Chapter 1
Basic Cancer Biology

1.1 Overview of Cancer

Cancer has been recognized since early times, but treatment protocols and medications have lagged, by millennia, the initial observations of the disease. The tragic cases of childhood and teenage cancer notwithstanding, most cancers develop in the aging population, consistent with the nature of metabolic, genetic and other alterations discussed below and in various chapters. Epidemiological data show that, behind heart disease, cancer is the second leading cause of death worldwide, and many expect that in time cancer will overtake heart disease as the leading cause of mortality. Some 150 years ago it was demonstrated that cancer is composed of cells with morphology differing from that of normal cells. With information becoming available from numerous areas in biology and medicine, and capitalizing on major advances in technology, great strides were made in the twentieth century in unraveling many of the complexities of cancer, work that is continuing at an accelerating pace in the twenty-first century. It is now recognized that by far the majority of all cancers arises from environmental factors, metabolic disturbances, somatic mutations, and other pathophysiological processes (discussed throughout the book), while the remaining ones are attributable to germline mutations and are thus inheritable (familial).

In the early development of vertebrates, the embryonic stem cells undergo differentiation into the three primary cell layers, ectoderm, endoderm, and mesoderm. These, in turn, differentiate to give the 200-plus cell types of the human body comprising the myriad organs and supporting structures. The tissues can be categorized into four main groups, the epithelium, mesenchyme, nervous system and reticuloendothelial system, which in time can become subject to the development of cancer. It is also believed that normal cells throughout the body are continually in the process of undergoing changes that can result in cancer; fortunately, these events are spread over many years. From this it follows that, while one may die from cancer, individuals will often die from other causes before the cancer develops sufficiently to cause death. Clearly, the changes alluded to, as well as their rate of formation, depend on

Y. Xu et al., *Cancer Bioinformatics*, DOI 10.1007/978-1-4939-1381-7_1

many variables such as genetic background, diet, environmental factors, etc. With tobacco smoke as the best documented example, one can convincingly argue for the importance of one's lifestyle in enhancing or diminishing the possible development of cancer.

Cancer has been considered by many investigators as a genetic disease, generally involving sequential random mutations and epigenetic changes. There is, however, now a school of thought being actively pursued by many scientists that the origins of cancer lie in cellular and micro-environmental perturbations that, in turn, can lead to genetic alterations or selection of such alterations. Indeed, cancer is now recognized as a very heterogeneous disease, even within the same type of cancer, and it may emerge that its origins can be attributable to a number of causes.

As discussed below and throughout the book, there are many metabolic/cellular micro-environmental disturbances and combinations of genomic alterations that can lead to cell transformation. Once established, or when being established, many other mutations accumulate in the tumor cells, each giving rise to clonal expansion. Regardless of the initiating cause(s) of cancer, there will be in time genetic alterations, e.g., mutations, amplifications, deletions and translocations, that facilitate growth, inhibit apoptosis (programmed cell death) and escape from immune destruction. The cells harboring metabolic alterations, micro-environmental changes and mutations that provide a growth advantage and best meet the other requirements for continued tumor survival will prevail, and the processes of natural selection and survival of the fittest and most adaptable become crucial for these cells. Thus, while Darwinian principles were originally proposed to explain the evolution of organisms, a similar rationale appears to underlie tumor progression. These events may lead to cellular heterogeneity, particularly since new mutations can arise due to loss-of-function of negative cell cycle regulators such as *P53* and perhaps even by gain-of-function of positive cell cycle regulators such as *RAS*, leading to persistent cell division and a statistical chance of errors in replication.

The following quotation (Eifert and Powers 2012) nicely summarizes the current thinking on the genetic component and alludes to the challenges ahead. "*Diversity and complexity are hallmarks of cancer genomes. Even cancers that arise from the same cell type can harbor a range of different genetic alterations that facilitate their unrestrained expansion and eventual metastasis. As a result, the behaviour of individual tumours—how they progress and eventually respond to therapy—can be varied and difficult to predict*." Cancer development, survival and growth are, however, also heavily influenced, if not caused, by many of the aberrations in cancer metabolism and the microenvironment in which the tumor is located. Indeed, as alluded to above and discussed later in this book, some of these non-genetic alterations may become driving forces for the possible formation and/or survival of cancer. Another quotation is germane to a more holistic perspective of cancer (Nakajima and Van Houten 2013). "*The tumor must be recognized as an evolving ecosystem, adapting constantly to oxygen and nutrient availability*".

Large scale cancer genome sequencing is occurring at a rapid pace, and already the data are showing the extraordinary genomic complexity of tumors. It is common to find thousands, tens of thousands, or even hundreds of thousands of mutations and other genetic changes in a typical epithelial tumor. A working hypothesis was

that only a limited number of the genetic alterations are necessary to initiate and/or propagate tumor formation in a single cell and that this genetically altered cell undergoes clonal expansion with increasing genetic changes. The few early key alterations are said to be "driver mutations" that confer a growth and survival advantage, in effect leading to the conversion of a normal cell, or one that is on the road to transformation from non-genetic causes, to one that is transformed and capable of sustained growth. The multitude of additional mutations are denoted as "passenger mutations" that are not required for tumor growth or survival. As discussed later, the driver mutations, at least for certain cancers, may occur sequentially, but whether there is any order to the process, whether there are many genes that can participate and how the genetic changes relate to phenotypic changes are not known (Ashworth et al. 2011).

The remainder of this chapter is focused on a succinct review of some of the aspects of cancer that are deemed important in its formation and growth. These sections will set the stage for the chapters dealing with *omics*-based cancer studies elsewhere in this book.

1.2 Hallmarks of Cancer

In 2000 Hanahan and Weinberg (2000) proposed six hallmarks of cancer to provide a framework for a better understanding of the basic molecular and cellular principles responsible for the development and maintenance of neoplasia, hallmarks that were extended in 2011 to a total of eight (Hanahan and Weinberg 2011). It is worthwhile to briefly review these hallmarks since they offer a rational understanding of the necessary changes that are required of normal cells to make the transition to a state of perpetual growth and survival. Suffice it to mention at this point that most of the following alterations can be attributed to one or a combination of the following: metabolic changes, hypoxia, extracellular matrix (ECM) alterations, epigenomic changes or somatic mutations, including chromosomal rearrangements, of key players in or regulators of the growth promoting or cell cycle pathways.

1.2.1 Sustained Proliferative Signaling

Unlike normal cells that tightly regulate their cell division, transformed cells have the ability to perpetuate growth-promoting signals and become refractory to growth-inhibiting processes. A variety of molecular mechanisms can contribute to sustained signaling for cell division, including the following examples: hyaluronic acid fragments (see Chap. 6), a constant supply of growth-promoting signals originally designed for tissue repair, constitutively activated (gain-of-function) growth factor receptors, a constitutively activated component of the cellular pathway for cell division, and the constitutive inactivation (loss-of-function) of growth-inhibiting components of the pathway for cell division A.

1.2.2 Evasion of Growth Suppressors

There are a number of negative regulators of the cell cycle, e.g., *RB* (retinoblastoma) and *P53* (tumor protein of 53 kDa) being two of the best known and studied, that must be overcome or evaded to ensure continued division of the aberrant cells. These two so-called tumor suppressors function in large part in responding to extracellular and intracellular signals, respectively. These important suppressors of growth are part of larger complex networks that in some manner serve to introduce redundancy in the regulation. In this vein, it should be mentioned that the ECM is important in modulating the balance of growth factors and growth suppressors. For example, when the ECM is altered from a highly elastic state to a one that is stiffer, the efficaciousness of growth factors can increase by 100-fold (see Chap. 4).

1.2.3 Resisting Cell Death

The cellular process of apoptosis (cell death or cell suicide) serves to rid the body of damaged or aged cells and is a powerful barrier to the development of cancer. *BAX* and *BAK* are two important mitochondrial membrane proteins that act to begin the process of apoptosis by disrupting the mitochondrial membrane and thus releasing cytochrome c; this in turn leads to the activation of caspases, a family of proteases key in releasing the apoptotic effectors. In opposition to this pathway are anti-apoptotic members of the *BCL2* family of proteins such as *BCL2*, *BCLB* and *MCL1*. Tumor cells have developed several mechanisms for overcoming the apoptotic pathway including the loss of *P53* function (a common alteration in cancer cells) and others that are actively being studied.

1.2.4 Enabling Replicative Immortality

Located on the ends of chromosomes, telomeres, composed of hexanucleotide repeats, are shortened as cells undergo progressive divisions. In time, after multiple divisions the telomeres become sufficiently shortened that cells are no longer viable, leading to senescence and eventual cell death. This seems to be the major reason that non-immortalized cells have a finite number of divisions and thus a finite life span. Telomerase is the enzyme responsible for adding these protective repeat segments of DNA to chromosomes, but it is present at progressively lower levels as cells divide. In contrast, cancer cells maintain relatively high levels of telomerase, thus ensuring that telomere shortening is minimized. In addition to the maintenance of telomere length, telomerase is now believed to also have other cellular functions related to growth.

1.2.5 Activation of Invasion and Metastasis

Carcinoma, the most common form of cancer and the main focus of this book, arise from epithelial cells that are engaged with neighboring cells and with the ECM. The protein E-cadherin is a well characterized cell-cell adhesion molecule, while interactions between cells and the ECM are regulated by other proteins (see Chap. 10). The processes of invasion and metastasis require several steps. First, the transformed cells must become disengaged from their interactions with other cells and with the ECM. This involves down-regulation of E-cadherin accompanied by metalloproteinases and cysteine cathepsin proteases, many of these being supplied by immune cells near the primary tumor. In addition, stromal cells neighboring the tumor, in response to signals from the cancer cells, secrete proteins facilitating invasiveness. This set of events is termed the *epithelial-mesenchymal transition* and also includes the ability of cancer cells to inhibit apoptosis. Second, the now loosely attached transformed cells undergo intravasation into blood and lymphatic vessels in their vicinity. Third, colonization to a distant site(s) then requires successful travel via the blood or lymph followed by the process of extravasation. Finally, growth of the cancer cell(s) at the new site completes the process of metastasis. Each of these processes requires many alterations in cell function that are systematically being investigated (see Chaps. 10 and 11).

