence of histidine. By comparing the amount of revertant colonies in the presence and absence of a test substance, it

is possible to estimate relative mutagenicity. To include metabolites of the test substance, homogenized rat liver extract, containing a range of cytochrome P450 oxidases, may be added.

Some mutations originate in the absence of exogenous mutagens

Mutations are not exclusively caused by exogenous mutagenic substances or radiation. Every day, DNA in a normal, healthy cell suffers about 20,000 instances of "spontaneous" damage which, if not repaired correctly, may lead to secondary mutations. Double stranded DNA is made to last; it is much more stable than RNA or proteins. Still, it works in an environment that is chemically quite reactive. Temperature, oxygen, ion-rich aqueous environment (compare conditions for the formation of rust!) and a plethora of reactive groups of surrounding macromolecules result in frequent chemical reactions. In the following, the most frequent modifications:

Hydrolysis
The combination of aqueous environment and relatively high temperature leads to the occasional splitting of a covalent bond by water. As some bonds are more susceptible to hydrolysis than others, two main forms of damage are observed:
Depurination: by hydrolysis of the N-glycosidic bond between purine and deoxyribose, the DNA of every human cell loses about 5000 purine bases every day. The result is an *abasic site*. If the next DNA replication is faster than repair, the abasic site functions as a wild card, allowing any base to be incorporated in the opposite strand.
Deamination: If cytosine loses its amino group by hydrolysis, what remains is uracil. This happens about 100 times per day in each cell. While cytosine pairs with guanine, uracil pairs with adenine. As uracil has no place in DNA, it is very efficiently eliminated by a specialized enzyme, uracil DNA glycosylase, followed by resynthesis of the strand. Only if this repair system, a form of base excision repair, is defective, deamination of cytosine results in a G→A point mutation at the opposite strand.

Oxidation
Oxygen is highly reactive, especially in the form of a hydroxyl radical, causing more than 10,000 base modifications per cell and day. Some of these modifications cause mutations. Oxidation of guanine to 8-oxo-guanine causes it to pair with A instead of C. Other oxidative base modifications, like the formation of thymine glycol, result in stalled replication forks. Oxidative damage to bases is already found under physiological conditions. It is strongly enhanced in chronic inflammation, which entails continuously elevated levels of reactive oxygen species from activated macrophages and neutrophils. In addition, resulting cell death drives regenerative proliferation, which acts to fix oxidative DNA damage by misincorporation during DNA replication.

Methylation
Controlled methylation of DNA on the C5-atom of cytosine in a CG doublet ("CpG" configuration) is an important mechanism in gene regulation. In addition, unregulated methylations occur at different atoms of all bases by transfer of methyl groups from S-adenosyl methionine. Some of these methylations cause misincorporation. While, e. g., the more frequent methylation of guanine on N7 does no harm, the rare methylation of O6 causes guanine to incorrectly pair with T instead of with C.

Replication errors

Obviously, while mutations happen at other points in time, they occur most easily during DNA replication. Later, we will look at a special, complex type of replication error: a lagging chromosome leads to formation of a separated micronucleus. This sequestration retards the chromosome's next replication, causes its fragmentation and results in chaotic reassembly. Here, we focus on misincorporation of individual bases. Normal nucleotides may be misincorporated at incorrect positions; in addition, "faulty", modified nucleotides, which are always present in small amounts due to reaction equilibria, are incorporated occasionally. To keep the resulting rate of mutations as low as possible, several layers of quality control are implemented. The first layer is the so-called proofreading activity of the DNA polymerases themselves. In most cases, a misincorporated nucleotide is immediately removed by the 3'→5' exonuclease activity of the polymerase before elongation proceeds. The second layer of corrections is provided by the mismatch repair (MMR) system.

Mismatch repair (MMR)

Replication errors include misincorporations, but also incorporation of too many (insertions) or omission (deletions) of nucleotides, producing small loops. Insertions or deletions are even more critical than simple misincorporations, as most of them imply reading frame shifts. With the large number of nucleotides polymerized in each DNA replication, even a low rate of error results in many incorrect nucleotides. Mismatched, that is, incorrect, additional or missing nucleotides cannot form correct hydrogen bonds with their counterparts, locally disturbing the geometry of the DNA double helix. This discontinuity in structure is easily recognized by molecular means; it is much harder to decide which of the two strands to correct. Mismatch repair systems of diverse organisms have the ability to make an educated guess on which of the two strands is old (this one is usually the correct one) and which one has been newly synthesized (usually containing the replication error). It is not yet entirely clear how the human system senses strand hierarchy. In bacteria, the system detects methylation to identify the old strand.

The mismatch repair system was originally characterized in its simpler bacterial form. In *E. coli*, the core of the system consists of three proteins named *mutS*, *mutL* and *mutH* (as its genome is hyper-*mut*able with a defect in one of these genes). *MutS* recognizes the discontinuity in DNA geometry: after binding to this spot, a *mutS* dimer scans the DNA in both directions in an energy-dependent process, looping out the surveyed intermediate. As soon as one of the *mutS* proteins finds a methylated base, the complex is completed by additional *mutL* and *mutH* dimers. *MutH* is an endonuclease, cutting the strand opposite the methylated base. This solves the strand recognition problem: the cut marks the "new" strand in the vicinity of the mismatch. With the help of additional proteins, the stretch of DNA between cut and mismatch is removed and resynthesized.

The human mismatch repair system uses the same type of components. It is, however, more complex, in that it employs several homologs for *mutS* and *mutL*, while no correlate has been identified for endonuclease *mutH*. Human *mutS* homologs are termed MSH2-MSH6 (MSH1 has been described in yeast only). Four identified *mutL* homologs are designated MLH-1, MLH-2, PMS-1 and PMS-2 (PMS stands for post-meiotic segregation, as the MMR system has additional functions in meiotic crossing over). While the bacterial system works with homodimers, the human system uses heterodimers. A typical complex used to correct misincorporations consists of MSH2/MSH6 with MLH1/PMS1. The human mechanism for

strand identification does not use methylation as a cue. Its exact nature remains unclear, although it seems to rely on nicks such as those between Okazaki fragments.

The molecular machinery responsible for DNA replication and proofreading works with stunning precision, yet cannot be perfect. After all checks and corrections, what remains is an error rate of about 10^{-6} mutations per cell division and gene. While this sounds very small, we subject our genes to a lot of cell divisions over our lifetime: approximately 10^{16}. Cumulatively, in all of our cells over our entire lifetime, this adds up to 10^{10} mutations in each of our genes.

Our cells operate in a constant drizzle of mutations

In their entirety, repair mechanisms remove the vast majority of all instances of DNA damage or replication errors, preventing the fixation of mutations. In summary, while our cells are inundated with DNA damage due to exogenous and "endogenous" causes, cellular safeguards reduce that to a drizzle of mutations. As most of these occur in different cells, this is no problem unless we connect too many divisions in series. Hence, a limit on the number of divisions a somatic cell can undergo is essential to prevent accumulation of too many mutations in a single cell. One way to implement such a limit is telomere shortening (explained in section 4). In its development from stem cell to terminally differentiated cell, a typical somatic cell undergoes only 10 to 15 divisions.

Can we avoid mutations? Can we avoid cancer?

In principle, environmental carcinogens can be avoided, while "endogenous" causes of mutation like hydrolysis, oxidation, replication errors etc. are unavoidable. So, what is the relative contribution of environmental carcinogens, and is it possible to avoid them? The first question cannot be answered precisely, and many different estimates have been published. One middle of the road-group of epidemiologists concludes (Trichopoulos et al., Scientific American 275 (Sept.): 50-57, 1996) that only a quarter of all malignant neoplasms is attributable to "endogenous" causes, while 75% is due to environmental factors. The bulk of these 75% can be attributed to smoking (30%) and unhealthy eating habits (30%). The premier effect of smoking becomes immediately obvious when looking at the development of death rates from cancer from the thirties until today. The tremendous increase in the death rate by lung cancer, first in men, then, with a delay of 30 years, in women, mirrors the development of smoking habits during the last century.

A single mutation is not enough to cause cancer

For most malignant tumors, incidence increases with age. Plotting, e. g., incidence of colorectal carcinoma against age yields a typical, exponentially increasing function. This type of function may be explained as the result of several unrelated, rare events coinciding in a single cell. From that specific function, it is possible to calculate that colorectal carcinoma is the product of six unrelated events. Such an "event" may be either activation of a proto-oncogene or the loss of a tumor suppressor. Coincidence of six such events in a single cell is extremely unlikely at the beginning of life, but its probability increases continuously with age due to accumulation of mutations, especially in stem cell lineages. The basis for this complex requirement can be found

in the layers of checks and control mechanisms governing proliferation and behavior of cells, each of which has to be altered separately by genetic or epigenetic changes.

There are exceptions to this rule. We will soon discuss one of these exceptions, retinoblastoma, a tumor of the retina, which develops almost exclusively in infants.

The multi-step model: mutation and selection

The model currently guiding our thinking about carcinogenesis was developed by Vogelstein and Kinzler based on observations in colorectal carcinoma. It views carcinogenesis as a multi-year process of "microevolution" based on steps of mutation and selection. In the development of colorectal carcinoma, the first event may inactivate the tumor suppressor APC (*adenomatous polyposis coli*) in a stem cell in a colonic crypt. The phenotype of this cell changes only minimally. The cell may divide somewhat more frequently, and there may be a problem with asymmetric stem cell division, potentially increasing the number of stem cells in the crypt (see section 12 on colorectal carcinoma for more details). Over time, maybe two years, a nest of APC-less cells with stem cell properties develops. They might even crowd other stem cells out of their cryptal stem cell niche. As we have seen before, mutations occur in proliferating tissue with a certain frequency, due to environmental and to "endogenous" causes. The more APC-less stem cells, the higher the probability that an additional mutation hits one of these cells. This second mutational event might be the activation of Ras by a point mutation. The resulting cell has an additional selective advantage and starts to form an "island" of offspring within the continuous epithelium. A further 18 months later, a third hit may occur in one of these cells, e. g., inactivation of the tumor suppressor p53. An island of daughter cells with three problems grows within the previous island of two-problem cells. The mass of cells, too big to remain within the epithelial plane, starts to form an adenoma growing into the lumen of the colon. Two years later, a fourth mutation occurs in one of the adenoma cells. And so forth, and so on…

Each mutational event increases the selective advantage of the respective cell clone, and releases inhibitions to show "bad manners". After maybe ten years, a sixth independent mutational or epigenetic event generates the first malignant cell.

"Tumor-evolution" does not stop at this point. During further developments, the tumor keeps accumulating additional mutations. This may influence its tendency to metastasize, the route of metastasis, or the tumor's response to therapy. The tumor thus splits up into genetically diverse subclones which may show different biological properties.

Because of genetic heterogeneity, the same type of carcinoma can show different properties in different patients

Cancer patients with the same type of malignancy, grading and staging are known to show vast differences in their course of disease. While one patient with colon cancer reacts well to therapy, living free of complaints for many years, another quickly suffers a relapse, passing away shortly after. At first glance this looks incomprehensible. Yet, as discussed above, the biological properties of the neoplasms are determined by specific mutations, and the tumors of the two patients are caused by different sets of mutations.

In some instances, the presence of mutations in certain genes helps to predict whether specific therapeutics (e.g., monoclonal antibodies against the EGF receptor) are likely to help the patient. One of the present developments in cancer therapy is to diagnose causative mutations in individual tumors, and to use this information to apply rational, "custom-made" therapy to the patient (personalized medicine).

Beyond mutations: tumor promoters

Not all carcinogens are mutagens. This became clear in early experiments where two types of cancer-promoting agents were identified, tumor initiators and tumor promoters. After treating the skin of a mouse first with a substance from the tumor initiator family, and then with one from the tumor promoter family, the mouse developed a carcinoma. In reverse –first promoter, then initiator— nothing happened.

Tumor initiators, it was found, are mutagens. Tumor promoters, in contrast, drive proliferation without causing mutations themselves. Applying only a tumor promoter, even repeatedly, does not cause progression to malignancy. Only if cells are first seeded with mutations by a tumor initiator, the tumor promoter exerts its negative effect by multiplying these damaged cells, their mutations included. This increases the probability that additional mutations will hit cells that already carry genetic defects. In summary, tumor promotors increase the number of cells with pre-existing mutations without being mutagenic themselves.