1.2.6 Induction of Angiogenesis

The high energy requirements of tumors, both primary and secondary, necessitate a good blood supply for continuing availability of oxygen, nutrients and precursors for fuel-generating metabolic pathways. Angiogenesis refers to the sprouting of new blood vessels from existing ones, i.e. those produced during embryogenesis. This process is regulated by the protein, vascular endothelial growth factor-A (*VEGFA*), which acts through tyrosine kinase receptors to ensure the continued biosynthesis of new vessels. Except in a few physiological and pathological states, e.g. cancer, angiogenesis is quiescent in the adult, being inhibited in large part by thrombospondin-1.

1.2.7 Evasion of Immune Destruction

During evolution humans have developed a most sophisticated immune system, often discussed in two categories, the innate and the adaptive. The immune system is believed to be highly effective in protecting the body from the growth of transformed cells, both virally and non-virally induced. From this argument, one can argue that the cancers that do emerge have, in some manner, escaped immune surveillance or have developed the ability to counter an immune attack, particularly from T helper cells and natural killer cells, as discussed in details in Chap. 8.

1.2.8 Reprogramming Energy Metabolism

In the 1920s Otto Warburg reported that cancer cells increase their rate of glycolysis many fold over that of non-cancer cells. This reprogramming event occurs even in the presence of an ample supply of oxygen that would normally dictate that the end-product of glycolysis, pyruvate, would be converted to acetyl-CoA that in turn would enter the tricarboxylic acid (TCA) cycle (also known as the citric acid or Krebs cycle), eventually accounting for the conversion of oxygen to carbon dioxide and the generation of ATP. The putative regulatory factors responsible for this altered course of glucose metabolism will be discussed later, one hypothesis to account for the Warburg effect being that intermediates in glycolysis can be shuttled into other metabolic pathways for the biosynthesis of amino acids and nucleosides, required components for protein and nucleic acid synthesis, respectively. The important role of the glucosaminoglycan, hyaluronic acid, cannot be overlooked in cancer metabolism. This topic is briefly mentioned in Sect. 1.10 below and greatly elaborated on in Chap. 6.

1.2.9 Other Considerations

In addition to these delineated eight hallmarks of cancer, Hanahan and Weinberg also discussed processes defined as enabling characteristics of cancer: (a) genome instability and mutation, and (b) tumor-promoting inflammation. They concluded that the reduced cellular efficiency in genome maintenance and repair ultimately increases the rate of developing viable phenotypes of the cancer cells. The presence of immune cells in tumors prompted studies into their possible functions. Tantalizing results show, paradoxically, that the immune cells, normally charged with protecting the body, can aid tumor growth by secreting growth factors, pro-angiogenic factors, survival factors, and others that contribute positively to the survivability and growth of the tumor as discussed in detail in Chap. 7.

1.3 Proto-oncogenes, Oncogenes and Tumor Suppressor Genes

As discussed earlier, the cancer genome tends to contain numerous mutations and genomic rearrangements, but a central question is: *Are these causal for cancer or important for cancer growth and survival*? The introduction of the concept of an oncogene in the 1960s clearly represented a major breakthrough in defining an intellectual framework for studying cancer. It has provided useful guiding information in elucidating cancer mechanisms, particularly cancer drivers. However, this well-accepted concept seems, unfortunately, to have also restricted the thinking of

cancer researchers somewhat since it requires that an oncogene must be the mutated or overexpressed form of a proto-oncogene, which is defined as genes involved in cell growth and differentiation. Originally attributed as being responsible for the origin of cancer, recent thinking by many has shifted the role of oncogenes from that of the originator to genetic alterations that arise during cancer evolution and selection of mutations that permit continued proliferation and survival.

1.3.1 The Rous Sarcoma Virus

The story begins with the elucidation of an avian retroviral oncogene prompted by the studies of Peyton Rous in the early 1900s at the Rockefeller Institute (now the Rockefeller University) in New York City. Interested in avian cancer, Rous was given a chicken harboring a sarcoma by an upstate chicken farmer who had read of his research at Rockefeller. Rous excised the tumor, then ground and filtered it to remove the cartilaginous residue. He found that upon injecting the soluble filtrate into certain strains of tumor-free chickens a sarcoma would develop. This represented a major breakthrough, demonstrating for the first time that this form of cancer was transmissible in chickens.

1.3.2 Proto-oncogenes and Oncogenes

Following many years of intense research by numerous investigators, the transmissible agent was identified as the (appropriately named) Rous sarcoma virus (*RSV*). Of interest to us in this section was the recognition that the oncogenic element in the retroviral genome was a mutated version of a highly conserved and essential gene in human cells, *SRC*. This gene encodes a tyrosine kinase that functions in a cellular growth pathway; the mutation of the gene in the retroviral genome renders the gene product constitutively active, thus the explanation for tumorigenicity in infected chickens. It appears that during a cycle of infection some time ago, *RSV* commandeered the normal cellular *SRC* gene, i.e. a proto-oncogene (also referred to as a cellular oncogene), from the infected bird and incorporated it into its genome. A subsequent mutation in the *SRC* gene was sufficient to render the protein constitutively active such that proliferation signaling occurred in the absence of proper growth signals. The mutation was responsible for the conversion of the proto-oncogene to an oncogene. To date, over 30 retroviral oncogenes have been identified, most of them being in rodent and avian viruses (Vogt 2012). [*N.B. While we do not know the exact constituents of the filtrate that were injected into the chickens, it surely contained some macromolecular constituents and probably cells associated with the sarcoma. Later studies demonstrated, however, that it was the presence of the viral SRC gene that produced the tumorigenicity*.]

In normal cellular function, the gene encoding almost any regulatory protein involved in cell growth or survival can undergo the proto-oncogene-to-oncogene conversion by certain mutations or amplification that result in constitutive activity promoting, say, cell division without the requirement of external or even internal growth signals. Even the growth factors or their receptors can be considered oncogenic if, for example, there are mutations in their genes that increase their expression. In addition to mutations, genomic rearrangements can also produce oncogenes if the proto-oncogene is translocated to make a fusion gene that is no longer regulated and possibly gives a constitutively active fusion protein. This is the basis of the Philadelphia chromosome that is responsible for many cases of chronic myelogenous leukemia (CML). With this definition one can consider hundreds of proto-oncogenes that have the potential to become oncogenes. Suffice it to say that the proto-oncogene-to-oncogene conversion does not necessarily lead to cancer; rather, it can be considered a signal.

1.3.3 Conversion of Proto-oncogenes to Oncogenes

There are several genetic alterations that convert proto-oncogenes to oncogenes, most of which were mentioned or alluded to above.

Mutations, often single base changes (point mutations), leading to gain-of-function of positive regulators of the cell cycle, e.g., growth factor receptors or *SRC*.

Chromosomal instability such as loss of portions of chromosomes or rearrangements, e.g., inversions, translocations, deletions and insertions, resulting in a gain-of-function of positive regulators. An example of such a translocation is the fusion of the *ABL* gene on chromosome 9 to chromosome 22 where it is fused to the *BCR* gene yielding a fusion protein of *BCR-ABL* where *ABL*, normally highly regulated, exhibits constitutive activity.

Gene amplification resulting in abnormally high expression of growth factor receptors (or growth factors) that function in a pathway leading to cell division, e.g. the *HER2* receptor in breast cancer.

Viral infection/insertion may also contribute to some forms of cancer, e.g., the human papilloma virus (HPV) and cervical cancer.

1.3.4 Tumor Suppressor Genes

We now turn our attention to the topic of tumor suppressor genes. These genes and their protein products refer to ones that function to prevent the progression of the cell cycle if conditions at some checkpoints are not met, e.g., DNA damage is detected and not repaired. For a tumor suppressor gene to lose its function, it

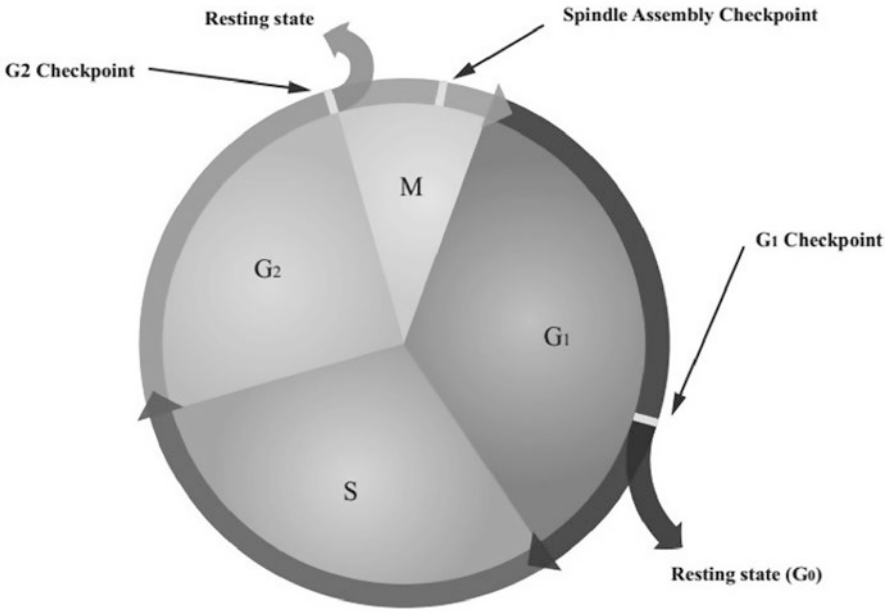

Fig. 1.1 A schematic view of the cell cycle showing the resting state (G_o), the first gap phase (G_1), the synthesis phase (S), the second gap phase (G_2) and the phase where mitosis occurs (M). A complete cycle can require some 18–24 h, although some cancer cells complete the cell cycle in less time

requires the loss of both copies of the gene while the loss of one copy increases the risk of cancer development. Within this class are the familiar *BRAC1* and *BRAC2* genes involved in a familial form of breast and ovarian cancers, and the *APC* (adenomatous polyposis coli) gene responsible for most cases of familial colorectal cancer. Of the many other tumor suppressor genes in the human genome, we will next discuss two well-studied examples that function in the cell cycle (Fig. 1.1).