Examples for tumor promoters are estrogens with respect to mammary carcinoma, androgens for prostate cancer. The proliferative stimulus by a chronic ulcer may have a tumor promoter effect in the development of gastric cancer, and regenerative proliferation in chronic viral hepatitis may promote development of hepatocellular carcinoma.

3. TUMOR SUPPRESSORS

3. 1. REGULATION OF THE CELL DIVISION CYCLE

Phases of the cell division cycle

In bacteria, replication of the circular genome, starting from a single origin of replication, and cell division happen at the same time, avoiding a lot of problems. But during evolution, genomes increased in size, eventually becoming too big to be handled in the form of a singular, circular macromolecule (*E. coli* genome: 4 million base pairs, human genome: 3 billion base pairs). So the DNA was chopped into segments. DNA replication had to be started simultaneously from multiple origins and had to be moved ahead of mitosis. A bunch of new tools and control mechanisms had to be introduced to make sure that each piece of DNA gets copied, but only once, and that each daughter cell later is assigned exactly one copy of each segment of replicated DNA.

Consequently, the phases of the eukaryotic cell division cycle are:
G1- originally for gap, more optimistically for growth

S- DNA replication (synthesis)
G2- between completion of DNA replication and mitosis
M- for mitosis

"G1" is only used for actively proliferating cells. As an alternative, G_0 denotes a prolonged resting state ("quiescence"). Cells in G_0 are idle, ready to reenter the cell cycle in case growth signals come in. In the meantime, they reduce their metabolic rate, as there is "nothing to do right now anyway".

Cyclin-CDK complexes

Once all requirements for the next step in the cell division cycle are met, a "master switch" in the form of a cyclin-CDK complex is activated.

The simplest form of the control system directing all those steps in eukaryotes can be found in bakers' yeast, *Saccharomyces cerevisiae*. This organism has a single master switch, a kinase, which is always present. Yet, it remains inactive unless a second protein, termed cyclin, is expressed, which binds to it. The kinase was named according to this mechanism: cyclin-dependent kinase or CDK. The CDK can be activated by different cyclins at different points in time. A G1/S cyclin is expressed for the general decision to activate the division process. An S-cyclin is necessary to actually start DNA replication. A mitotic cyclin is expressed to induce mitosis proper. The cyclin-CDK complex is active as long one of these cyclins is present. Once the respective cell cycle step is completed, the cyclin is quickly broken down by controlled proteolysis, which inactivates the CDK. Individual cyclin-CDK complexes differ in substrate specificity. Phosphorylated target proteins translate the master switch's decision into action.

The human cell cycle

The human cell cycle control system functions along the same lines. While the system in yeast uses a single CDK, several related CDKs share the work in humans. The following cyclin-CDK complexes are important for cell cycle progression (CDKs are denoted by numbers, cyclins by letters):

G1 phase:	cyclinD-CDK4 –phosphorylates pRb
	cyclinD-CDK6 –phosphorylates pRb
G1-S transition:	cyclinE-CDK2 –phosphorylates pRb
S phase:	cyclinA-CDK2
G2-M transition:	cyclinA-CDK1
M phase:	cyclinB-CDK1

Levels of CDK regulation, checkpoints

Information feeds into the CDK master switches on at least three levels:
- cyclin presence/absence
- activating and inhibiting phosphorylations
- protein regulators binding to the complex or to CDK alone

Examples for protein regulators are p21Cip1, p27Kip1 or p16INK4A. A CDK is thus an integrator of signals that is active only after a set of independent conditions are met. If that is not yet the case, these master switches are able to arrest the cell cycle at so-called checkpoints. Factors like cell mass, nutritional supply or cell energy levels feed into the control system, but of special importance is the integrity of the genome: potential DNA damage and the attachment of all DNA segments (chromatids) at the mitotic spindle. The following checkpoints have been defined:

- *G1 DNA damage checkpoint* (arrest in case of DNA damage)
- *DNA replication checkpoint* (arrest in case of stuck replication forks)
- *G2/M DNA damage checkpoint* (arrest in case of DNA damage following replication)
- *Spindle checkpoint* (arrest until all chromatids are attached and under tension)

Cell cycle arrest in these situations prevents problems from spiraling out of control, like replication of damaged DNA, distribution of incompletely replicated DNA or misallocation of chromatids (aneuploidy). Cell cycle arrest creates a time window to correct the underlying problem: damage can be repaired, replication can be completed, the last chromatids can be correctly attached. In case a problem cannot be overcome within a reasonable time frame, the checkpoint mechanism usually induces apoptosis. This protects the organism from the danger of cells harboring defective versions of the genome.

To implement a checkpoint, an array of proteins is required. If these proteins are affected by mutations themselves, the cell cycle cannot be arrested in response to a problem. Ongoing DNA replication or mitosis in the face of such problems leads to further mutations or aneuploidy.

We will take a closer look at the G1 DNA damage checkpoint in the section on p53.

3. 2. RETINOBLASTOMA PROTEIN (pRb)

The protein pRb, named for the retinal tumor retinoblastoma, is the central substrate of cyclin-CDK complexes active in G1. As long as it is not phosphorylated, pRb masks and inactivates transcription factor E2F at the promoters of its target genes (there are several types of E2F, which heterodimerize with dimerization partner DP). The more growth signals come in, the more steps are completed to activate CDK4 and CDK6: cyclin D is expressed, along with small amounts of cyclin E. Phosphorylation of CDKs and protein regulators are adjusted. Cyclin D-CDK4 and -6 phosphorylate pRb at several sites. After a finishing touch via phosphorylation by cyclin E-CDK2, pRb undergoes dramatic conformational change, falling off E2F. E2F is now free to activate its target genes:

- E2F, its own gene, for a positive feedback amplification loop
- cyclin E (enhancing positive feedback)
- CDK2 (enhancing positive feedback)
- c-Myc, c-Myb (transcription factors inducing additional groups of genes)
- cyclin A (cyclin A-CDK2 is necessary to fire origins of replication)
- DNA polymerase α
- PCNA (sliding ring tethering DNA polymerase to DNA)
- thymidylate synthase, thymidine kinase (for thymidine synthesis)

- dihydrofolate reductase (for *de novo* synthesis of purines and thymidine)
- CDK1 (required for subsequent mitosis)
- APAF1 (for emergency exit via apoptosis in case of problems)
- caspase 3 (for emergency exit via apoptosis in case of problems)
- p14ARF (sensor for non-physiological cell cycle activation, see section on p53!)

Proteins encoded by these target genes have functions essential for S or M phases, e. g., *de novo* synthesis of deoxynucleotide triphosphates (dNTPs) required for DNA replication. In summary, phosphorylation of pRb by G1 cyclin-CDK complexes results in activation of all the preparatory steps required for cell division. At the same time, factors required for apoptosis are supplied as inactive precursors, allowing the cell to take that path in case it runs into serious problems. This has to be done before chromosomes start to condense, as gene expression is shut down during mitosis.

pRb as the paradigm of a tumor suppressor

PRb is the poster child of a tumor suppressor. In the default situation, it masks and inhibits E2F, thereby arresting cells in G1. Thus, pRb acts as a parking brake for all cells. Growth signals act to release this brake via expression of cyclins D and E, enabling CDKs to change pRb conformation by phosphorylation. Many tumor suppressors may be understood as brakes acting on the cell cycle.

Rb gene mutations frequently cause a loss of braking function

Physiologically, the pRb brake can be released only by growth signals. What if the Rb gene is affected by mutations? Deletions, premature stop codons, reading frame shifts and many point mutations interfere with the ability of pRb to mask E2F. In all these cases, mutations remove the parking brake, again feigning the presence of a growth signal.

Typically, tumor suppressors are dual-circuit brake systems: the real problem arises once both of the two alleles fail

Defects typically result in something NOT functioning. Applying this truism to mutations, this means that the large majority of them are loss-of-function mutations. An affected allele is inactivated. However, the human diploid genome contains two copies for each gene, and the braking function of pRb derived from the second allele is still sufficient to keep the cell in G1. In other words, loss of Rb is recessive. Only after the second allele is inactivated, the cell enters the division cycle.

This principle was first recognized by Alfred G. Knudson, who observed children with retinoblastoma, a tumor of the retina. Retinoblastoma is unusual in several respects: it affects almost exclusively infants; it frequently affects both eyes, and it is sometimes hereditary. Retrospectively analyzing case histories of children with retinoblastoma, Knudson identified two groups: the first group had more than one primary tumor (three independent tumors, on average, usually affecting both eyes), was younger at diagnosis (14 months), frequently developed secondary tumors later in life (e. g., osteosarcoma) and in this group, retinoblastoma was occasionally familial. The second group had a single primary tumor in one of the eyes and was older at diagnosis (30 months) with no indication of secondary tumors or hereditary factors.

Knudson interpreted these data as follows: the first group, he reasoned, started their lives with a defective, at that time hypothetical, "Retinoblastoma" allele. All somatic cells of the developing child harbor the defective copy. Over time, in the constant "drizzle of mutations", some cells lose the second allele, originating a primary retinoblastoma tumor. On average, this happens three times per child. Why just in the retina? Some of the checks and controls implemented in human cells do not seem to apply in retinal cells. For the development of this malignant tumor, loss of the tumor suppressor pRb in the environment of an infant's growing eye ball seems to be sufficient. What about other tissues? Obviously, safeguards are better there, requiring mutations in additional loci. This explains the time lag until secondary tumors are seen. Why are parents usually unaffected? In most cases, the first hit seems to occur in parents' germ cells; the parents themselves are healthy. In the second group of children, retinoblastoma is due to two sequential independent mutation events in the same line of retinal cells. This takes more time, so that the children are older. Once the retinoblastoma has been removed, there is no increased risk for other tumors. In both groups, the two alles of the Rb gene had to be inactivated by two independent hits. By his "two-hit-hypothesis", Knudson defined the concept of a tumor suppressor: a gene that contributes to tumor formation by inactivation of both of its alleles.

In this regard, the braking systems of our cells resemble those of our cars. Hitting the brake pedal leads to compression of brake fluid in the master brake cylinder. Via a hydraulic system, the pressure is transmitted to the individual wheel brake cylinders, where by moving the piston, brake pads are pressed against the brake disk. Yet, hydraulic systems are prone to leaks: fluid escapes, air enters the system. With easily compressible air in the system, the brake pedal goes to the floor, and pressure is not transmitted to the wheels any more. In the days of single circuit braking systems, this caused many accidents. So, the automotive industry developed dual-circuit braking systems, with two independent hydraulic systems. If one of them leaks, it is still possible to brake via the second.

Tumor suppressor genes are sometimes called antioncogenes. A third term is "recessive oncogenes", as the phenotype is only observed after both alleles have been lost. Many antioncogenes have been identified by a phenomenon termed *loss of heterozygosity*. After a first "hit" inactivating the first allele, the cell is heterozygous for the respective locus. In children with hereditary retinoblastoma, all somatic cells are heterozygous. What is frequently observed in tumors, is that the second, normal allele has disappeared. There are several ways for that to happen: a large deletion eliminating the second allele altogether; gene conversion ("repair" of the normal allele using the defective allele as template); somatic uniparental disomy (loss of the normal chromosome compensated by a doubling of the defective one). Usual diagnostic procedures, e. g., PCR followed by sequencing, cannot differentiate between these possibilities: they "see" one defective type of locus, be that a single allele, or two identical defective alleles. This is the technical explanation for the somewhat strange term "loss of heterozygosity".

Epigenetic inactivation of tumor suppressors

Any one cell expresses only a small slice of the human genome's 23,000 protein-encoding genes. In any tissue, the majority of genes are silenced permanently by epigenetic changes resulting in heterochromatin formation. The process starts by CpG methylation, the methylation of atom C5 in those cytosines that are followed by guanines (with an equivalent configuration

at the opposite strand). This is followed by histone deacetylation and tight packaging of the respective locus with additional specialized chromatin proteins. Sometimes, this process goes awry, shutting down genes that should have remained active. If that happens to an antioncogene, the result is functionally equivalent to a loss by deletion. The factors determining erroneous shutdown are insufficiently understood, but specific tumor suppressors are known to be affected frequently. These include cyclin-CDK regulators p27 Kip1, p16 INK4a and p15 INK4b (see below), breast cancer tumor suppressor BRCA1 and mismatch repair system component MLH1. The epigenetically inactivated tumor suppressor is passed on to successive cell generations like an inheritable trait. While the locus has to be opened for DNA replication, DNA maintenance methylase quickly methylates the newly synthesized strand, renewing the gene's inactivated state.

Tumor suppressors functionally related to pRb: inhibitors of cyclin-CDK complexes (CKIs)

As cyclin-CDK complexes are required for cell cycle progression, inhibitors of these complexes act as brakes on the cell cycle. There are two distinct groups:

1. Group: Members bind to cyclin-CDK complexes. They help assemble cyclin D-CDK4/6 complexes, but inhibit them, probably depending on the phosphorylation state of the inhibitor; complexes containing CDK2 or CDK1 are inhibited:
-p21 CIP1/WAF1
-p27 Kip1
-p57 Kip2

2. Group: Members bind to CDK4 and CDK6 alone, preventing their interaction with cyclin D:
-p16 INK4a
-p15 INK4b
-p18 INK4c
-p19 INK4d

3. 3. CENTRAL TUMOR SUPPRESSOR p53

Its central position in carcinogenesis derives from the fact that p53 seems to be the single most frequently mutated gene in human tumors. P53 knockout mice are born healthy, at normal Mendelian frequency. Thus, p53 is not required for development of a normal mouse. However, these mice later develop malignant tumors at a relatively young age.

One might compare p53 to the fire department: if we abolish it today, that does not necessarily imply immediate catastrophe. Yet, we can be virtually assured of a disaster sometime in the future. P53 is a transcription factor enabling the cell to respond to specific damaging conditions with the goal to minimize negative outcome.

Structure and degradation of p53

P53 has three domains: a N-terminal degradation/activation domain, a central DNA binding domain and a C-terminal tetramerization domain. P53 is expressed in virtually all cells. As long as the cells do not encounter any problem, p53 is degraded promptly via the ubiquitin-proteasome pathway. As a consequence, healthy cells contain hardly any p53.

In the ubiquitin system, three enzymes cooperate: E1, E2 and E3. E1 is ubiquitin-activating enzyme: in an energy-dependent process, it binds the small, ubiquitously expressed protein ubiquitin to an –SH group. It then hands the activated ubiquitin to one of several ubiquitin conjugating enzymes, termed E2. E2 in turn attaches to one of many E3 proteins. Usually, E3 proteins are called ubiquitin ligases, which is not entirely correct, as they don't have any enzyme activity themselves. Only the complex of E2 plus E3 enables transfer of ubiquitin to the substrate; E2 does the transfer; E3 does the binding to a regulable degradation domain of the substrate. In the case of p53, this is the N-terminal domain.

The main E3 protein in charge of p53 is Mdm2 (*mouse double minute 2*, a historical designation frequently substituted by Hdm2 *–Human double minute-* on the grounds that humans cannot have a mouse gene). As long as it not phosphorylated, the N-terminal degradation/activation domain of p53 binds efficiently to Mdm2, which, together with its E2 partner, ubiquitinates p53 several times over. Polyubiquitination is the signal required for proteasomal degradation. If the N-terminal domain of p53 is phosphorylated, binding to Mdm2 is impaired, resulting in much slower degradation.

Point mutations frequently found in human carcinomas predominantly affect a select few codons of the DNA binding domain, most of them arginines (e. g., Arg248, Arg249, Arg273). These amino acids seem to be of special importance for the function of p53, or, in other words, substitution of these amino acids particularly favors carcinogenesis. By X-ray structure analysis, these amino acids were found to either make direct contact with DNA, or to have stabilizing functions in the immediate vicinity of the p53-DNA interface. Substitution of these amino acids interferes with DNA binding, preventing or impairing the function of p53 as a transcription factor.

Activation of p53

Defined conditions of cellular stress cause activation of p53, including DNA damage, hypoxia and non-physiological forms of cell cycle activation.

The textbook example is a DNA double strand break. The break induces chromatin changes in its vicinity. Histone H2A is replaced by H2AX in nearby nucleosomes, and heterodimers of proteins termed "Ku" mark the ends of DNA strands. These changes recruit several protein kinases, including DNA-dependent protein kinase, ATM and the checkpoint kinase Chk2, which proceed to phosphorylate p53, H2AX and a range of additional proteins.

[ATM stands for Ataxia Teleangiectasia Mutated. A defect in the gene encoding this kinase causes a syndrome combining extreme sensitivity to radiation with a predisposition for lymphoid malignancies as well as cellular and humoral immunodeficiency]

Phosphorylation of p53 has two effects:

- Less efficient binding to its E3 ubiquitin ligase Mdm2. As this results in reduced degradation (or increased half-life), cellular p53 levels rise strongly.
- A switch of the N-terminal degradation domain into a transactivation domain. The negatively charged phosphates facilitate recruitment of RNA polymerase.

Target genes of p53

Here, we consider only a few out of a large number of p53-induced genes:

- Mdm2: At first glance, it looks strange that p53 induces its own E3 ubiquitin ligase. Yet, keep in mind that the goal of p53 is to overcome a problem. Once that's been achieved, accumulated p53 has to be eliminated. Viewed from this perspective, it is useful to supply considerable amounts of Mdm2 in advance. As long as p53 is phosphorylated, ubiquitination by Mdm2 will remain inefficient, anyway.
- p21 CIP1/WAF1: inhibitor of cyclin D-Cdk4 and cyclin E-CDK2, causing cell cycle arrest in the G1 phase
- 14-3-3σ as a tool to mediate G2 arrest
- several genes involved in DNA repair
- Bax and related apoptosis-promoting genes of the Bcl2 family, like PUMA and Noxa

P53 activation results in cell cycle arrest with repair, or in apoptosis

Let's continue with our example of a DNA double strand break: a single break in a cell is sufficient to activate p53. Phosphorylated p53 activates its target gene p21, which in turn inhibits the cyclin-CDK complexes required to phosphorylate pRb. Transcription factor E2F remains masked, arresting the cell in G1. This gives the cell an opportunity to repair its DNA, preventing the propagation of incomplete chromosomes. However, the problem that caused p53 activation cannot always be solved. A p53-induced G1 arrest dragging on for weeks or longer is called senescence. Under certain circumstances, senescent cells are eliminated by macrophages.In summary, p53 is a crucial tool of the G1 DNA damage checkpoint. This mechanism works well as long as p21 and pRb are functional. What if one of these tumor suppressors gets inactivated?

Let's consider this second type of emergency, assuming the loss of both Rb alleles by mutations. In the absence of pRb, E2F is constitutively active. Obviously, E2F is activated with each cell division, but this physiological activation is only temporary. In the absence of pRb, in contrast, E2F remains active indefinitely, forcing its transcriptional targets, including p14ARF, higher and higher. P14ARF binds and inactivates Mdm2, the E3 ubiquitin ligase for p53. Without Mdm2, intracellular levels of p53 increase progressively, activating p53 target genes. In this case, induction of p21 remains without effect (inhibition of cyclin-CDK complexes is useless in the absence of pRb), but apoptosis-promoting target genes Bax, PUMA, Noxa are expressed more and more, finally pushing the cell over the cliff. From the organism's perspective, this is a good solution. The single apoptotic cell is easily replaced. In return, this eliminates the risk that a cell with defective cell cycle control develops into a malignant neoplasm.

The name p14ARF stands for *alternative reading frame*. It is encoded by an odd gene on chromosome 9: in this small gene, the main exon 2 is translated in two different reading frames,

giving rise to two entirely different proteins. One of these proteins is p16 INK4A, the other p14ARF. Both of them are crucial tumor suppressors. If the common exon is affected by a deletion or mutation, the cell loses two tumor suppressors by a single event. The locus on 9p21 is thus sort of an Achilles' heel. In many tumors, it is in fact mutated.

Consequences of mutations in p53

Individual mutations of p53 differ with respect to their biological consequences and may affect at least four aspects to varying degrees: mutations may impair DNA binding, impair proper protein folding, cause mislocalization in the cytoplasm and interfere with tetramer function.

As mentioned before, frequent point mutations near the surface of the p53 molecule cause the protein to lose its DNA binding activity. Some specific, less frequent substitutions within the dense core of the DNA binding domain have an additional effect: they cause the DNA binding domain to "pop" (like corn to popcorn) by massive conformational change. Frequently, this has two consequences: an increase in half-life of the now difficult-to-degrade protein, as well as mislocalization in the cytoplasm (probably, nuclear localization signals are buried in the protein blob).

The tetramerization domain remains unaffected by mutations of the DNA binding domain. Mutated p53 monomers, even the products of "popping" alleles, continue to participate in forming tetramers. Tetramers are assembled in all permutations from both mutated and wild-type monomers, the latter stemming from the remaining normal allele. Exacerbating the problem is the fact that popped versions frequently have a longer half-life. How many mutant subunits a tetramer can tolerate depends on individual mutations. For many mutations, especially the popping type, already a single mutant subunit is enough to inactivate the tetramer. The only tetramers that work, those made up of four healthy subunits, become exceedingly rare (one sixteenth of all tetramers if mutant and wild-type have the same stability and even less if the mutant is degraded more slowly). Yet, only those are able to do the job. Hence, this type of single allele-mutation in essence kills the cell's p53. For some mutations affecting only the DNA binding interface, like the frequent Arg273His mutation, this effect is less severe.

Thus, many, but not all mutations inactivate most of the p53 in that cell. This makes p53 an exception. While mutations in antioncogenes are typically recessive, many mutations in p53 cause the mutated allele to dominate over the remaining normal allele. We call that a *dominant negative* effect.

If a cell's p53 has been lost, double strand breaks have deleterious consequences. Deprived of the ability to arrest in G1, the cell does not succeed in repairing the break immediately. The broken chromatid is replicated in parts. The peripheral chromatid fragment, the one not attached to a centromer, is likely to be lost during the next mitosis. Following replication, "illegal" DNA ends ("legal" ends are indicated by telomers) are "repaired" by end-to-end ligation of sister chromatids. This "solution" to the problem is an illusion: in the following mitosis, a tug-of-war begins with three possible outcomes: 1) daughter A wins both fused chromatids, daughter B gets nothing 2) daughter B wins both chromatids or 3) the DNA breaks once again. In the latter case, the entire process, termed breakage-fusion-bridge cycle, starts anew. From one cell generation to the next, genetic imbalances accumulate, over time severely mutilating the

genome with massive aneuploidy and scrambled chromosomes. In summary, p53 is the "warden of the genome's integrity". Once it is lost, the cell's genome deteriorates quickly.

Li-Fraumeni syndrome

If a mutated p53 allele is passed on in a family, members will develop tumors at an early age. A typical constellation: a young person suffering from sarcoma, e. g., osteosarcoma, with two direct relatives who also developed tumors at an early age, e. g., carcinoma of the adrenal cortex, mammary carcinoma or a brain tumor.

The status of p53 may affect a tumor's response to therapy

P53 is of particular importance in tumor therapy. Many therapeutic options, e. g., radiation or components of chemotherapy, work by causing DNA damage. These measures seem to be much more efficient if the treated tumor's p53 is still functional. Massive DNA damage causes p53 to induce apoptosis in these cells. Malignancies that tend to have intact p53, like seminoma, Wilms' tumor or ALL in children, generally respond to therapy much better.

Virus proteins inactivating pRb and p53

Several DNA viruses independently developed proteins to inactivate the two central tumor suppressors pRb and p53:

virus	pRb	p53
adenovirus	E1A	E1B-p55
human papilloma virus	E7	E6
SV40 (simian virus)	T	T ("large T")

DNA viruses rely on contributions from the cell, like deoxyribonucleotides, to replicate. While they cannot replicate in a cell in G1, they obviously gain a selective advantage if they find a way to push the cell into S phase. This can be done by inactivating pRb, but cells would then enter apoptosis via p14ARF and p53. Therefore, the virus only wins if it also succeeds in inactivating p53.

3. 4. FURTHER TUMOR SUPPRESSORS

Other tumor suppressors are covered in the sections on colon carcinoma and breast cancer:

Colorectal carcinoma:
APC
"HNPCC"

Mammary carcinoma:
BRCA1
BRCA2

The state of a cell's DNA repair systems is of paramount importance for carcinogenesis. Cells use several repair systems to prevent ubiquitous DNA damage from being transformed into fixed mutations. P53 or ATM may be classified under the heading "repair". Many proteins required for DNA repair show properties of tumor suppressors.

4. TELOMERES AND TELOMERASE

Is this a legal end?