Some cells in the body are dividing frequently, but most are in a resting or quiescent state denoted as G_o. A signal for cell division, such as growth factors initiating an intracellular signaling cascade or even the presence of a constitutively active oncogene in the signaling pathway, begins a process that takes the cells from the quiescent state to the first gap phase (G_1). Cyclin D family members are expressed and the proteins interact with cyclin-dependent kinase (*CDK*) complexes. The retinoblastoma (*RB*) gene, a tumor suppressor gene, encodes a nuclear protein that is a negative regulator of cell division, constantly maintaining cells in G_1 provided it is associated with another nuclear protein, the transcription factor *E2F*. The action of the *CDK* complex is to hyperphosphorylate *RB*, leading to dissociation of the *RB-E2F* complex. Freed of the inhibitory effects of *RB*, *E2F* acts to up-regulate itself, another cyclin, and enzymes required to carry out replication of the genomic DNA. These events, along with others, will lead to the progression from G_1 into the

S or synthesis phase of the cell cycle where DNA synthesis occurs. This progression is limited, however, by another protein *P53* (or TP53, tumor protein 53) that oversees DNA fidelity, along with other roles to be discussed later. Among its many actions *P53* can induce the activation of genes for DNA repair, cause cell cycle arrest or send the cell into apoptosis if DNA repair is not successful. Successful entry into the S phase results in the replication of ~3 billion base pairs of DNA accomplished with a variety of enzymes including ones capable of proofreading and repair of errors. As summarized, estimates suggest that the mutation rate in cells is some 10^{-12} to 10^{-9} per nucleotide in each cell division and that, of the 10^{14} cells comprising the average human, there are about 10^{16} division cycles during a lifetime (Duesberg 1987; Loeb 1989). Needless to say, these estimates are dependent on many factors and assumptions, and a wide range can be found in the literature. Although not discussed here, there are checkpoints and inhibitors at the $S \rightarrow G_2$ and the $G_2 \rightarrow M$ boundaries (G_2 is the second gap phase and M is mitosis, i.e., cell division).

RB and *P53* are, in effect, gatekeepers that prevent cells from dividing unless signaling pathways for growth impact on the nucleus and the cells have high fidelity, e.g., no damaged DNA. These negative regulators of cell division are thus critical components of cell oversight. From this perspective it is not surprising that mutations interfering with *RB* and *P53* functions could allow continual cell division and transcription of faulty DNA. Suffice it to say that the *RB* and/or *P53* genes are frequently mutated in various cancers.

1.4 Emerging Results on Cancer Genomes, Tumor Heterogeneity and Cancer Evolution

The emergence of rapid deep sequencing technology has provided an unprecedented opportunity to sequence large numbers of cancers for comparison with DNA sequences obtained from normal controls. In an interesting twist of fate, DNA sequencing is no longer the rate-limiting step in cancer genomics. Rather, it is the ability to analyze the copious amounts of data that are forthcoming from many laboratories and factory-like sequencing centers. From this perspective, the timing is good for bioinformaticians to enter cancer research with the possibility of adding substantively to our knowledge relating genomic changes to phenotypic changes in cancer patients.

1.4.1 What Is Being Learned?

The results from cancer genome sequencing are providing considerable information on mutations and the myriad other genomic changes, e.g., chromosomal gains, losses and rearrangements, present in most cancers (Stratton et al. 2009;

Pleasance et al. 2010; Garraway and Lander 2013; Alexandrov and Stratton 2014; The Cancer Genome Atlas Research Network 2011a, b, 2012a, b, c, 2013a, b, c, 2014; Alexandrov et al. 2013; Kandoth et al. 2013; Vogelstein et al. 2013). In one study 3,281 tumors from 12 different types of cancer (11 solid tumors plus acute myeloid leukemia) were analyzed for point mutations and small insertions and deletions (Kandoth et al. 2013). In this sampling 617,354 somatic mutations were identified: 398,750 missense; 145,488 silent and smaller numbers each of nonsense, splice site, non-coding RNA, non-stop read-through, frame-shift insertions/deletions (indels) and in-frame indels. *P53* was found to be the one most frequently mutated, and the lipid kinase gene, *PIK3CA* (phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha), was the second. Not surprisingly, many mutations appeared in genes encoding transcription factors; cell cycle regulators; signaling pathways, including receptor tyrosine kinase, *MAPK*, *PI3K*, *TGFβ* and *WNT*/β-catenin; and ECM related genes as detailed in Chap. 4.

Genome sequencing has also provided some surprising observations, but ones that are consistent with the emerging view that cancer is not just a disease of the genome. Sequencing studies by two labs (Mack et al. 2014; Parker et al. 2014) on three subtypes of ependymoma brain tumor found the following. One subtype had an intrachromosomal translocation yielding what appears to be a 'driver mutation' for cancer, and another subtype had abnormal epigenetic alterations. Of particular interest, however, was the finding that another subtype was devoid of gene mutations and aberrant epigenetic changes. These results emphasize the complexity of cancer and importantly the role of non-genomic changes driving cancer formation.

1.4.2 Driver and Passenger Mutations

In recent years there has been considerable interest in identifying the 'driver' mutations and separating them from the 'passenger' mutations. Although as discussed later (Chap. 5), there is a current movement to consider those crucial mutations as ones that were selected as necessary to maintain proliferation and survival of the developing cancer cell(s) and may not necessarily be causal to cancer. This information will of course direct many of the treatment modalities for specific cancers. Many mutations, particularly in older individuals, are known to exist before the occurrence of cancer and are believed to have nothing to do with the onset or continuation of cancer (Tomasetti et al. 2013). These innocuous mutations arise from the high number of cell divisions and the inherent errors that occur in proofreading and repair, as well as mutations from environmental causes that do not produce 'drivers' of cancer. One estimate is that there are about 140 genes that, with appropriate mutations, can become drivers (Vogelstein et al. 2013).

1.4.3 Major Findings

Some of the major findings from genomic sequencing of various tumors have been delineated and provide much insight into, if not cancer initiation, then at least its progression (Vogelstein et al. 2013). Some of these principles are listed below; however, we have qualified them as more likely being responsible, through natural selection, for unlimited growth and survival, not necessarily causal.

a. Solid tumors have an average of 33–66 somatic nonsynonymous mutations, predominantly single-base changes that are expected to alter the resulting proteins; however, a limited number of mutations are capable of sustaining cancer proliferation and survival. [*N.B. Vogelstein et al. (*2013*) claim that the majority of human cancers result from two to eight sequential mutations occurring over 20–30 years, each of which confers about a 0.4 % growth advantage*.]
b. There are about 140 such genes that if mutated can contribute to cancer, either via initiation, proliferation or survival.
c. Three cellular processes are regulated by these essential genes: cell fate determination, cell survival and genome maintenance.
d. Although the pathways altered by key mutations in different tumors are similar, each individual tumor is distinct.
e. Heterogeneity exists in the cells of tumors, and this can affect therapeutic effectiveness.

1.4.4 Metastasis

There is considerable interest in the delineation of the various changes that can drive a primary tumor to metastasis (see Chaps. 10 and 11), the cause of over 90 % of cancer mortality (Irmisch and Huelsken 2013). This is an important aspect of cancer research that bioinformatics can address when more data are available from the ongoing sequencing of cancer genomes and transcriptomes. In addition to the genetic changes referred to above, many alterations in metabolism, hypoxia, and other cellular processes exist that tend to drive cancer growth. These are covered in depth in Chap. 10.

1.4.5 Cancer Heterogeneity

It is important that the role of cancer heterogeneity be pointed out. Pathologists and clinicians have known for years that solid tumors are heterogeneous with regard to cellular morphology and patient responses to treatment. Thus, cancer heterogeneity, first proposed several decades ago (Nowell 1976), is an important aspect of cancer and an area that is being addressed (Meacham and Morrison 2013;

Burrell et al. 2013; Vogelstein et al. 2013). It is now appreciated that considerable heterogeneity exists in any given cancer, both at the molecular and cellular level. Cancer is clearly many diseases, and even individual tumors within similar types of cancer may be unique. A given tumor is likely composed of a dominant clone and several subclones, each of which may grow at different rates and respond differently to treatment(s). This intratumor heterogeneity impacts on the evolution of cancer and the natural selection of clones more favorable for sustained growth, survival and ability to colonize distant sites (extravasation and metastasis). Cancer heterogeneity, evolution and natural selection are emerging as significant features in our understanding of cancer growth and control (Klein 2013; Burrell and Swanton 2014; Lawrence et al. 2013) and are areas in which bioinformatics can provide considerable insight as more data become available.

1.5 An Early Sequential Model of Cancer Development

One of the early models to explain the development of cancer strictly from genetic changes is referred to as a "sequential model" based on a series of mutations. It is now well recognized that this model may be overly simplistic, but it is presented to introduce the concept and several genes that, when mutated, can function to aid propagation of cancer. Based on extensive studies of benign and malignant colorectal cancer (hereafter referred to simply as colon cancer), Fearson and Vogelstein proposed a sequential pathway for the development of malignancy, a pathway often referred to as the canonical pathway (Fearson and Vogelstein 1990). Colon cancer can be categorized into two forms, sporadic and familial, having respective frequencies of about 80 and 20 %. Sporadic colon cancer can be further divided into a form arising from mutations and/or chromosomal instability and a form attributable to microsatellite instability. These two forms of sporadic colon cancer exhibit frequencies of approximately 80–85 % and 15–20 %, respectively.

The canonical pathway was proposed to arise from a sequential or linear set of genetic alterations. In the majority of cases the *APC* gene, located on chromosome 5q, was found to undergo a mutation that reduced or abolished its activity and contributed to the formation of a benign lesion or early adenoma. While the protein encoded by *APC* has a number of biological actions including its role in the *WNT* pathway, cell adhesion (via E-cadherin), mitosis and cytoskeletal regulation, it is the former that has attracted most attention. Forming part of a complex with glycogen synthase kinase-3β and axin, loss of *APC* activity results in β-catenin escaping degradation and thus constitutively activating the *WNT* pathway that regulates numerous genes, some of which are involved in the cell cycle. We mention in passing that mutations in *APC* have been identified in many if not most cases of familial colon cancer.