From a biological perspective, DNA is an "endless" macromolecule; ends are the result of double strand breaks. Fittingly, the genome of bacteria is circular. In eukaryotes, where the genome is too big for that arrangement, genome segments can't avoid having ends. To mark these ends as permissible ends, they have a special structure based on hundreds of repeats of the short sequence TTAGGG. The end of a chromosome forms a loop. After looping back, one of the two strands, the "G-rich strand" which forms a large overhang, displaces its anteceding self (producing a displacement loop) by specific base pairing. This generates kind of a "lasso structure" which is stabilized by specific proteins, in effect hiding the DNA end from the cell. In its entirety, this structure is termed telomere. Telomere structures mark "legal" ends; all other ends are interpreted as double strand breaks that need to be repaired.

The end-replication problem

Chromosomal ends cause problems in replication. The strand synthesized in the direction towards the end is no problem, as it can be synthesized continuously. The direction from the end is synthesized discontinuously in the form of Okazaki fragments. First, a small stretch of RNA is synthesized, as only RNA polymerase can start from scratch. The RNA then functions as primer for DNA polymerase which produces the Okazaki fragment. Following synthesis of the next fragment, the RNA primer is removed and replaced by DNA. This works for all RNA primers except for the last one, the one at the very end of the chromosome. This can be removed, but it cannot be replaced, as there is no "next" Okazaki fragment. The single-stranded remainder is not stable and is eventually degraded. With each round of DNA replication, the telomere is shortened by 50 to 100 base pairs.

Telomerase

Over successive generations, this process would lead to disaster. At some point in time, the lasso structure is lost, and still later, important genetic information might go. To prevent that, evolution developed telomerase, an enzyme complex that helps circumvent the problem. Telomerase uses a loop of built-in RNA as a template to elongate the chromosome's 3'end, attaching the sequence TTAGGG many times over. By that, it creates a template of discretionary length for additional Okazaki fragments, thereby overcoming the end replication problem. In essence, telomerase has reverse transcriptase activity, as it uses a RNA template to

create DNA. Therefore, its main catalytic unit is designated hTERT (human telomerase reverse transcriptase).

Replicative senescence

Telomerase is active only in germline cells –the only cells handed down indefinitely--, in clonally expanding B- and T-cells and weakly in some stem cells. In all other somatic cells, telomerase is not expressed any more. This helps to limit their proliferative potential and contributes to cell senescence. As soon as, following a number of division-related shortenings, a cell's first telomer is too short to permit the lasso structure, the cell gets aware of the DNA end, interpreting it as a double strand break. Via activation of p53 and p21, the cell enters a sustained cell cycle arrest, referred to as replicative senescence. This barrier limiting the proliferative potential of somatic cells is an important mechanism to protect us from cancer.

Cancer cells reactivate telomerase

A cell losing p53 loses this protective mechanism, too. Lacking their molecular replication countdown clock, these cells keep dividing, spending all remaining telomere capital. The process ends in chromosomal instability (CIN), with chromatid end-to-end joining and breakage-fusion–bridge cycles. Most cells enter a crises, where they lose essential genes and die. Over time, a few cells manage to reactivate their hTERT gene. Telomerase now stabilizes the present, more or less catastrophic state of these cells' genome by adding telomeres to whatever DNA ends are around, camouflaging them as "legal" ends. This stops the vicious breakage-fusion–bridge cycle and leaves a viable cell with a grotesquely mutilated genome that forms the basis of the next phase of tumor expansion. In 85 to 90% of human tumors, telomerase is reactivated (the rest have acquired an alternative pathway for telomere maintenance that we will not go into here). Reactivated hTERT thus acts as an oncogene, contributing to the unlimited proliferative potential of tumor cells.

Pharmaceutical industry is working hard to develop telomerase inhibitors. As the vast majority of somatic cells do not express telomerase, there is hope for a novel class of anti-cancer therapeutics with limited toxicity.

5. PROTO/ONCOGENE ACTIVATION

The path of discovery resulted in two or more names for one protein

Confusingly, many of the classical proto-oncogenes/oncogenes have two or more names. Many of them were first discovered in their oncogenic form in retroviruses causing tumors in animals. Only later were they recognized to have physiological counterparts in our body. Consider the two oncogenes of avian **er**ythro**b**lastosis virus (AEV), *erbA* and *erbB*. ErbA later was recognized as an oncogenic form of the thyroid hormone receptor, ErbB as that of the epidermal growth factor (EGF) receptor.

Some proteins, like Myc, have kept their viral designation. In cases where it might be unclear which form is meant, the physiological form is indicated by the prefix c- (for cellular), the viral

one with v- (e. g., c-Myc, v-Myc). Traditionally, in literature dealing with molecular aspects of cancer, the gene is written *c-myc* (lower case and italics), the protein, c-Myc. Today, this is colliding with a newer convention, according to which human genes should be capitalized throughout: c-MYC (instead of *c-myc*).

So, c-ErbB sounds very different from EGFR or HER-1 (human EGF receptor-1), but in fact they are all the same thing.

The EGF signal transduction pathway harbors typical proto-oncogene jobs

The EGF/PDGF signal transduction pathway contains several classical proto-oncogenes. The pathway is activated in many tumors and is a target for intervention, as we have several drugs to block it. Epidermal growth factor is a misnomer: almost all epithelial cells express EGF receptors. Many mesenchymal cells express PDGF receptors. From the specific receptor onward, both pathways are identical. **EGF** crosslinks two **EGF receptor** proteins, which react by mutual phosphorylation by their intracellular tyrosine kinase domains. The phospho-tyrosines bind to an **adaptor**, which then recruits a guanine exchange factor (GEF) to the membrane. GEF interacts with the **Ras** protein, exchanging GDP for a GTP. GTP-Ras activates several kinase cascades: one comprises **Raf**, MEK and ERK, the other MEKK, SEK and Jun-K (never mind the full names). In each case, the last of these kinases transfer to the nucleus, phosphorylating transcription factors **Elk** and **Jun**. The transcription factors in turn activate genes encoding early growth factor response transcription factors: **Fos** and Jun. Via this short-term positive feedback loop, activated Jun produces more Jun (the heterodimer Fos/Jun is called AP-1, activating protein-1). In combination, Elk, Fos and Jun activate a host of additional genes required for implementation of the growth signal.

All proteins printed in **bold** in the preceding paragraph have occasionally been found to contribute to tumor formation in their oncogenic form. If activated by mutations, these oncogenes create the illusion of lots of EGF around the cell, and the cell proceeds to proliferate.

How proto-oncogenes are activated

Class I: Growth factors

Growth factors may act oncogenic if their expression is deregulated; usually, the signal protein itself is not altered. In many cases, the malignant cell expresses both growth factor and receptor: an autocrine feedback loop.

The oncogene of a retrovirus causing sarcomas in monkeys and apes, *sis*, conforms to platelet derived growth factor (PDGF). PDGF is expressed by many epithelial cells, with neighboring stroma cells expressing the PDGF receptor. In humans, deregulated expression of PDGF contributes to a small subgroup of Ewing sarcomas. Ewing sarcomas are made up of small, round, undifferentiated cells, most commonly in the bone of teenagers. Of the different translocation events involving the EWS gene, one results in a chimaeric transcription factor that leads to deregulated expression of PDGF.

In Hodgkin's disease, overproduction of growth factors by relatively low numbers of Hodgkin and Reed-Sternberg cells causes secondary accumulation of normal lymphocytes, histiocytes

and granulocytes, over time inducing huge lymphomas. The histologic type of the disease is determined by the individual cocktail of growth factors produced (e. g., nodular sclerosing if the mix contains an unhealthy dose of fibroblast stimulator TGFβ). More than 99% of lymphoma cells in Hodgkin's disease are normal, non-malignant cells just doing what they have been told to do.

Many epithelial cell express one or more of the four types of human EGF receptor (HER1-HER4). These receptors may be activated by a range of growth factors from the EGF family, including EGF, TGFα, heregulin and several others. Airway epithelia express heregulin and all of the four receptor types. Normally, heregulin and its receptors remain segregated: receptors are sorted to the basolateral membrane, while heregulin is secreted via the apical membrane towards the lumen. As long as the epithelium is intact, tight junctions between the cells prevent heregulin from meeting its receptors. This changes as soon as a small tear or scratch produces a discontinuity: via this hole in the epithelium, heregulin diffuses to the basolateral side, stimulating cell proliferation until the epithelial sheet is resealed. This highly useful mechanism proves counterproductive, however, following development of bronchial carcinoma. As soon as cells detach from the epithelium, heregulin is able to contact its receptors, establishing an autocrine feedback loop stimulating proliferation. Drugs like cetuximab (Erbitux®), panitumumab (Vectibix®, both of them monoclonal antibodies against HER1) or erlotinib (Tarceva®, an inhibitor of the EGF receptor's tyrosine kinase) are used in non-small cell bronchial carcinoma to counter this type of autostimulation, as well as stimulation by EGF produced by stroma cells.

Kinases of proto-oncogene classes II and III

As a group, kinases may be activated by mutations with relative ease. The main reason for that is that many kinase domains are active as long as they are on their own; additional parts of the protein are required to moderate their activity. The signal transducer c-Src is a good example to illustrate this principle. C-Src is switched off by folding the protein up like a clasp-knife, deforming the kinase domain. This happens via phosphorylation of a C-terminal tyrosine by another kinase. The phospho-tyrosine is then able to bind tightly to the SH2 (Src homology 2) domain at the N-terminus of c-Src, folding up the protein. The kinase can be reactivated by a phosphatase that removes the C-terminal phosphate. With that, it is easy to imagine mutations activating c-Src: deleting the C-terminus is the easiest way, but point mutations affecting either the tyrosine or the SH2 domain would do the trick just as well. Generally speaking, kinases are frequently activated by removal/inactivation of inhibiting protein parts.

[Remember, gain-of-function mutations are rare; loss-of-function mutations are the rule. So, why is activation of kinase oncogenes quite common? Because the gain-of-function mutation of a kinase is actually a loss-of-function mutation of an integrated inhibitor!]

Although less obvious, a similar situation applies to the EGF receptor. Its extracellular domain normally prevents continuous contact between two receptor molecules. Physiologically, this inhibition is overcome by EGF, which forces the two receptors together. Two types of mutations are known to have a similar effect: deletion of the entire extracellular domain, or a point mutation at the cytoplasmic end of the transmembrane helix (in HER2/Neu). In both cases, these mutations create the "illusion" of the presence of lots of EGF or related growth factors.

Along the same lines, many kinases may be activated to oncogenes. Here, let us just mention stem cell factor receptor c-Kit, c-Raf and c-Abl. In humans, c-Abl is most frequently mutated by the Philadelphia translocation, which we will scrutinize a few paragraphs downstream.

EGF receptor HER2 is known to contribute to carcinogenesis by deregulated expression, without any change in the structure of the protein. This happens in about a quarter of all cases of breast cancer. If the membrane is packed with receptors, these are able to phosphorylate each other even in the absence of ligand. By that, the cell generates a continuing proliferation stimulus. In the past, this type of breast cancer had a worse-than-average prognosis. In the meantime, it is possible to interfere with this growth signal by use of the monoclonal anti-HER2 antibody trastuzumab (Herceptin®), which also directs immune mechanisms to these cells.

Class III: Ras: activation of a G-protein

The Ras protein is drifting along the inner surface of the plasma membrane, anchored by a farnesyl and a palmitoyl moiety. Ras always carries a guanosine nucleotide, GDP as default, GTP if switched on. When a growth signal comes in, a guanine exchange factor (GEF) forces GTP out and allows GTP in. With that, Ras is active for a certain time period. This period is limited by an integral enzymatic GTPase activity, formed by a sort of built-in "pliers" that "pinch off" the last of the three phosphates of GTP. This GTPase activity is fairly inefficient compared to typical enzyme activities. It is accelerated with the help of a GTPase activating protein (GAP), but the pliers themselves are an integral part of Ras. The two jaws of the pliers are glycine 12 and glutamine 61. If one of these two amino acids is substituted by something else, the GTPase is dead, and Ras cannot be switched off any more.

The classical mutation with this effect, the replacement of Glycine 12 (codon: GGC) by either Valine (GTC) or cysteine (TGC), is seen frequently in, e. g., lung cancer. In both cases, a G is replaced by a T. This may be explained by benzo[a]pyrene adduct formation with guanine, causing misincorporation of dATP at the opposite strand during DNA replication (see section on mutations). In summary, there is a direct causal chain from smoking to Ras activation.

In patients with non-small cell lung cancer, one application of personal medicine is to check whether Ras is activated in carcinoma cells. If it is, the entire EGF signal transduction pathway is constitutively active. In this case, it is useless to try expensive anti-EGF drugs like monoclonal antibodies or EGF receptor kinase inhibitors that act upstream of Ras.

Class IV: Myc: ways to activate a transcription factor

c-Myc is a transcription factor that needs to be activated for a cell to enter S phase. Mutations can activate Myc in two ways:

Amplification:

In this context, amplification means a process whereby the cell ends up with multiple copies of a gene instead of the usual two (of our diploid genome). Causes and mechanisms of this strange phenomenon are insufficiently understood. Two types of amplification can be demonstrated by fluorescent in situ hybridization (FISH):

1. Homogeneous staining regions: a chromosome carries a region containing multiple copies of the same gene, in our example, the MYC gene.

2. Double minute chromosomes: Apart from normal chromosomes, a number of additional small chromosomes appear, containing multiple copies of the gene in question.

[Mdm2, *mouse double minute-2*, obtained its name for this phenomenon. It was identified in this form in a mouse tumor cell line. Amplification of the E3 ubiquitin ligase of p53 is biologically equivalent with a loss of p53. Therefore, *mdm2* is a proto-oncogene that may be activated by amplification.]

How are these strange rearrangements generated? It was observed that single lagging chromosomes in mitosis occasionally result in separate micronuclei at the time nuclear membranes reform. These micronuclei frequently possess a low number of nuclear pores and show inadequate import of components required for the next round of DNA replication. Replication is delayed compared with that in the proper nucleus. On entry of the cell into the next mitosis chromosome condensation then causes shattering (**chromothripsis**) of the micronucleus-chromosome which still has numerous open replication forks. Chromosome pieces are later stitched together by non-homologous end joining (NHEJ) repair, resulting in absurdly rearranged single chromosomes and or small circular double minutes. In addition, fork stalling may lead to "inventive" repair attempts involving microhomology-dependent priming and serial template switching, which may explain the emergence of multiple copies of a gene on a reassembled chromosome.

Each single one of the many MYC genes amplified in one of these two ways is structured normally. Yet, there are so many of them, that if each one functions "normally", vast amounts of c-Myc transcription factor are present at any point in time, continuously driving cell proliferation.

Translocation:

Translocations involving the MYC gene on chromosome 8 as well as one of the three immunoglobulin loci (heavy chains on chromosome 14, light chains on chromosomes 2 or 22) are typical for Burkitt's lymphoma, a neoplasm of B lymphocytes. In B cells, translocations concerning these loci are facilitated by the immunoglobulin rearrangement process, a molecular random generator involving cutting and pasting of DNA. As a result, the myc gene is placed next to one of the immunoglobulin loci. The normal MYC gene is thus driven by the nearby immunoglobulin enhancer, a regulatory DNA element that is meant to drive strong expression of antibody chains. Result: normal c-Myc protein is produced in excessive amounts.

What is true for c-Myc is true for many other transcription factors as well: oncogenic activation usually implies overproduction of structurally normal protein.

Class V: deregulated expression of cyclin D

Some translocations habitually involve typical loci. Some genes were even named accordingly, e. g., bcr (*breakpoint cluster region*). Genes identified by analysis of B cell-typical translocations were named *bcl* (for *B cell leukemia*). *Bcl-1*, identified as the chromosome 11 partner gene in mantle cell lymphoma t(11;14) translocations involving the heavy chain locus, was later recognized to encode cyclin D. Due to the translocation, cyclin D is expressed even in the absence of any growth signal, promoting rapid proliferation.

Further typical translocations

Bcl-2, identified in t(14;18) translocations, has retained this name. It was later found to encode an anti-apoptotic protein localized at the mitochondrial surface. Although formally analogous to the two translocations mentioned before, this translocation does not result in faster proliferation, but prevents the cells from dying. Resulting follicular lymphoma is highly differentiated and grows slowly, with a natural course of many years. While this in itself is positive, the trouble is that follicular lymphoma is extremely hard to treat. Our therapeutical options work much better for cells that are rapidly proliferating.

t(9;22)- Philadelphia translocation

This translocation involves tyrosine kinase *c-abl* on chromosome 9 and a gene with unknown function, *bcr*, on chromosome 22 (not the immunoglobulin light chain locus; *bcr* stands for *breakpoint cluster region*). It differs from the translocations mentioned above in one important point: here, the breakpoint is located **within** the two genes. Two fusion genes result from this reciprocal translocation, where the "front ends" are derived from one gene and the back ends from the other. Functional consequences are known only for the one that at the front looks like *bcr*, at the back like *c-abl*, including its kinase domain. Expression of the fusion gene results in a fusion mRNA and a fusion protein. This protein normally does not exist in humans; functionally, it works like a deregulated Abl kinase, causing chronic myelogenous leukemia (CML).

Pharmacology cross reference: Novartis developed a more or less specific inhibitor against this fusion kinase (imatininib, Gleevec®). Results of CML treatment with imatinib are excellent.

Mutations with long-range effects

The mutations described up to here have the advantage of being educational, as it is easily comprehensible how they promote the development of an anarchist cell clone. Unfortunately, the majority of mutations causing any individual tumor refuse to be educational. Many mutations out there in the vast expanse of the DNA alter unidentified binding elements of transcription-activating or repressing proteins, affecting genes far away that may be only spatially close by some loop in the DNA. We are able to detect the resulting overexpressed oncogene or underexpressed tumor suppressor, but we don't stand a chance of identifying the causative point mutation 27 genes further up. As another example, mutations affecting non-coding RNAs are able to alter expression patterns of far-away genes.

6. SUMMARY: TYPES OF MUTATIONS IN HUMAN CARCINOGENESIS- CLASSIFICATION ACCORDING TO FORMAL CRITERIA

Three types of mutations contribute to carcinogenesis in humans:

1. Point mutations, deletions, insertions
Examples:
Proto/oncogenes: EGF receptor, Src, Ras: proteins that cannot be switched off
Tumor suppressors: pRb, p53: inactive proteins

2. Amplifications
Examples: c-Myc, Mdm2: strongly overexpressed normal proteins

3. Translocations
Examples:
Breakpoint between genes: c-Myc, cyclin D, Bcl-2: strongly overexpressed normal proteins
Breakpoint within genes: Bcr-Abl: fusion proteins

7. VIRUSES AND CARCINOGENESIS

Contribution of viruses to carcinogenesis in humans

Several viruses are known to promote specific forms of malignant neoplasia:

HPV	cervical carcinoma
EBV	Burkitt's lymphoma, Hodgkin's lymphoma, nasopharyngeal carcinoma
HBV, HCV	hepatocellular carcinoma
HHV-8	Kaposi's sarcoma
MCV	Merkel cell carcinoma
HTLV-1	adult T cell leukemia

This does not happen via a single uniform mechanism. Among the tumor-promoting viruses relevant for humans, DNA viruses predominate. DNA viruses have an incentive to drive cell proliferation, as S phase equips them with all the materials required to replicate their genome. To achieve that, viruses developed different tools. The tools of **human papilloma virus (HPV)** have been mentioned before, in the section on p53: proteins E7 an E6 inactivate the two fundamental tumor suppressors pRb and p53, respectively. "E" stands for "early", as these proteins start being expressed soon after infection of cells. Papilloma viruses infect squamous epithelium. Many of them cause only benign warts; about a third of the total of about 100 serotypes is sexually transmitted. Only few "high-risk" serotypes, including 16, 18, 31, 33, 35, 45, cause the majority of all cervical carcinomas. The process starts with a chronic infection of the squamous epithelium of the cervix, driving cell proliferation. While the virus normally remains episomal, it is found integrated into the host genome in cervical carcinoma. This is an accident, not part of the plan as with retroviruses or with hepatitis B virus. Integration leads to the highest expression levels of E7, as the counteracting E2 gene is frequently inactivated in the process. Vaccination against the most common high-risk virus types is able to prevent chronic infection and, by extension, cervical cancer caused by these types, if girls are immunized before they start sexual relations. The vaccine consists of the L1-encoded ("late") capsid protein, which is produced by recombinant DNA technology and self-assembles to virus-like particles. Pap screening remains important, though, as present vaccines (e. g., Gardasil®, Cervarix®) do not cover all relevant virus types.

Epstein–Barr virus (EBV) is a virus of the herpes family (human herpesvirus 4, HHV-4). The first infection usually occurs in childhood or during the teenage years, as the virus is transmitted via saliva ("kissing disease"). It may either go undetected, or cause a mild disease with a sore throat, fever, fatigue and lymphadenopathy: infectious mononucleosis or Pfeiffer's disease. The virus replicates first in the epithelial cells of the pharynx and oral cavity and then infects

mucosa-associated B lymphocytes. The B cells are activated and start to proliferate, producing lots of virions. As long as it takes to mount an immune response, the virus spreads rapidly to infect a large proportion of all B cells. A system of several viral proteins maintains the B cells in this expansion phase; the cells do not enter terminal differentiation and do not by themselves enter apoptosis, both aspects of immortalization. Generally, viral infections are controlled via cytotoxic T cells, which recognize viral peptides presented by MHC-I proteins of infected cells. In EBV infection, too, T cells hunt down infected B cells and kill most of them over a time span of weeks or months. Slowly, swelling of the lymph nodes recedes. Alas, the T cells never succeed in finding all infected B cells. In some B cells, the EBV enters a temporary so-called latent cycle, where it shuts down replication, including expression of most of its genes, effectively hiding from the immune system. A steady state is reached where the virus not eliminated, but is held in check. By their forties, more than 95% of people have been infected with EBV. Several malignancies, like Burkitt's lymphoma, Hodgkin's lymphoma or nasopharyngeal carcinoma, typically are EBV positive. Together with additional findings, this suggests a contributing role of the virus. On the other hand, that contribution is probably small; otherwise, malignant lymphoma would be much more common. In sub-Saharan Africa, Burkitt's lymphoma is more common than elsewhere. Here, the explanation may be a combination of EBV infection and malaria. Malaria by itself stimulates additional B cell proliferation while inhibiting T cell responses. The additive B cell stimulatory effects increase the chance for misrearrangement resulting in one of the typical translocations where the *c-myc* gene is positioned next to an immunoglobulin locus.

The mechanisms of the respective contributions of **hepatitis B and C viruses (HBV** and **HCV)** to hepatocellular carcinogenesis are insufficiently understood. In the majority of cases, carcinoma is preceded by many years of active hepatitis and cirrhosis. Reactive oxygen species and other products of inflammation are likely to cause DNA damage and secondary mutations. The effect of these mutations is enhanced by the tumor promoter effect of ongoing regeneration. For HBV, two additional mechanisms have been proposed. In spite of its DNA genome, HBV is essentially a retrovirus that is able to integrate into the host cell genome, potentially destroying antioncogenes or activating oncogenes. One of the four genes of the tiny HBV genome (3200 base pairs) encodes protein X, named for its unknown function (the other three encode HBs antigen, HBc antigen and polymerase). A range of mechanisms have been proposed on how protein X might favor hepatocellular carcinogenesis, but consensus has not been reached so far.

Kaposi's sarcoma, which was extremely rare before the emergence of AIDS, as well as primary effusion lymphoma of serous cavities, are caused by **herpesvirus type 8 (HHV-8)**. With normal immune defense, the virus is non-pathogenic; only after massive immune suppression, e. g. by HIV, HHV-8 contributes to transformation of cells. HHV-8 seems to have incorporated several host genes, including versions of Bcl-2 and Cyclin D, as well as typical E2F target genes, including thymidylate synthase, thymidine kinase and dihydrofolate reductase. These tools would seem to make HHV-8 a powerful tumor virus, but the immune system is obviously very efficient in eliminating infected cells.

Merkel cell carcinoma, a rare neuroectodermal skin tumor derived from the sensory Merkel cell nervous organ, frequently seems to be caused by **Merkel cell polyoma virus (MCV)**. Similar to SV-40, it expresses a large T protein that can be expected to inactivate pRb and p53. In tumor cells, a helicase-deficient copy of the virus is found integrated in the host genome. Again, this frequent virus is usually held in check by the immune system.

Human T cell leukemia virus or **human T-lymphotropic virus (HTLV-1)** is a retrovirus endemic in southwest Japan, the Carribean, parts of South America and Central Africa, that may be transmitted sexually or by breast feeding. It contributes to adult T cell leukemia by expanding the population of T cells, thereby increasing the probability of additional mutations.