Mutations to the oncogene *KRAS* on chromosome 12p12.1 have been frequently identified in intermediate adenomas, mutations that result in producing a defective GTP-binding protein that is involved in the mitogen-activated protein kinase

(*MAPK*) and other growth-promoting pathways. Again, constitutive activation of the protein *KRAS*, e.g., by inactivating the intrinsic GTPase activity that converts GTP to GDP leading to *KRAS* inactivation, results in a constant enhancement of growth-promoting pathways as well as a loss of cell polarity that could reduce cell adhesion.

Other allelic losses (via mutations or chromosomal loss) are often found in late adenomas and carcinomas with the *P53* and *DCC* (deleted in colorectal cancer) genes on chromosomes 17p13.1 and 18q21.1, respectively. *P53* is considered a tumor suppressor gene and serves as a gatekeeper for cells exiting G_1 to the S phase of the cell cycle. It promotes repair of damaged DNA, e.g. from errors in replication or environmental stress, and if repair is not successful terminates cell cycle progression and leads to apoptosis. Clearly, inactivation of this key cell cycle regulator could have major detrimental effects on the normal fidelity expected in the cell cycle. The *DCC* protein is a transmembrane protein that serves as a receptor for proteins involved in regulating axon guidance in the nervous system and also seems to participate in cell motility, signaling and overcoming apoptosis.

Although attractive in its simplicity, the sequential model is now believed by many investigators to function more as developing cancer cells are undergoing an "evolution and natural selection" phase to obtain a genomic background that perpetuates growth and evades apoptosis and immune destruction.

1.6 Epigenetics and Cancer

Most of the research on cancer has heretofore dealt with the role of genetic changes, i.e. alterations in the sequence of DNA, that lead to changes in normal cellular functions that regulate proliferation, survival, angiogenesis, metastasis and others. Recent studies have, however, documented that epigenetic changes are also important in the initiation and progression of cancer (Beck et al. 2012; Shen and Laird 2013; Timp and Feinberg 2013; Waldman and Schneider 2013; Suva et al. 2013). Such changes are attributable to modifications of chromatin and chromatin packaging, as emphasized by the appearance of mutations in genes involved in DNA methylation, histone modification and chromatin remodeling, with a number of mutations found to be tumor-specific.

Composed of nucleic acids and proteins, there are potentially many possibilities for epigenomic changes in chromatin. The protein core around which genomic DNA is wrapped is composed of a histone octomer with two copies each of four distinct histones, forming a nucleosome; these, in turn, form a helical arrangement. As summarized (Shen and Laird 2013), there are multiple sites for alterations that control the level of transcriptional activity, including DNA methylation, histone modifications and variants, interacting proteins, noncoding RNAs and nucleosome positioning.

Certainly one of the more prevalent alterations is that of DNA methylation, catalyzed and maintained by several DNA methyltransferases yielding primarily

cytosine-5 methylation of CpG dinucleotides. Such enzymes can be considered as 'writers' since they in effect make the epigenetic mark. Another chemical change is that of histone modification, including methylation, acetylation and phosphorylation. The histone modifications are catalyzed by histone methyltransferases and demethylases, acetyltransferases and deacetylases, and kinases and phosphorylases. Those enzymes that remove the covalent tag are referred to as 'erasers'. Also, numerous histone variants have been identified. Another mode by which the epigenome can be altered is that of nucleosome positioning and remodeling, accomplished by sequence-specific binding proteins. These are important in selecting the form of chromatin, euchromatin (open form) and heterochromatin (a more closed form), thus enhancing or inhibiting the availability of readers, writers, erasers and other chromatin-binding proteins.

From this abbreviated overview, mentioning most but certainly not all of the factors responsible for defining and maintaining the epigenome, sequencing and functional studies have shown that mutations can occur in essentially all of the genes required to form and maintain the epigenome (Fullgrabe et al. 2011; Shen and Laird 2013; Timp and Feinberg 2013). Importantly, many of these mutations have been documented to be related, or at least correlated, at one level or another to tumorigenesis. This area will undoubtedly emerge as an important component of cancer as more results become available.

1.7 Cancer Cell Metabolism

1.7.1 Meeting the Energetic Requirement of Cells

Of the many types of foods ingested, the body uses three major classes as fuels for its energetic needs: carbohydrates, lipids (fats) and proteins. The chemical compositions and structures of these vastly different biomolecules, ranging from simple sugars to complex polysaccharides, fatty acids to triacylglycerols (triglycerides) and peptides to high molecular weight proteins. Yet, many of the different metabolic pathways converge at a common intermediate, acetyl-coenzyme A (acetyl-CoA) or a downstream intermediate, leading to the biosynthesis of adenosine triphosphate (ATP), an important source for cellular energetic needs.

The average sedentary adult requires about 2,000 Calories (Cal) per day to meet the normal requirements to maintain overall homeostasis, i.e., for heart, brain, lung, kidney and other organs to function. This daily requirement for any given sedentary individual can vary as much as ±400 Cal since it is influenced by age, gender and metabolic factors. Over 80 kg of ATP are required to meet this daily basal caloric need; however, the body contains only about 0.25 kg (Tymoczko et al. 2013). Thus, ATP is constantly being utilized and resynthesized to meet daily needs. For someone who is physically active, the caloric requirement rises dramatically, and consequently ATP biosynthesis must increase as well. [*N.B. The calorie, more specifically the gram or small calorie, is defined as the energy required to*

increase the temperature of 1 g of water 1 °C at standard atmospheric pressure (this corresponds to about 4.2 J). Biochemists and nutritionists, on the other hand, use a "large" or "kilogram" calorie, i.e. the Cal, that is equivalent to 1,000 "small" calories (about 4.2 kJ).]

Metabolism of carbohydrates, proteins and lipids yields approximately 4 Cal/g, 4 Cal/g and 9 Cal/g, respectively. Thus, per unit weight ingested, lipids provide more than twice the Cal (or energy) than do carbohydrates and proteins; however, we rely on all three for energetic needs, particularly lipids and carbohydrates (and certainly circulating glucose for minute-to-minute cellular needs). Glucose can be metabolized anaerobically (absence of oxygen) to yield small amounts of ATP and aerobically (presence of oxygen) to obtain greater amounts. Fatty acids, most being obtained from lipolysis of triacylglycerols, can undergo β-oxidation, giving acetyl-CoA that can enter the TCA cycle for ATP production, and glycerol that can enter the hepatic gluconeogenic pathway and be converted to glucose for metabolism. Proteins are constantly turning over, and some of the amino acids so derived can serve as precursors for glucose synthesis (gluconeogenesis) or for synthesis of pyruvate or intermediates in the TCA cycle (see below).

Cancer cells also utilize carbohydrates, lipids and proteins to generate the ATP that is required to meet the energetic needs for proliferation, metastasis and survival. In order to appreciate the metabolic derangements unique to cancer, it is first necessary to understand, even if superficially, the pathways of normal metabolism, which are briefly treated in the following section with reference to the changes that occur in cancer.

1.7.2 Glucose Metabolism

Glucose metabolism will be discussed first since it is quite distinct in cancer cells compared to normal cells. For both normal and cancer cells, circulating glucose in the bloodstream enters cells via one or more glucose transporter proteins (*GLUT*s) and is then rapidly phosphorylated to glucose 6-phosphate by either of two ATP-dependent enzymes, hexokinase or glucokinase, thus ensuring retention within the cell (Reaction 1.1).

$$\text{glucose} + \text{ATP} \rightarrow \text{glucose}\,6-\text{phosphate} + \text{ADP} + \text{H}^{+} \tag{1.1}$$

There are three metabolic paths for glucose 6-phosphate within the cell. If the cell does not require ATP, glucose 6-phosphate can be metabolized to a high molecular weight polysaccharide of repeating glucose units, glycogen (Reaction 1.2). The three subsequent reactions are catalyzed by the enzymes phosphoglucomutase, UDP-glucose pyrophosphorylase and glycogen synthase, respectively, where UDP is uridine-diphosphate.

$$\text{glucose}\,6-\text{phosphate} \rightarrow \text{glucose}\,1-\text{phosphate} \rightarrow \text{UDP}-\text{glucose} \rightarrow \text{glycogen} \tag{1.2}$$

The two other metabolic paths for glucose 6-phosphate are of most interest here: the glycolytic pathway and the pentose phosphate pathway. Glycolysis represents a series of enzymatic reactions that convert the phosphorylated 6-carbon glucose into two molecules of pyruvate, a 3-carbon product. The pathway yields, in addition to two molecules of pyruvate, two molecules each of ATP, NADH and other reaction products for each molecule of glucose entering the pathway (Reaction 1.3 and Fig. 1.2).

$$\text{glucose} + 2\text{ADP} + 2\text{NAD}^+ + 2\text{P}_\text{i} \rightarrow 2\text{pyruvate} + 2\text{ATP} + 2\text{NADH} + 2\text{H}^+ + 2\text{H}_2\text{O} \quad (1.3)$$

Pyruvate can be converted to acetyl-CoA (Reaction 1.4), the main substrate for the TCA cycle, to lactate (Reaction 1.5) or to oxaloacetate (Reaction 1.6). The enzymes catalyzing these reactions are, respectively, pyruvate dehydrogenase, lactate dehydrogenase and pyruvate carboxylase.

$$\text{pyruvate} + \text{CoA} + \text{NAD}^+ \rightarrow \text{acetyl} - \text{CoA} + \text{CO}_2 + \text{NADH} + \text{H}^+ \quad (1.4)$$

$$\text{pyruvate} + \text{NADH} + \text{H}^+ \leftrightarrow \text{lactate} + \text{NAD}^+ \quad (1.5)$$

$$\text{pyruvate} + \text{CO}_2 + \text{ATP} + \text{H}_2\text{O} \rightarrow \text{oxaloacetate} + \text{ADP} + \text{P}_\text{i} + 2\text{H}^+ \quad (1.6)$$