Classical retroviruses causing tumors in animals

While the classical retroviruses do not contribute to carcinogenesis in humans, they have been of extraordinary importance in the identification of human proto/oncogenes. Following infection of a cell, retroviruses produce a DNA copy of their RNA genome with the help of reverse transcriptase. After synthesis of the second strand by host DNA polymerase, the proviral DNA genome is inserted into the host DNA. Transcription of this unit produces mRNAs for viral proteins, as well as genomic RNA for new virions. Transcription is activated by a viral promoter present at both ends of the provirus, the long terminal repeat (LTR). Protein-encoding genes are grouped into *gag* (group-specific antigen, nucleocapsid proteins), *pol* (polymerase/reverse transcriptase/integrase) and *env* (envelope) genes.

Retroviruses facilitate carcinogenesis via two principal mechanisms, insertional mutagenesis and integrated oncogenes.

Insertional mutagenesis denotes the genetic change resulting from a proviral integration event. If the provirus inserts near a gene encoding a growth or transcription factor, the viral LTR may cause massive overexpression of that protein. Mouse mammary tumor virus (MMTV) does not contain any oncogene. Yet, if it inserts near a Wnt gene in mammary cells, this leads to overproduction of this growth factor. The resulting autoendocrine growth-stimulating feedback loop contributes to the development of mammary carcinoma in mice.

The second mechanism by which retroviruses cause tumors is by **integration of an oncogene** into the viral genome. This is a rare event. The structure of most known oncoviruses may only be explained by several independent recombination events, including reverse transcription of a proto-oncogene mRNA. Once a replication-competent, oncogene-containing virus emerges, that oncogene is overexpressed in all infected cells. The term "oncogene" was coined for these "viral", tumor-causing genes at a time when little was known about growth regulation in vertebrate cells. Their connection to genes involved in growth regulation of vertebrate cells was recognized only later, leading to the designation "proto-oncogenes" ("proto" is Greek for "pre") for these physiological counterparts. Usually, the oncogenes in tumor-causing retroviruses are present in a mutated form, encoding proteins that are constitutively active.

8. HYPOXIA AND TUMOR ANGIOGENESIS

Up to here, we looked at genetic alterations leading to the "first" malignant cell. However, this cell is by no means the endpoint of tumor evolution. Critical for further tumor development is the supply of nutrients and oxygen to the proliferating cells. Autonomous proliferation of tumor cells rapidly leads to a point where diffusion distances grow too long to supply centrally located cells. Resulting hypoxia sets in motion two processes of importance to further tumor development.

The first consequence of hypoxia is **induction of p53**. If p53 is still functional in the cell clone in question, it induces G1 arrest first, and, over time, apoptosis in hypoxic cells in the tumor's center. Thus, in this area, strong selective pressure against functional p53 is created: a cell with newly mutated p53 can take advantage of this situation, continuing to proliferate in the face of arrested or dying competitors. Hypoxia thus promotes the next step in tumorigenesis, selecting a cell clone that is more malignant by the loss of p53. The respective cells have lost the G1 DNA damage checkpoint and show chromosomal instability, rapidly accumulating additional genetic problems. This selection may account for the fact that p53 is the most frequently mutated protein in human tumorigenesis.

The second mechanism is mediated by hypoxia-induced factor (HIF-1), which we encountered when considering erythropoietin. While all of the erythropoietin worth mentioning is induced in the kidney, HIF-1 is active in all cells of our body. In the presence of adequate oxygen, HIF-1 is hydroxylated, ubiquitinated and degraded. Hypoxia, in contrast, stabilizes HIF-1, which proceeds to activate its target genes with the overall goal of adapting the cell to hypoxic stress. Induction of NO synthase leads to maximal dilatation of nearby blood vessels. Induction of glucose transporters GLUT1 and GLUT3 as well as hexokinase and lactate dehydrogenase enzymes helps to generate ATP by anaerobic glycolysis. Last, not least, HIF-1 activates the gene encoding **vascular endothelial growth factor (VEGF)**. Secreted VEGF diffuses along the extracellular matrix in all directions, ultimately reaching endothelial cells of the nearest vessels. Endothelial cells react by sprouting new capillaries in the direction of the VEGF concentration gradient. On arrival, these new vessels improve the precarious conditions in the central tumor area, where cells in the meantime may have lost p53. With that, the tumor vessels fulfill a requirement for the next step in the natural history of a malignant tumor: only if there are vessels, the tumor is able to spread via the blood.

Pharmacology cross reference: The intention to block this step in tumor progression led to the development of the humanized monoclonal anti-VEGF antibody **bevacizumab** (Avastin®), which is used to treat, e. g., metastasizing colorectal carcinoma and non-small cell lung cancer. Unwanted side effects are bound to include problems with wound healing --critical in tumor patients who may require surgery on short notice--, bleeding episodes and teratogenicity.

9. INFILTRATION AND METASTASIS

Carcinogenic events considered so far mainly result in deregulated proliferation. Yet, deregulated proliferation is not equivalent with malignity or metastasis. Proliferation in itself only results in benign local growth and tumor formation. As we will see, the same formal types of mutations affect genes responsible for attachment, migration and survival in remote areas of the body. Over time, these changes shift the nature of a tumor, e. g., starting as *carcinoma in situ*, via a locally infiltrating and destructive stage to one of distant metastasis.

Infiltration and metastasis are promoted by specific molecular alterations:

- Reduced expression or direct mutation of cellular adhesion molecules like E-cadherin or N-CAM allow cells to detach from their tissue.

- Loss of cellular polarity and changes in affinity for components of the extracellular matrix. Epithelial cells are polarized, with receptors for laminin and fibronectin only at the basolateral membrane. Imagine an epithelial cell "swimming" on the surface of the extracellular matrix like a boat on the water. This unequal distribution is maintained by sorting and transport proteins within the cell. Changes in expression of proteins governing cell polarity may redistribute these receptors all over the cell surface. Being able to establish contacts with extracellular matrix at all cell surfaces, the cell submerges into the matrix like a submarine.

- The ability to degrade extracellular matrix. Invasion of extracellular matrix is facilitated by previous local degradation. This is made possible by different families of matrix-degrading enzymes. Of particular importance are matrix metalloproteinases (MMPs) like collagenase or stromelysin. Membrane-type matrix metalloproteinase (MT1-MMP) may be overexpressed as a result of tumor suppressor APC inactivation (see section on colorectal carcinoma) or activation of β–catenin, as its gene (like the *c-myc* gene) is regulated by the β–catenin /TCF4 complex.

- Increased locomotion. We have already considered signaling molecules promoting proliferation. Likewise, extracellular signaling proteins exist which promote cellular movement and migration. One of these is *hepatocyte growth factor* (yet another misnomer) or *scatter factor*. The receptor for scatter factor is encoded by the c-*met* proto-oncogene. Tumor cells overexpressing c-Met demonstrate increased motility and spreading characteristics.

- Autonomy from tissue signals normally required to keep a cell alive. The molecular basis of this autonomy is insufficiently understood. However, it is clear that cells need extracellular information not only to proliferate, but even to survive. The default setting for cells is to enter apoptosis as soon as they are not required anymore. The "you are still required"-signal is usually present in the home tissue of the cell. As soon as the cell leaves that tissue, the signal ceases, causing the cell to undergo apoptosis, e. g. by induction of a pro-apoptotic protein of the Bcl-2 family like Bmf. This mechanism has been designated *anoikis*, the Greek word for homelessness. From experimental settings, we know that for each cell that succeeds in establishing a metastasis in a remote tissue, thousands of cells having entered the blood stream die. In addition to the loss of pro-apoptotic proteins, tumors may show overexpression of anti-apoptotic proteins like Bcl-2. Bcl-2 overexpression has been reported in follicular lymphoma, hormone-refractory carcinoma of the prostate (90-100%), malignant melanoma (90%) or estrogen receptor-expressing mammary carcinoma (80-90%).

The list does not imply that each of these aspects requires "dedicated" mutations. Emigration of cells out of epithelial tissue is a physiological program, activated in defined situations. During embryogenesis, neural and melanocyte progenitors leave the epithelial neural tube and migrate to the periphery. In adults, this process, termed *epithelial-mesenchymal transition* (EMT), is activated to enable wound healing. A small cut in a finger, e. g., is first closed by a clot consisting of fibrin and platelets. Via EMT, cells detach from the edge of the epithelium, migrate through the extracellular matrix below the clot just like mesenchymal cells, and on arrival morph back to form a new epithelial layer. Morphing back is called mesenchymal-epithelial transition (MET). In other words: under defined conditions, a complex physiological program enabling epithelial tissue to start a local infiltration process may be switched on and

off. One program switch seems to be the protein twist, a transcription factor that in experiments has been shown to be able to induce aspects of EMT in epithelial cells. Local infiltration by a carcinoma may be explained as a mistimed activation of this program. The causes of this activation remain to be elucidated. Interactions between epithelial cells and stroma cells have been shown to contribute; mutations affecting program switches in either cell type may play a role, either.

10. IMMUNE RESPONSES AGAINST TUMORS

Immune responses can be successful against malignancies

This proposition can be shown by the following experiment: Cells from the tumor of a mouse are transferred to another individual of the same inbred mouse strain. If, previous to this transfer, the second mouse is first immunized with irradiated cells of the tumor, it is able to successfully eliminate a dose of tumor cells that is lethal for mice without pretreatment.

This protective effect is T cell dependent: it does not exist in mice deficient in T cells, and it can be extended to other mice by adoptive transfer of T cells.

This protective effect is antigen-dependent, either. It does not work against another tumor from the same mouse strain. Conclusion: tumors express antigenic peptides which may be targeted by tumor-specific T cells. Such peptides, termed *tumor rejection antigens*, are presented on MHC-I proteins.

Tumor rejection antigens have also been defined for human tumors: specific peptides, presented on MHC-I, that are recognized by specific T cells isolated from a patient. Why do these anti-tumor cells exist at all? Tumors are part of immunological *self*, and T cells recognizing *self* should have been eliminated in the thymus by negative selection! For each tumor rejection antigen, there is an explanation for the fact that the respective T cell clones weren't eliminated:

1. unique peptides: peptide sequences directly resulting from mutations. These peptides had not been part of immunological *self*; they were newly generated in the tumor.

2. ectopically expressed antigens otherwise only expressed in immunologically privileged zones (e. g., testis). Normally, antigens expressed in immunologically privileged zones only are not shown to the immune system; the respective T cells were never deleted.

3. cell-specific differentiation antigens (e. g., enzymes for melanin production in melanocytes). These little-expressed antigens may have remained under the radar of negative selection.

4. strongly overexpressed proteins (e. g., HER-2/Neu in breast cancer). The multitude of identical peptides presented may turn an originally weak interaction with the T cell into a strong combined Velcro effect.

5. proteins with abnormal post-translational modifications (e. g., hypoglycosylated mucin). The tumor protein looks different from the normal protein.

6. virus proteins (e. g., HPV E6 and E7)

It is common to find immune responses against tumor rejection antigens in manifest tumors, but only in exceptional cases they succeed in fighting back the tumor (occasionally, spontaneous remissions of malignant melanoma have been observed). However, tumor elimination at a pre-diagnostic stage cannot be excluded; such an event would never come to our attention. So, the big question is: is the immune system effective against malignant cells?

Immune surveillance against neoplastic cells

The term "immune surveillance" was coined by Frank MacFarlane Burnet (1899-1985, clonal selection hypothesis) in the early days (1967) of immunology, to express the expectation/hope, that the immune system be able to recognize and eliminate malignant cells. On the other hand, it is abundantly clear that malignant tumors are the second-frequent cause of death, so if it exists at all, the effectiveness of tumor surveillance seems to be limited.

Let us first look at cautionary points. The most frequent tumor types in humans and mice occur with roughly the same incidence in individuals with and without T cell defects. Those tumors that actually are more frequent, are virus-induced. The difference is easily explained by the fact that viruses encode "foreign" peptides, inviting T cells to attack infected cells. These results correlate well with data from AIDS patients. AIDS patients suffer more frequently from anal cancer, which is caused by sexually transmitted HPV, while the frequency of colorectal cancer at other locations is similar as in healthy controls. Also Kaposi's sarcoma, the signature neoplasia of full-blown AIDS, is caused by a virus, HHV-8. On the other hand, individuals with T cell defects or AIDS patients have severely reduced life expectancy from infections, which limits the validity of these observations.