The other metabolic route for glucose 6-phosphate is that of the pentose phosphate pathway (Fig. 1.3). This pathway consists of an oxidative phase in which glucose 6-phosphate is converted to ribulose 5-phosphate by several enzymes acting sequentially, glucose 6-phosphate dehydrogenase, lactonase and 6-phosphogluconate dehydrogenase (Reaction 1.7). This is an important reaction since it regenerates NADPH and associated reducing power. The second phase is a complex oxidative component consisting of a number of enzymes that yields ribose 5-phosphate (Reaction 1.8, catalyzed by phosphopentose isomerase), fructose 6-phosphate and glyceraldehyde 3-phosphate. Of the three pentose phosphates, ribose 5-phosphate (a 5-carbon sugar phosphate) is needed for the synthesis of nucleic acids, and the other two, fructose 6-phosphate and glyceraldehyde 3-phosphate, can enter as intermediates in the glycolytic pathway.

$$\text{Glucose}\,6-\text{phosphate} + 2\text{NADP}^+ + \text{H}_2\text{O} \rightarrow \text{ribulose}\,5-\text{phosphate} + 2\text{NADPH} + 2\text{H}^+ + \text{CO}_2 \quad (1.7)$$

$$\text{Ribulose}\,5-\text{phosphate} \leftrightarrow \text{ribose}\,5-\text{phosphate} \quad (1.8)$$

From the point of entry of glucose into cells, the glycolytic pathway is composed of ten enzymatic reactions, all occurring in the cell cytoplasm, to give the 3-carbon product pyruvate. Pyruvate, in turn, can undergo one of several enzymatically catalyzed steps with its conversion to either of the following. (1) lactate: This reaction reduces pyruvate and occurs independent of the availability of oxygen; it is catalyzed

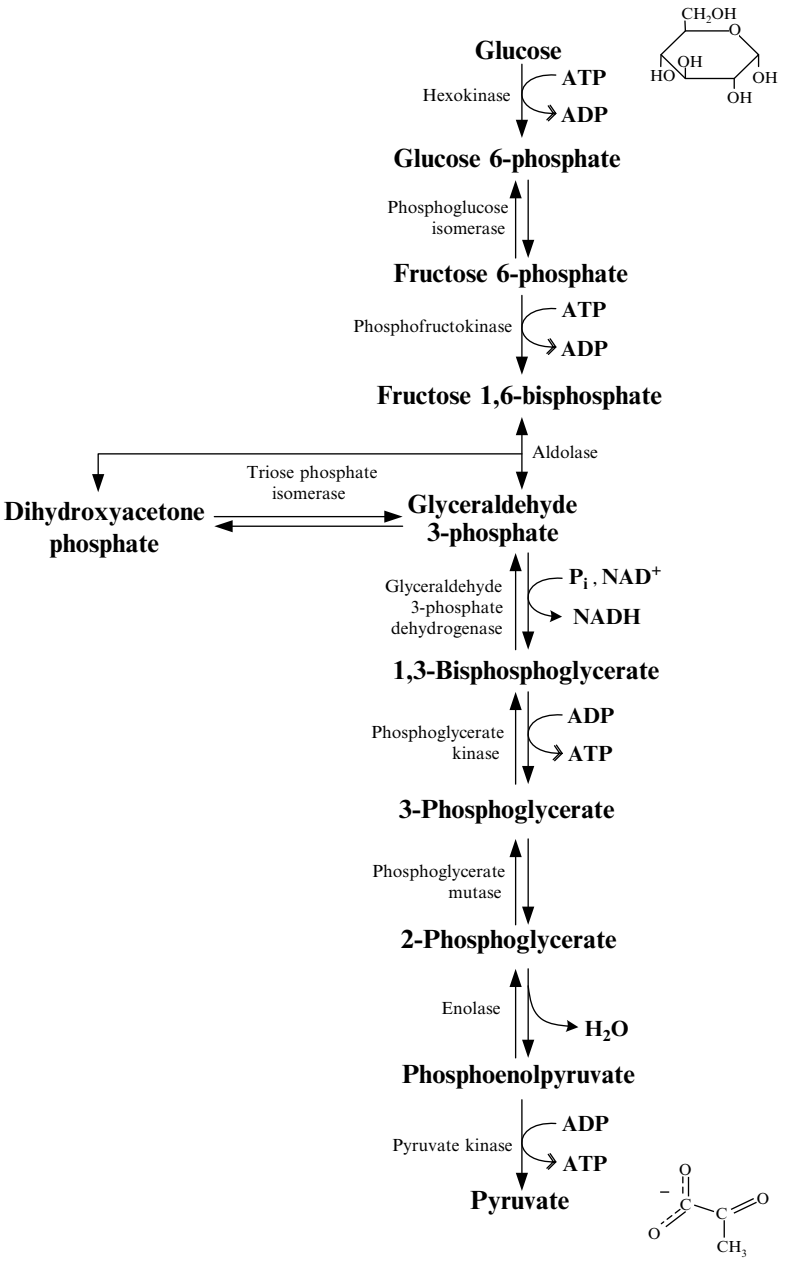

Fig. 1.2 The glycolytic pathway from the entry of glucose into a cell and the subsequent reactions that yield pyruvate. Note the conversion of the 6-carbon structure, glucose, into two molecules of pyruvate

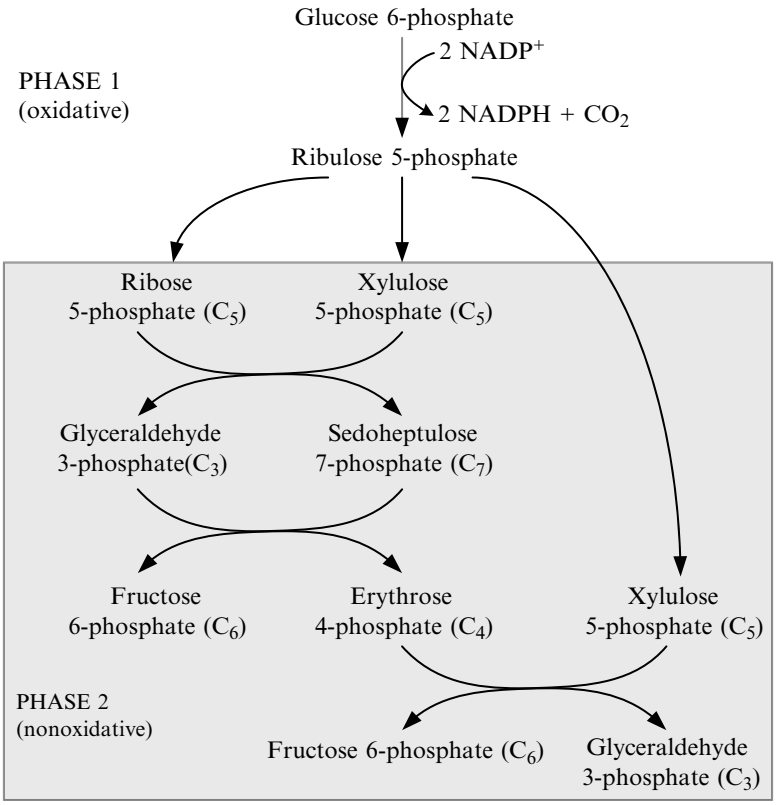

Fig. 1.3 The pentose phosphate pathway showing the conversion of glucose 6-phosphate to three pentose phosphates, ribose 5-phosphate, fructose 6-phosphate and glyceraldehyde 3-phosphate

by the enzyme lactate dehydrogenase. (2) oxaloacetate: This metabolite, obtained from the carboxylation of pyruvate by the enzyme pyruvate carboxylase, is an intermediate in the TCA cycle and also a precursor for the synthesis of glucose via a metabolic pathway, gluconeogenesis (the synthesis of glucose from non-glucose precursors). (3) acetyl-CoA: In the presence of oxygen the enzyme pyruvate dehydrogenase, in an oxidation reaction, catalyzes the conversion to acetyl-CoA, releasing CO_2, of which the body must rid itself, and reducing NAD^+ to NADH, thus the enzyme is catalyzing an overall oxidation-reduction reaction. Acetyl-CoA is an important intermediate for several pathways and serves as a convergent point for metabolism of carbohydrates, lipids and proteins.

Pertinent to our discussion, one major fate of acetyl-CoA is its entry into the TCA cycle, located in the mitochondrion and composed of eight enzymes (Fig. 1.4). This is an important component of metabolism, particularly when energy is needed and available to the cells. The entering acetyl group on acetyl-CoA is oxidized, i.e., loses electrons, and forms CO_2 (two molecules for each acetyl-CoA entering the pathway) in a series of oxidation-reduction reactions that

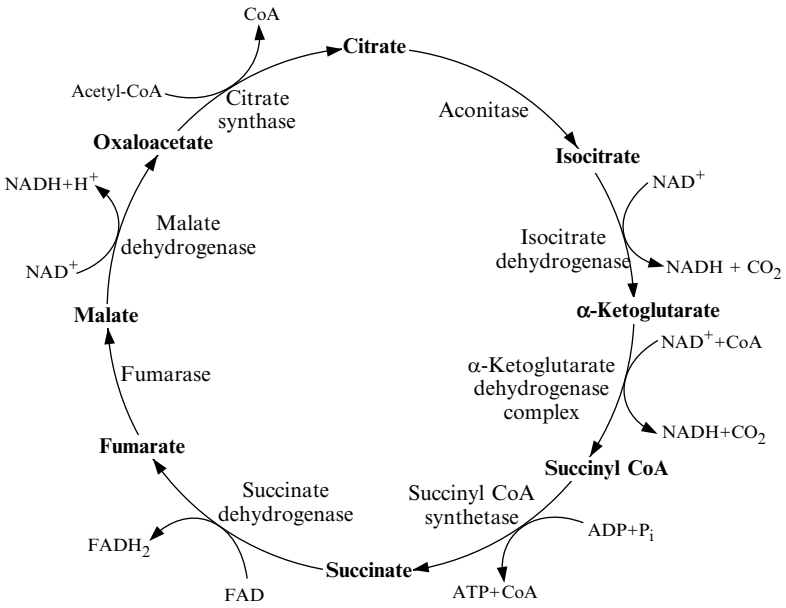

Fig. 1.4 The TCA cycle showing the entry of acetyl-CoA

generate high energy electrons from carbon sources. Continuing the mitochondrial reactions, oxidative phosphorylation refers to an important series of complex reactions that culminate in the synthesis of ATP. The reactions allow electrons from NADH and $FADH_2$ to transfer to oxygen (O_2) that is converted to water in the presence of hydrogen ions. This flow of electrons pumps protons into the region between the inner and outer mitochondrial membranes from the mitochondrial matrix (cf. the schematic in Fig. 1.5). The proton gradient is responsible for driving the synthesis of ATP from ADP and P_i, a reaction catalyzed by ATP synthase. Accounting for the ATP required to transport NADH into the organelle, a net production of 30 molecules of ATP for each molecule of glucose metabolized is realized. As mentioned earlier, even sedentary individuals require some 2,000 Cal per day, an amount that can increase significantly during vigorous exercise, and this requires some 80 kg of ATP biosynthesized per day, most of this from *de novo* synthesis by recycling ADP into ATP.