Yet, other data support the concept of a quite powerful role of the immune system, at least for certain types of malignancies:

- In patients with organ transplants, especially liver transplants, who remain immunosuppressed for long periods, the frequency of skin tumors increases over time. This is true for squamous cell carcinoma, but also for basalioma and melanoma.

- Several strains of knockout mice, in which central functions of the immune system were inactivated, show an increase in tumor incidence when challenged with the carcinogen methylcholanthren. This includes knockout mice for RAG2, IFNγ, IFNγ-R, STAT1 and perforin. In another mouse model, an immune response involving $CD4^+$ T cells and macrophages was able to eliminate premalignant senescent hepatocytes in which a p53 response had been induced by transfection of oncogenic ras.

- A retrospective analysis of the course of disease in patients with advanced ovarian cancer indicated a decisive role of tumor-infiltrating lymphocytes (TILs). On average, patients with many TILs survived much longer.

- Most importantly, the impressive success rate of immune checkpoint blockers in many tumor types suggests that the immune system frequently mounts an effective response, but that malignant cells are equally adept at keeping this response in check.

The fact that tumors emerge at all is no valid argument against tumor surveillance, nor is the absence of a visible immune reaction in many tumors. Even with highly efficient immune surveillance, observed tumors would have been selected by their ability to escape it.

Granted an efficient immune surveillance against malignant cells, mechanisms exist to escape it. Remember the set of conditions necessary to mount a cytotoxic T cell response: non-self peptides have to be presented on MHC-I in combination with costimulatory proteins like B7 (CD80 or CD86). In the absence of costimulation, the presented peptide induces peripheral anergy in naïve T cells, in other words, tolerization. In tumors, the activating combination is rare, as the vast majority of presented peptides represent *self* and as cancer cells usually don't express costimulators. In those cases where all conditions have been met and an effective T cell response is under way, tumor cells ceasing to express a crucial element, e. g. the specific MHC-I molecule presenting the tumor rejection antigen, gain a selective advantage. Even if anti-tumor T cells are at the scene, there are ways for tumor cells to escape. Sometimes, tumor cells produce immunosuppressive cytokines like TGFβ or IL-10, which directly inhibit T cell function. A frequent and successful strategy for malignant tumor cells is to activate inhibiting receptors present on active infiltrating T cells ("hit the T cell's off-button"). These include the so-called "immune checkpoint" proteins CTLA-4 and PD-1, which are structurally related to CD28 (careful: this "immune checkpoint" is not in any way related to the cell cycle checkpoints). If PD-1 is activated by its ligand PD-1L, signal transduction via the T cell receptor is interrupted. Many cell types constitutively express PD-1L, and many tumor cells do so, too. By hitting PD-1 on a cytotoxic T cell, the tumor cell may escape in the nick of time, even if a matching T cell receptor is present.

Apart from T cells, an alternative model for early elimination of tumor cells invokes natural killer (NK) cells, from the innate immunity tool box. While investigated intensely in laboratory settings, the mechanism's importance in vivo is unclear. Tumor cells may activate NK cells either by loss of MHC expression, or by expression of specific stress-induced NK cell activating ligands like MICA (*MHC-I chain-related A*), which may reflect cell stress resulting from oncogenic transformation.

Malignant melanoma is a model where immune effects seem especially promising

In 2003 (New Engl. J. Med. 348: 567), two recipients of renal transplants were reported to have developed secondary malignant melanoma transferred with the transplanted organ. The donor had had a malignant melanoma surgically removed 16 years before she died from other causes. For all those years, her immune system obviously had been able to hold micrometastases in check that had spread throughout her body.

Interventions to de-repress an existing anti-tumor immune response

A resounding success have been attempts to prevent tumor cells from hitting the CTLA-4 or PD-1 immune checkpoint buttons in situations where cytotoxic anti-tumor T cells had already been generated. The concept first proved successful in the melanoma model, for which response rates up to 60% were reported, but has since been extended to many other malignancies.

Pharmacology cross reference: Immune checkpoint blockers include the following antibodies that have been approved for melanoma and are being investigated for many additional tumor types:

- **Ipilimumab** binds to CTLA-4, blocking its interaction with B7.
- **Nivolumab** binds to PD-1, preventing its activation by tumor-expressed PD-1L.
- **Pembrolizumab** blocks PD-1 in the same way.

Obviously, this strategy only works if there is already a T cell response. If that is not the case, it might be possible to induce one:

Interventions with the goal of inducing anti-tumor immune responses

For melanoma, a number of tumor rejection antigens have been defined:

- MAGE antigens (*Melanoma AntiGen Encoding* gene; not expressed in normal adult tissue with exception of the testis)
- peptides from the tyrosinase enzyme (catalyzing the first step in melanin synthesis)
- gp100 (melanocyte differentiation antigen)
- MART1 (*Melanoma Antigen Recognized by T cells*; melanocyte differentiation antigen)
- gp75 (melanocyte differentiation antigen)

Melanoma patients were immunized with these antigens according to various protocols. Two types of protocols demonstrated encouraging results in some studies: first, direct immunization with antigens or peptides from antigens, combined with different adjuvants and/or cytokines (GM-CSF, IL-2, IL-12); second, immunization with autologous dendritic cells, loaded in vitro with these antigens; in this case, costimulatory proteins like B7 help to activate naïve T cells. For both strategies, some studies report response rates in the range from 20 to 30%, others find no effect at all.

In planning immunity-boosting interventions against cancer, an important concern is not to prompt an immune reaction against the tissue from which the tumor originates. In the strategies attempted to treat melanoma, this danger is exemplified by the occurrence of vitiligo.

Pharmacology cross reference: With Sipuleucel-T, a therapy along these lines has been approved for prostate cancer in the US, while it was retracted in the EU. Antigen-presenting cells are extracted from the patient and loaded/stimulated with the help of a fusion protein consisting of GM-CSF and prostatic acid phosphatase. Matured APC presenting phosphatase peptides on their MHC proteins are re-infused into the patient to stimulate a prostate-specific immune response.

11. GENETIC PLASTICITY OF TUMORS UNDER THERAPY: MECHANISMS OF RESISTANCE

The same mutational mechanisms that generated the first malignant cell continue to modify the natural development of the disease. In particular, they are the basis for the ability of subclones to develop resistance to chemotherapy.

Alkylating agents are a frequently used drug class in chemotherapy. A typical example is cyclophosphamide (e. g., Endoxan®). Their essential ability is to bind to macromolecules by forming covalent bonds with atoms providing free electron pairs. Like aflatoxin, they bind, e. g., to O6 or N7 of guanine. One out of a range of cellular resistance mechanisms involves glutathione, abbreviated "GSH", a small molecule consisting of three amino acids with the crucial feature of a cysteine-SH group that fits the bill of free electron pairs. Every alkylating molecule inactivated is one less to damage DNA. A family of enzymes, glutathione S-transferase, catalyses inactivation of alkylating agents by GSH-coupling. Tumor cells that succeed in overexpressing glutathione S-transferase become resistant to alkylating agents. (We will consider another mechanism, inactivation of the MMR system, in the section on colorectal carcinoma.)

Methotrexate, from the drug class of "antimetabolites", causes its anti-tumor effect by inhibition of nucleotide synthesis. In purine synthesis, as well as in synthesis of thymine from uracil, transfer of single carbon groups is dependent on tetrahydrofolic acid, which is transformed to dihydrofolic acid in the process. The enzyme dihydrofolate reductase (DHFR, already mentioned as an E2F target gene) regenerates the tetrahydrofolate pool. Methotrexate resembles folic acid, binding and blocking DHFR. If a tumor cell succeeds in increasing the number of intracellular DHFR molecules to an extent that part of them is not blocked any more by therapeutic methotrexate concentrations, it grows resistant to methotrexate. This frequently happens by amplification of the DHFR gene.

Vinca alkaloids like vincristine (e. g., Oncovin®) are plant products from Catharanthus roseus (formerly Vinca rosea). Vincristine binds to tubulin monomers, thereby preventing microtubulus polymerization and blocking spindle formation. One way to develop resistance is a point mutation that slightly modifies the monomer, interfering with vincristine binding but not with polymerization. Another escape has very negative effects on the chances of successful treatment: amplification of the MDR1 (multidrug resistance) gene. We already encountered this gene in the context of stem cells, which express it stronger than other cells. MDR1 encodes a transmembrane protein, P-glycoprotein, with the ability to pump "alkaloid-like" molecules out of the cell. A tumor cell overexpressing this pump grows resistant not only to vinca-alkaloids, but also to other drug classes, including **anthracyclines** (e. g., Doxorubicin, Adriamycin®). Anthracyclines, produced by Streptomyces bacteria, block DNA replication by intercalating between bases of DNA and inhibiting topoisomerase II.

Glucocorticoids are an important component of chemotherapy against lymphatic leukemias and lymphomas, as they are frequently able to induce apoptosis in these cells. Cells may become resistant by loss of function-mutations or by downregulation of the receptor.

These examples illustrate the following principles:

- The tumor evolution process of mutation and selection continues under chemotherapy, contributing to the development of resistance.

- Polychemotherapy is an absolute necessity. As cells are able to "invent" antidotes to any chemotherapy agent, the combination of several agents lowers the chances that a single cell succeeds in making all the necessary inventions in time to survive. Only a cell that succeeds in 1) amplifying glutathione-S-transferase, 2) amplifying MDR and 3) inactivating the glucocorticoid receptor within a short time frame is able to survive a protocol combining

cyclophosphamide, doxorubicin, vincristine and prednisolone (the classic CHOP protocol used to treat lymphoma). The odds of that are much lower than those of developing resistance against a monotherapeutic agent.

- At the same time, this implies that a recurring tumor is genetically diverse from its predecessor at the time of diagnosis. Chances for successful treatment are low, as in the meantime, the relapsing tumor has developed resistance against the majority of well-established drug classes.

12. SOME MOLECULAR FEATURES OF COLORECTAL CARCINOMA

Tumor suppressor APC

APC stands for adenomatous polyposis coli. Members of families with this hereditary syndrome (synonym: familial adenomatous polyposis, FAP) develop hundreds of colonic polyps, some of which invariably lead to malignancy. The genotype/phenotype pattern is similar to those of hereditary retinoblastoma or Li-Fraumeni-syndrome: affected patients inherit a defective APC allele; in cancer cells, the second allele is inactivated, too. Once more, the importance of APC is not limited to the eponymous rare disease; in about 60% of all colorectal carcinomas, APC function has been lost by somatic mutation.

[Disambiguation: the abbreviation APC is also used for anaphase-promoting complex –the ubiquitin ligase complex inhibited by the spindle checkpoint-- and for antigen-presenting cells]

The loss of APC is thought to promote cancer via several independent pathways, due to independent biological functions of the protein.

In its first function, APC acts like a brake on the Wnt signal transduction pathway, which is important to maintain the population of intestinal epithelial cells. Specifically, this is achieved by antagonizing β-catenin. β-Catenin helps to "chain" adhesion protein E-cadherin to the cytoskeleton. Surplus β-catenin not required for this mechanical function is earmarked for ubiquitination and proteasomal degradation by phosphorylation. Phosphorylation, which happens only in the absence of extracellular growth factor Wnt, is dependent on a kinase complexed with APC, glycogen synthase kinase 3β. In the presence of Wnt, the protein complex containing APC and the kinase is inhibited; unphosphorylated β-catenin escapes degradation and accumulates, entering the nucleus. At the promoters of target genes, it removes a repressor from transcription factor TCF4/LEF (T cell factor-4/ lymphoid enhancer factor), activating transcription of genes required for proliferation, including Cyclin D and c-*myc*. The physiological importance of this signaling pathway is illustrated by TCF4 knockout mice, which die soon after birth due to a lack of intestinal epithelial cells.

Wnt is produced by the stroma cells at the base of intestinal crypts. It drives prolifation of epithelial cells in the crypt. The farther the epithelial cells move out, the less Wnt they see, until they stop proliferating. In summary: in the presence of Wnt, intestinal cells proliferate. This is mediated by a transcription-activating function of stabilized β-catenin.

In the absence of Wnt, cells stop proliferating. This requires breakdown of surplus β-catenin with the help of APC.

If both APC alleles in a cell are inactivated by mutations, the cell loses its ability to inactivate β-catenin, even in the absence of Wnt. Accumulation of β-catenin causes persistent expression of cyclin D and c-Myc. In other words: the loss of APC simulates persistent Wnt signaling.