1.7.3 The Warburg Effect and Other Metabolic Alterations in Cancer

Working at the Kaiser Wilhelm Institute in Berlin, now the Max Planck Institute, Otto Warburg made the significant observation in the 1920s that cancer cells utilize more glucose than normal cells (Koppenol et al. 2011). Further, it was found that

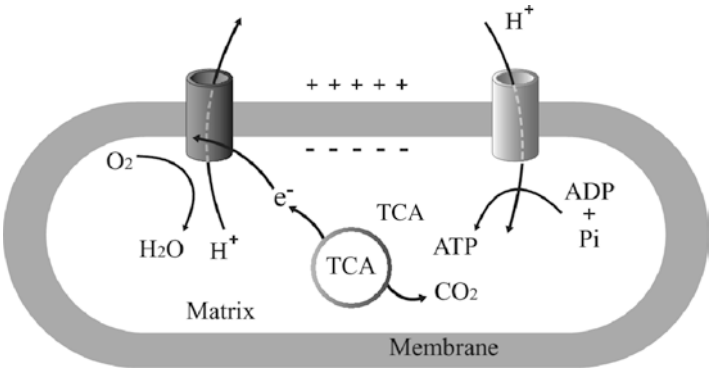

Fig. 1.5 A simplified and schematic representation of oxidative phosphorylation leading to the synthesis of ATP. The electron transport chain, shown as a *dark cylinder*, is responsible for transferring electrons from NADH and $FADH_2$ to oxygen and creating a proton gradient. The energy of the proton gradient that is used by the enzyme ATP synthase, indicated by a *gray cylinder*, to drive the synthesis of ATP from ADP and P_i. Adapted from (Tymoczko et al. 2013)

glucose was converted to lactate (lactic acid) via glycolysis (see Fig. 1.2 and Reaction 1.5). The confounding aspect of this finding, however, was that the increased level of glycolysis occurred even in the presence of oxygen. Under these conditions, i.e., ample oxygen, one expects aerobic respiration in which glucose is directed to pyruvate which is then converted to acetyl-CoA, not lactate.

Another surprising observation by Warburg was that aerobic respiration was like that of normal tissues, but it failed to prevent lactate formation. This is in contrast to aerobic metabolism in general since the well-known and accepted Pasteur effect leads to a reduction in lactate production in the presence of oxygen. In spite of these results, Warburg nonetheless believed that the pathway of aerobic respiration was damaged; it is now known however that it is the regulation of glycolysis that differs from normal cells in cancer cells. Warburg's experiments were initially conducted using thin slices of Flexner-Jobling rat liver carcinoma, and they were later confirmed with a number of human carcinomas.

The singular finding of enhanced anaerobic glycolysis in the presence of oxygen has led to numerous investigations in the subsequent years. Although a full explanation of the Warburg effect and its ramifications in cancer are still unfolding, a number of recent investigations have yielded many exciting and provocative observations that offer some rationale of why a less efficient energy-generating pathway, anaerobic metabolism, may be preferred over the more efficient aerobic respiration (Ferreira 2010; Cairns et al. 2011; Dang 2012; Bensinger and Christofk 2012; Icard and Lincet 2012; Oermann et al. 2012; Soga 2013). Moreover, elucidation of the distinctions between normal and cancer cell metabolism provides potentially new avenues to explore for therapeutic regimens (Jang et al. 2013). [*N.B. As a side note, the Warburg effect forms the basis of imaging by means of positron emission tomography (PET) in which patients receive 2-fluoro-2-deoxy-*d*-glucose (FDG), a radiolabeled (^{18}F) and non-metabolizable form of glucose that becomes concentrated in*

cancer cells at a higher level than in normal cells thus enabling imaging to occur. Tumors in highly metabolically active tissues such as liver and brain, however, are often difficult to detect because of the high background level.]

Through various mechanisms, lactate itself has been found to enhance angiogenesis, cell migration and escape from immune surveillance. Also, the increased lactate production reduces pericellular pH resulting in the activation of apoptosis in neighboring normal cells, the protection of the cancer cells by inhibition of the immune system and an elevation of a number of proteases, including metalloproteinases that can facilitate escape of tumor cells from their local environment, a requirement for metastasis to occur. In addition, the increased uptake of glucose, and hence the amount of glucose 6-phosphate, ensures a plentiful supply of substrate for the pentose phosphate pathway (cf. Fig. 1.3 and Reactions 1.7 and 1.8), products of which can be converted to nucleotides for nucleic acid synthesis or serve as intermediates to the glycolytic pathway.

Reinforcing the importance of metabolic changes in cancer, an exome sequencing study (175,471 exons from 20,661 genes) uncovered recurring mutations in the *IDH1* gene (Parsons et al. 2008). The protein encoded by *IDH1* is isocitrate dehydrogenase, an enzyme that converts isocitrate to α-ketoglutarate in the TCA cycle (cf. Fig. 1.4). Subsequent research by a number of investigators studying different cancers, reviewed by Garraway and Lander (Garraway and Lander 2013), showed that the mutations in *IDH1* led to gain-of-function in the enzyme and that, moreover, the enzyme product was an enantiomer of 2-hydroxyglutarate. This unexpected metabolite was found to inhibit α-ketoglutarate-dependent enzymes, including prolyl-4-hyroxylases that are important in regulating hypoxia inducible factor (*HIF*). Such *IDH1* mutations, surprisingly, correlated with the CpG island methylator phenotype; further, *IDH1* and *IDH2* (the mitochondrial homolog) mutations were found to be mutually exclusive with *TET2* mutations, the gene product being a methylcytosine dioxygenase that catalyzes methylcytosine to 5-hydroxymethylcytosine in DNA. Such unexpected observations and correlations reinforce the importance of metabolic alterations in cancer and emphasize the need for careful bioinformatic approaches when comparing large datasets; totally unexpected and potentially important new information can be forthcoming.

Another player that has emerged is the amino acid glutamine, and of interest is a role of the oncogene *MYC*, as well as other oncogenes and tumor suppressors, in regulating glutamine metabolism. Glutamine can function as a carbon source in the process of energy production; it can also regulate redox homeostasis, in large part through its role in the biosynthesis of the antioxidant glutathione. Lastly, glutamine can supply carbon and nitrogen to a number of cellular reactions. Regarding glutamine's role in energy production, the enzyme glutaminase is responsible for the conversion of glutamine to glutamate, the latter of which can be converted to α-ketogluterate that is an integral part of the TCA cycle (Fig. 1.3). This is particularly important in proliferating cells since citrate, another integral component of the TCA cycle (Fig. 1.3), is transported from the mitochondria to contribute to the synthesis of acetyl-CoA for lipid biosynthesis (Icard et al. 2012). Many other cellular proteins, including enzymes, oncogenes and tumor suppressors, are emerging as

having important roles in cancer metabolism (Chen and Russo 2012; Oermann et al. 2012). While most of these are not discussed further, it is expected that they and others, as well as presently unknown regulators and processes, will materialize as important contributors to the altered metabolic status of cancer cells.

Hypoxia-inducible factor (HIF), notably *HIF1* (discussed in greater detail in Sect. 1.8), can escape its normal degradation under normoxic conditions due to mutations in certain enzyme-encoding genes, e.g., succinate dehydrogenase, fumarate hydratase, or prolyl hydroxylases or tumor suppressor proteins, e.g., von Hippel-Lindau (*VHL*), as well as higher cellular levels of metabolic intermediates, e.g., lactate, oxaloacetate and pyruvate (Cairns et al. 2011). The presence of *HIF1* under normal concentrations alters the expression of a number of genes participating in glycolysis, such as those for phosphofructokinase, hexokinase-II, pyruvate kinase M2, lactate dehydrogenase-A and glucose transporters, thus enhancing glycolysis; other genes are also affected that lead to reduced amounts of pyruvate entering the TCA cycle. *YC* and *HIF1* both activate the expression of the lactate dehydrogenase gene, *LDHA*, that favors the conversion of pyruvate to lactate; *MYC*, by suppressing two microRNAs (miR-23A and miR-23b), stimulates glutaminase gene expression resulting in a replenishment of intermediates in the TCA cycle (Oermann et al. 2012).

Another important metabolic component is that of the Ser/Thr kinase, AMP-activated protein kinase (*AMPK*), that serves to regulate metabolism and energy homeostasis. This regulatory kinase, depending upon the cellular conditions, can enhance or inhibit cancer cell growth (Faubert et al. 2014). In addition to these well documented changes, there are also other changes in cancer that impact on metabolism, but in the interest of brevity these will not be discussed.

Many years after the discovery of the Warburg effect, Warburg himself was still discussing the importance of mitochondrial alterations in giving a reduced ability of ATP synthesis via oxidative phosphorylation. Yet, more recent studies have shown unequivocally that cancer cells are not deficient in oxidative phosphorylation, at least for some cancers. On the other hand, some form of mitochondrial dysfunction or uncoupling has recently been noted. This involves elevated expression of certain uncoupling proteins (*UCP*s) that would lead to a reduction in effectiveness of the mitochondrial membrane potential. While not discussed herein this would result in a reduction in mitochondrial ATP synthesis, thus enhancing the cell's need for increased aerobic glycolysis.

Another aspect of the Warburg effect and mitochondrial function involves reactive oxygen species (ROS). An increase in oxidants such as ROS that are not countered by an increase in antioxidants, leads to oxidative stress in a cell. Since in mitochondrial respiration oxygen is the final acceptor for electrons in the formation of water, several ROS can arise: the superoxide radical ($^{\bullet}O_2^-$), the hydroxyl radical ($OH^{\bullet}$), and hydrogen peroxide (H_2O_2). These highly reactive species can damage all molecules, and proteins and DNA are particularly susceptible. There are enzymes to remove the free radicals, e.g., superoxide dismutase and catalase, but if ROS levels become too high, then cellular damage can occur. It is possible that the Warburg effect can reduce the level of ROS by increasing the amount of pyruvate produced

since pyruvate can scavenge peroxides that result from the action of superoxide dismutase; moreover, the pentose phosphate pathway generates NADPH that is required for the conversion of glutathione disulfide to glutathione, important in the inactivation of hyperoxide. Lastly, the mitochondrial uncoupling discussed above may reduce oxidative stress.

An additional component of metabolism was recently found with regard to ROS (Anastasiou et al. 2011). Cancer cells, like normal cells, must protect themselves from high concentrations of ROS. Proliferation of the transformed cells requires reducing power from NADPH to support the biosynthesis of nucleotides and lipids. NADPH also acts to maintain glutathione in the reduced state, necessary for homeostasis of ROS. This increased demand for NADPH, supplied in large part through the pentose phosphate pathway (Fig. 1.3), was found to be facilitated by ROS-mediated oxidation of a particular cysteine on an alternatively spliced form of the glycolytic enzyme, pyruvate kinase that functions to convert phosphoenolpyruvate to pyruvate. This alternatively spliced form is designated pyruvate kinase M2 (*PKM2*) and is expressed in many cancer cells. Oxidation of the cysteine leads to enzyme inactivation, thus diverting glucose metabolism into the pentose phosphate pathway. This metabolic switch helps ensure synthesis of adequate amounts of NADPH to meet the needs for cell proliferation and protection from excess ROS.

The Warburg effect leads to interesting and, at first, paradoxical effects on the pH of cancer cells. It seems reasonable to expect the intracellular pH to decrease with the higher levels of lactate (lactic acid) and other acidic intermediates in the glycolytic pathway being produced. Yet, the opposite occurs with the intracellular pH increasing from its normal value of approximately 7.2 to about 7.4 or even greater. While this may appear to be but a minor alteration, it nonetheless represents a significant decrease in the concentration of hydrogen ions. Conversely the extracellular pH, normally some 7.3–7.4, becomes acidified. This unusual reversal of hydrogen ion fluxes can be attributed to increased expression of plasma membrane-associated acid transporters such as H^+-ATPase, the Na^+-H^+ exchanger NHE1, and the H^+-monocarboxylate transporter, all of which lead to increased efflux of hydrogen ions from the cell interior into the extracellular milieu (Webb et al. 2011). The latter also transports lactate out of the cells. Cell surface carbonic anhydrases increase as well, these being enzymes that catalyze the important reaction by which carbon dioxide (CO_2) from respiring cells interacts with water to form carbonic acid (H_2CO_3); this in turn, forms bicarbonate (HCO_3^-) and a hydrogen ion (H^+) as shown below (Reaction 1.9).

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow HCO_3^- + H^+ \tag{1.9}$$

This simple reversible reaction can proceed non-enzymatically, but it is greatly accelerated by carbonic anhydrase. It shows how much of the carbon dioxide from respiring cells/tissues is converted to bicarbonate and how, in the lungs, carbon dioxide is formed that can be exhaled. In the vicinity of cancer cells overexpressing carbonic anhydrase there can be acidification from increased utilization of carbon dioxide, as well as the increase from hydrogen ions pumped from the cells, as further discussed in Chap. 8.

The small shift of the intracellular pH to one more alkaline can have profound effects on the cells. Numerous cellular pathways are altered by the pH change, in effect favoring cancer cell survival. For example, glycolysis, cell growth, and metastasis are enhanced while apoptosis is inhibited.

While this section has emphasized the alterations in carbohydrate metabolism, cancer cells also exhibit changes in other aspects of metabolism. For example, increased lipid biosynthesis often occurs in cancer (Yoshii et al. 2014), and lipids have been associated with maintaining redox potential in cancer cells, as well as enhancing tumor cell proliferation and survival (Santos and Schulze 2012). Altered amino acid metabolism and increased protein synthesis also accompany cancer development and growth. Recent studies have shown that *P53*, in addition to its known function as a tumor suppressor, is important in regulating glycolysis, oxidative phosphorylation, lipid metabolism, glutamine metabolism and ROS levels in non-transformed cells (Liang et al. 2013). Consequently, loss-of-function mutations in this gene can contribute significantly to the metabolic derangements in cancer,

In addition to the alterations in the cellular function of cancer cells, there are many other genes and pathways, some of which appear at first glance as being paradoxical, that can at least partially explain the Warburg effect. Of interest is the suggestion that epigenetics contribute to altered cell metabolism (Johnson et al. 2014). Importantly, what is emerging is a paradigm shift in our understanding and appreciation of the Warburg effect in that the metabolic perturbations may be important in driving tumor growth and survivability, not just the result of certain mutations that hinder carbohydrate metabolism. A comprehensive *omics* approach as discussed in this volume will contribute greatly to our understanding of this fundamental observation made many years ago that has withstood the test of time and countless studies, and along with genomic and proteomic investigations is surfacing again as a likely regulator, not a by-product, of cancer.

1.8 Emerging Roles of Hypoxia, Inflammation and Reactive Oxygen Species in Cancer

A general understanding now exists that hypoxia and inflammation are linked in cancer as well as in other pathological disorders. Hypoxia can lead to inflammation; in turn, inflammation can also lead to hypoxia, both of which can contribute to cancer formation and survival (Grivennikov and Karin 2010; Grivennikov et al. 2010; Eltzchig and Carmeliet 2011; Shay and Simon 2012; Ji 2014; Gorlach 2014). Adding to this pathophysiological interplay, ROS are associated with both hypoxia and inflammation, thus inextricably linking these three conditions and cellular components to cancer (Gorlach 2014; Costa et al. 2014). ROS, including the superoxide anion (O_2^-), hydroxyl radical ($HO^\cdot$) and hydrogen peroxide (H_2O_2), are highly regulated in cells through a combination of generation, e.g., mitochondrial metabolism, and elimination, e.g., via a variety of routes such as superoxide dismutases, catalase, glutathione peroxidase, thioredoxin and others. There are also reactive nitrogen species, but these are not discussed in this section. As briefly mentioned in Sect. 1.7 and

discussed further in subsequent chapters, ROS is elevated in cancer and is believed to contribute to its initiation and subsequent cell growth (Waris and Ahsan 2006; Lu et al. 2007; Liou and Storz 2010; Catalano et al. 2013; Costa et al. 2014).

Hypoxia, or low oxygen tensions, is defined as cellular environments in the presence of 2 % or less oxygen. This is compared to normal, healthy cellular environments of oxygen in the range of 2–9 % (except at high altitudes the air we breathe is 21 % oxygen). Normal physiological responses to overcome hypoxia in the body include increased blood flow and respiration. Under more chronic conditions of hypoxia, two related heterodimers, *HIF1α/HIF1β* and *HIF2α/HIF2β*, respectively, are key players in regulating the myriad cellular responses to low oxygen (Wilson and Hay 2011; Shay and Simon 2012).

In normoxic conditions an enzyme, oxygen-sensitive prolyl hydroxylase, hydroxylates two prolines in *HIF1α*, which results in recognition by an E3 ubiquitin ligase, the von Hippel-Lindau tumor suppressor. The polyubiquinated *HIF* is then targeted for degradation by the 26S proteosome, thus rendering it inactive at normal oxygen tensions. Of interest, the degradation, even under normal oxygen concentrations, can be overcome by mutations in several proteins and by certain signaling pathways. As discussed earlier, *HIF* so stabilized is involved in enhancing glycolysis and inhibiting oxidative phosphorylation.

Another enzyme (factor-inhibiting *HIF*) is also oxygen-dependent and can inhibit *HIF* (via hydroxylation of asparagines on either of the two α subunits). It is the combined action of these two enzymes that monitor and respond to oxygen deprivation. At low oxygen concentrations prolyl oxidation is reduced and the *HIF1α* subunit accumulates and associates with *HIF1β*. This *HIF1* heterodimer is then translocated to the nucleus where it binds to a hypoxia-response element, thus transcriptionally activating several genes including those encoding nuclear factor κB (*NFκB*), toll-like receptors (*TLR*s), *VEGFA* and other growth factors, glucose transporters, most of the glycolytic enzymes (see Fig. 1.2), some enzymes in the pentose phosphate pathway (see Fig. 1.3) and others. These *HIF*-mediated gene activations lead to changes in metabolism, one such adjustment being that ATP production is shifted from oxidative respiration to glycolysis. This is a result of *HIF*'s role in stimulating gene expression of pyruvate dehydrogenase kinase 1, an enzyme that inhibits pyruvate dehydrogenase, the enzyme responsible for the reaction, pyruvate to acetyl-CoA (see Reaction 1.4).

Inflammation refers to a rather detailed and multifaceted process of vascular tissue in response to noxious or harmful stimuli, which can include hypoxia. The disorder, recognized some 2,000 years ago in the west by Celsus and Galen, is characterized by swelling, pain, redness, heat and loss of mobility (or function of a joint). Some of the normal responses of the body to overcome the harmful stimulus include vasodilation of the surrounding vessels to permit more blood flow and increased vessel permeability to permit leukocytes (mainly macrophages and other immune cells), antibodies, fibrin and other components to escape the blood and serve in a protective manner at the site of inflammation. Pertinent to our discussion is the observation that chronic inflammation can lead to cancer, for example hepatitis

B or C viruses give rise to liver cancer, *Helicobacter pylori* infections can result in gastric cancer and tobacco smoking can induce lung and other forms of cancer, to mention but a few.