Simulation of Wnt signaling should also be possible with active APC if β-catenin is stabilized by other means. In fact, 10% of colorectal carcinomas have been found to overexpress β-catenin. In these cases, β-catenin has oncogene function.

The APC/β-catenin example illustrates how activation of a cell-specific growth signaling pathway is essential to promote cancer. This is possible either by loss of braking elements (inactivation of tumor suppressors) or by activation of stimulating elements (activation of proto/oncogenes).

A second function of APC has been described in mitosis. In cooperation with a host of other proteins, it helps to attach chromosomes to the mitotic spindle. In addition, it also contributes to anchoring the spindle to the cell membrane at a defined position. Loss of APC reduces the precision with which chromatids are distributed to daughter cells. Therefore, loss of APC, like loss of p53, results in chromosomal instability (CIN) with further deterioration of genome quality.

In addition, there is evidence that the attachment of the mitotic spindle is of special significance in stem cells. The asymmetric characteristics of stem cell division may be dependent on the correct positioning of its axis. In the absence of APC, the axis may drift, occasionally resulting in symmetric division producing identical daughter cells. Over time, this mechanism may increase the number of stem cells in a crypt, subjecting a larger number of cells with unlimited proliferation potential to the process of accumulating mutations. This latter function may be the reason why loss of APC is frequently the first step in colorectal carcinogenesis.

HNPCC tumor suppressors

A second familial syndrome predisposing affected members to colorectal carcinoma is hereditary non-polyposis colorectal cancer (HNPCC). As its name implies, it does not include the formation of multiple adenomatous polyps. The clinical criteria for HNPCC: colorectal carcinoma in three members of a family, with at least two generations affected and at least one of the tumors diagnosed before age 50. Other tissues, especially endometrium, are affected by the tendency to develop malignancies, too. In some families, endometrial cancer predominates. About 5% of all instances of colorectal carcinoma can be attributed to HNPCC. Frequently, these develop in the ascendent colon, while colorectal carcinoma is typically located in sigmoid or rectum. In contrast to FAP, where 100% of patients over time develop carcinoma, in HNPCC this number is around 80%. Once the predisposition syndrome is diagnosed, it is thus essential to closely monitor colon and uterus.

When searching for "the" HNPCC gene, it became clear that the phenotype segregated with different chromosomal loci in different affected families. HNPCC was found to be due to the inactivation of the mismatch repair (MMR) system (see section 2). Most frequently, MLH1 and MSH2 are mutated, but other components may be affected as well. Like with APC, MMR

defects are not only found in the hereditary form, but also in a considerable percentage of sporadic cases of colorectal, endometrial and gastric carcinoma.

MMR defects may be diagnosed by a specific form of genetic instability. "Microsatellites" are repeats of single or a few bases, like TTTTTTTT or ATATATATATATATATATAT, which are obviously hard to copy faithfully. With defects in the mismatch repair system, the number of repeats within a microsatellite tends to drift over cell generations. This phenomenon is termed microsatellite instability (MIN). The phenomenon is of diagnostic importance, as it is easier to determine satellite instability than to check all components of the complex repair system for mutations.

We already considered one resistance mechanism against alkylating agents: overexpression of glutathione S-transferase. Defects in the MMR system may induce resistance, too. Alkylating agents cause methylation or choroethylation of guanine on oxygen atom 6 (O6). As mentioned before (in section 2, "Mutations"), this causes a G•T-mismatch during the next replication. A dedicated protein exists to remove the O6 methyl group under physiologic conditions: O6-methyl guanine methyl transferase (MGMT). However, this protein works only once, as it is subject to "suicide inactivation" by accepting the alkyl group. Under therapy, the MGMT pool is instantly overwhelmed, leaving the majority of O6-alkylations in place. Following DNA replication, these are identified by their mismatch with T by the MMR system. The MMR system attempts to start repairs, but in the continuous presence of alkylating agents, repair is inefficient, and for reasons that are not completely understood (continuous checkpoint signaling resulting in apoptosis?), cells are worse off by trying. Somehow, the involvement of the MMR system kills the cell- the main way how alkylating agents work. In other words, for alkylating agents to work, MMR has to work. Tumor cells may grow resistant by inactivating their MMR system, blissfully ignoring G•T-mismatches. What results is a hypermutable MMR deficient cell clone that is resistant to alkylating agents.

13. SOME MOLECULAR FEATURES OF BREAST CANCER: BRCA GENES

Defects in tumor suppressors BRCA-1 and BRCA-2 contribute strongly to the development of mammary and ovarian carcinoma. They are essential components of several DNA repair systems, including one to repair double strand breaks. Human cells have two systems to restore continuity following a double strand break: non-homologous end joining (NHEJ) and homologous recombination (HR). The essential difference between the two systems: perfect repair can only be achieved by HR, but HR is only possible when the DNA has already been replicated. In all other situations, NHEJ must be employed, which leads to incomplete repair.

Non-homologous end joining

The NHEJ repair process is started by the appearance of free DNA ends. As a prerequisite for reattachment, it is usually necessary to pare back the ends, as these are frequently damaged. For example, nucleotides may have been chemically modified by the energy of ionizing radiation. Inevitably, this entails the loss of a few nucleotides; NHEJ is therefore the main cause of small deletions. As the overwhelming majority of the human genome is non-coding, this is still a good solution, much better, in fact, than losing all the genes on the non-attached peripheral fragment.

The signaling pathway of the G1 DNA damage checkpoint, with chromatin alterations (Ku, γ-H2AX) and activation of several kinases (DNA-PK, ATM, CHK2), has been described before (see section on p53). The kinases phosphorylate and activate repair proteins, like XRCC4 and NBS1 (XRCC stands for *X-ray Repair Complementing defective repair in Chinese hamster cells*; NBS for *Nijmegen breakage syndrome*). Damaged DNA ends are pared back by a complex with exo- and endonuclease activity, consisting of proteins MRE11-RAD50-NBS1. (In many cases, designations for components of repair systems have been derived from experiments in yeast. Strains with noticeable features, e. g. following treatment with radiation, were isolated to identify the involved genes, leading to names like radiation-sensitive—RAD or Meiotic REcombination deficient—MRE.)

The last step in NHEJ is catalyzed by a DNA ligase which binds to Ku-proteins via the adapter XRCC4 and restores the continuity of the double strand. Via this sequence of steps, NHEJ repairs the vast majority of all double strand breaks, leaving only a small deletion as a mark.

Homologous recombination

Yet, in some situations NHEJ is not able to do the trick. The strongest evidence in that regard comes from experiments with a protein central to HR, RAD51. Cells with intact NHEJ, but lacking RAD51, are not viable: when trying to replicate their DNA, they commit so many errors that they eventually die. These and related experiments indicate that HR is a necessary retouching mechanism for DNA replication. In DNA replication, double strand breaks frequently occur at the lagging strand immediately behind the moving replication fork, e. g., if the discontinuous Okazaki fragments coincide with a single strand break on the opposite strand. This situation is favorable for HR, as an identical strand –the nascent sister chromatid— is right at hand.

In essence, homologous recombination uses the sister strand as a template to repair the break. A 3'-overhanging single strand invades the sister double strand, displacing and putting itself at the position of its copy with the help of proteins including RAD51. Using the nascent sister chromatid's strand as a template, the broken strand can now be elongated beyond its original break point. This happens from both directions, bridging the gap at both strands, followed by strand resolution. The process results in perfect repair without deletions.

Step by step, the process works as follows. After a double strand break in the wake of a replication fork, HR recombination is initiated by 5' to 3' resection of the 5' ends by the MRE11-RAD50-NBS1 complex' exonuclease activity. This generates a 3' overhang at the opposite strand, liberating it to start strand invasion. Strand invasion requires RAD51, RAD52, the tumor suppressors BRCA1 and BRCA2 (*breast cancer*), as well as several additional proteins. The 3' end of the invading strand is then elongated by a DNA polymerase, progressively displacing the autochthonous strand. The same is done from the opposite direction. Remaining nicks are ligated. Eventually, the two interwoven strands are separated by cleavage and religation ("Holliday junction resolution"). The replication fork can resume working.

Loss of BRCA1 or BRCA2 function reduces HR efficiency dramatically. In this case, replication problems are likely to be patched up by other means (e. g., NHEJ or error-prone repair) at the cost of a steep increase in mutations or chromosomal aberrations. Obviously, this

is not compatible with development of an organism: knockout of BRCA1 or BRCA2 in mice is embryonally lethal.

Pharmacology cross reference: one strategy to treat carcinomas based on a loss of BRCA is to block backup repair systems. If neither homologous recombination nor backup systems are left to repair damage, the DNA problems accumulate to a degree that the cells cannot survive. **Olaparib** inhibits the enzyme PARP (Poly ADP Ribose Polymerase), which is involved in reporting DNA damage to repair systems. It is used in relapsed ovarian cancer with loss of BRCA1 or BRCA2 function if other forms of chemotherapy remain unsuccessful.

Defective DNA double strand break repair is of major importance in carcinogenesis: many malignant tumors, especially colorectal, breast, prostate cancer and carcinoma of the pancreas regularly show chromosomal instability (CIN), with loss of heterozygosity affecting numerous alleles. While frequent, defective DNA repair is not the only cause: mutations affecting genes with roles in cell cycle checkpoints or chromosome handling and transport lead to the same phenotype.

Tumor suppressors BRCA1 and BRCA2

BRCA1 and BRCA2's role is not limited to homologous recombination; they are also part of additional repair systems. Yet, well before their involvement in repair was recognized, the two proteins had been isolated by their association with breast cancer.

The age distribution of breast cancer incidence does not follow a simple exponential function; there are "too many" early cases. Genetic factors have been estimated to contribute to 5% of all breast cancers, yet to 25% of those diagnosed before age 30. Cosegregation analyses in families with increased incidence of breast cancer identified two genetic loci, termed BRCA1 and BRCA2. Subsequently, candidate genes were identified, enabling diagnostic procedures at a time when the two genes' functions were entirely unclear. Analogous to hereditary retinoblastoma or Li-Fraumeni syndrome, affected women had one loss-of-function allele in every cell; in the tumor cells, the second, normal allele had usually been lost (loss of heterozygosity). That explained the increased risk of these patients to develop another carcinoma in the second breast. BRCA1 and BRCA2 thus behave like typical tumor suppressors. The presence of a defective allele may be diagnosed and confronts the carriers with the difficult decision whether to undergo prophylactic mastectomy or to trust other prophylactic measures like tamoxifen, early ovariectomy and frequent specific check-ups.

Inherited mutations in BRCA1 strongly increase not only the probability of developing breast cancer (lifetime probability 80%), but also ovarian cancer (40%). Males, in contrast, have a slight increase in their (small) risk of breast cancer, as well as a moderate increase in the risk of prostate cancer, pancreatic cancer and melanoma. As a heterozygous male's increase in the overall risk to develop cancer is much smaller than that of a heterozygous female, mechanisms to account for this gender specificity are being investigated. There is evidence for a role of BRCA1 in X-inactivation— the inactivation of one of the two X-chromosomes in female cells. Loss of BRCA1 would thus, in addition to impairing DNA repair, redouble expression of all X-encoded genes, but only in females, not in males! Which of the affected genes might contribute to mammary and ovarian carcinoma is not clear. Another model is based on the fact that BRCA1 also functions as a E3 ubiquitin ligase, with estrogen receptor α as one of its substrates. Lack

of ERα inactivation might explain why females are disproportionately affected, and why this effect is most pronounced in tissues depending on estrogen for proliferation.

In humans, heterozygous as well as homozygous genetic BRCA2 defects occur; in both cases, tumor risk is increased. As with BRCA1, heterozygous BRCA2 defects confer high risk of breast cancer (lifetime probability 80%) and ovarian cancer (20%). Homozygous defects, in contrast, are one of the causes of Fanconi anemia.

Fanconi anemia is a rare autosomal recessive syndrome caused by the loss of one out of at least 12 genes. Symptoms usually start in childhood and include a variable spectrum of malformations, progressive pancytopenia (aplastic anemia) and a high risk of developing acute myelogenous leukemia (AML) or head and neck tumors. Cells of Fanconi anemia patients show genetic instability. The products of the twelve genes are likely to cooperate in a common pathway that helps cells to cope with some form of DNA damage. BRCA2 was found to be one of the causative genes. This seems to contradict experimental data from the mouse, where BRCA2 knockout mice die in utero. Yet, BRCA2 alleles of patients with Fanconi anemia still allow expression of part of the protein. Thus, expression of this protein part may allow survival.

***