It is now recognized that a number of mechanisms are involved in inflammatory-associated tumorigenesis (Grivennikov and Karin 2010; Grivennikov et al. 2010; Wu et al. 2014). Numerous signaling pathways lose their regulatory controls and result in pro-inflammatory gene expression related to cancer formation. Genes so activated include protein kinases, e.g., members of the *JAK* (Janus-activated kinase), *MAPK* (mitogen-activated protein kinase) and *PI3K/AKT* (phosphatidylinositol-3-kinase), thus impacting on cell proliferation. As discussed below, immune cells form an integral component of an inflammatory response. Moreover, as will be elaborated on later in the book, cancer cells develop an ability to escape immune destruction and instead use such cells, e.g., lymphocytes (T and B), macrophages, natural killer cells, neutrophils, and others, to produce cytokines that can function in a mitogenic or survival role for the developing, as well as established, cancer cells. For example, cytokines can activate transcription factors such as *STAT3* and *NFκB* that, in turn, can lead to the expression of many genes associated with tumorigenesis: angiogenic regulators, proliferation mediators and anti-apoptosis. Lastly, it has been shown that ultraviolet radiation to melanoma produces an inflammatory response that leads to metastasis (Bald et al. 2014), again documenting the important role of inflammation in cancer.

Recent studies have shown that hypoxia and inflammation are inextricably linked components of cancer. Solid tumors tend to be hypoxic and exhibit features of inflammation. For example, the presence of leukocytes in tumors was noted about 200 years ago. The main component of immune cells within solid tumors is now known to be macrophages, designated macrophage-associated tumors (TAMs). Hypoxia can give rise to inflammation; inflamed tissues are often hypoxic. Both hypoxia and inflammation trigger a series of biological responses that favor cancer growth. As described above, hypoxia of cancer cells, for example, leads to the transcriptional activation of *NFκB* and *TLR*s, as well as other genes encoding proteins involved in the endothelial-to-mesenchymal transition, metastasis, angiogenesis, cell proliferation (*HIF2α*, but not *HIF1α*, increases *c-MYC* activity) and activation of TAMs via secretion of chemokines and cytokines. In addition, hypoxia increases ROS and down-regulates DNA repair mechanisms. Similarly, leukocytes can be recruited and activated by the hypoxic cancer cells and, moreover, respond to hypoxia, also via *NFκB* and *TLR*s, by secreting chemokines and cytokines, as well as additional signals that enhance angiogenesis and other parameters favorable for cancer survival and growth. Necrotic cancer cells, acting through *TLR*s, also activate TAMs. Thus, rather than being benign or even negative aspects of cancer, hypoxia and inflammation participate in promoting cancer growth and metastasis.

Again the case is made for the need of incorporating *omics* approaches to aid in unraveling the many and often overlapping biological processes. The combination of experimental and computational biology is required to reduce the often confusing

and, at times, paradoxical findings into a rational framework. Then and only then can the intricacies of cancer be fully appreciated and individual therapeutic regimens devised.

1.9 Overcoming Apoptosis

For survival all cancer cells must overcome apoptosis, i.e., programmed cell death (Elmore 2007). Apoptosis refers to a series of cellular events including plasma membrane breakage, reduction in cell volume, swelling of mitochondria, and chromatin fragmentation. There are two major pathways involved in apoptosis, namely an intrinsic and an extrinsic cascade; in addition, a third pathway, activated by natural killer (NK) cells and cytotoxic T lymphocytes, serves to lead to apoptosis of targeted cells.

The intrinsic pathway will be discussed first. This pathway is initiated by a variety of non-receptor-mediated factors that activate intracellular signaling pathways. These initiators can be quite diverse and include external and internal factors such as toxins, radiation, free radicals, viral infections and others. Also, certain proteins, e.g., cytokines, can initiate apoptosis simply by their absence, the presence of which inhibits apoptosis. The tumor suppressor protein *P53* is very much at the center of regulating this pathway, as are mitochondria. An important class of proteins is the *BCL2* family that contains both pro-apoptotic members (*BAX*, *BAK*, *BID*, *BOK* and others) and anti-apoptotic members (*BCL2*, *BCLXL*, *MCL1* and others). The link between *P53* and the *BCL* pro-apoptotic proteins is not well understood, but the proteins are known to act on the inner mitochondrial membrane and open the mitochondrial permeability transition (MPT) pore with a loss of the mitochondrial transmembrane potential and the initial release of cytochrome c, *SMAC/DIABLO* and a serine protease *HTRA2*/*OMI*. The heme-containing protein, cytochrome c, interacts with *APAF1* to make the apoptosome, and this structure activates procaspase-9, a member of the caspase (**c**ysteine **asp**artyl-specific prote**ase**s) family of proteases, converting it to the enzymatically active form, caspase-9. The activated caspase-9 then activates the first of the so-called executioner pro-caspases, pro-caspase-3; this in turn continues the proteolytic cascade via the activation of pro-caspases 6 and 7. These proteases serve to cleave a variety of proteins termed death substrates that contribute to the destruction of the cell. Two other mitochondrial proteins released in apoptosis are *SMAC/DIABLO* and *HTRA2*/*OMI*, which function to inhibit *IAP* (inhibitors of apoptosis), that otherwise would antagonize caspase-9. Later in apoptosis several additional proteins are released from the mitochondria, *AIF*, endonuclease *G* and *CAD*, three proteins that are responsible for fragmenting DNA and chromatin condensation.

A distinct pathway, the extrinsic pathway, can mediate apoptosis via transmembrane receptors (referred to as death receptors) that belong to the superfamily of *TNF* (tumor necrosis factor) receptors. Members of this family include *TNFR1*, *FASR*, *DR3*, *DR4* and *DR5*, and these bind, respectively, *TNFα*, *FASL*, *APO3L* and *APO2* (or *TRAIL*) that associates with both *DR4* and *DR5* (also termed *TRAILR1*

and *TRAILR2*, respectively). The ligands bind to their cognate receptor ectodomain, thus triggering a conformational change transmitted through the membrane a cytoplasmic death domain, common to all death receptors. The death domain then binds and activates a *FAS*-associated death domain protein (*FADD*); the complex so formed is denoted as a death-inducing signaling complex (*DISC*). The role of *DISC* in apoptosis is to activate procaspase-8 (or in some cases procaspase-10) that, in turn, is responsible for activating pro-caspases 3, 6 and 7. This latter event represents a converging point for the intrinsic and extrinsic pathways. Moreover, a component of the intrinsic pathway can be recruited to enhance the extrinsic pathway. Here, *BID*, a member of the *BCL2* family, is activated by caspase 3 and functions to open mitochondrial channels, resulting in increased signaling for apoptosis.

The third pathway, also an extrinsic pathway but one requiring NK cells or cytotoxic T lymphocytes to initiate apoptosis, acts via two mechanisms. The most common one acts through the *FAS/FASR* interaction, and the other involves the proteases granzyme *A* and granzyme *B*. These enzymes enter the cell after perforin, a pore-forming protein that opens a channel into the targeted cell, and trigger apoptosis as follows. While both function, it appears that granzyme *B* is the more common of the two pathways. This protease exhibits specificity for cleaving proteins at Asp residues, and consequently serves to activate procaspase-3 and procaspase-10, as well as cleaving intracellular proteins. Granzyme *A*, acting independently of the caspase system, leads to DNA degradation by its actions on two proteins, DNAse *NM23H1* and *SET*, a nucleosome assembly protein.

These pathways, covered in greater depth elsewhere (Elmore 2007), represent challenges that cancer cells must overcome. The many mechanisms used by cancer cells to avoid apoptosis are discussed by Weinberg (Weinberg 2012) and reflect a multi-faceted approach to escape early destruction by the body. One of these responses by a variety of cancer cells is an inhibition, e.g., by overexpression of *MDM2* resulting in an inactivation of the *P53* pathway, thus diminishing its role in apoptosis, as well as permitting cells with damaged DNA to progress through the cell cycle. The same is true for the *RB* pathway, hence overcoming the negative regulation of the cell cycle exerted by this tumor suppressor protein.

Growth factors such as insulin-like growth factor *IGF1* are important in maintaining cell viability, and these may become overexpressed with a concomitant reduction in the expression or activity of the *IGF* binding proteins (*IGFBP*s) that otherwise would render them ineffective. Among the many intracellular signaling pathways activated by *IGF*, one important one for cancer cells is the activation of the *PI3K-AKT/PKB* pathway that results in anti-apoptotic signals. Another mechanism utilized by cancer cells is the overexpression of survivin, an inhibitor of caspases. An inhibitor of the extrinsic apoptotic pathway such as *FLIP* is often expressed to reduce apoptosis in cancer cells. These are but a few of the many changes that have been observed in cancer cells to overcome apoptosis. Indeed, cancer cells have devised multiple strategies for minimizing or even abolishing the three pathways used by normal cells for programmed death. Obviously there is much interest in the design of new drugs that act on these various steps found in cancer cells to aid their survival. This is another major area of interest where the use of *omics* can contribute significantly to new treatment options.

1.10 Contributions of the Extracellular Matrix and Stroma to Cancer

In addition to the changes in the cellular component associated with cancer, there is an important non-cellular component, the ECM, as well as the surrounding stromal cells, both of which have essential roles in the development and progression of cancer. Originally believed to be more of a static unit that maintains tissue integrity, it is now recognized that the ECM is vital to normal cellular function and has emerged as another key factor of cancer initiation and metastasis (Friedl and Alexander 2011; Jinka et al. 2010; Lu et al. 2012; van Dijk et al. 2013). Likewise, the neighboring stromal cells (e.g., fibroblasts), immune cells and endothelial cells (reflecting blood vessel formation), were originally believed to have no role in cancer, but there are now inconvertible data showing that these non-transformed cells contribute significantly to cancer progression (Bhowmick et al. 2004; Tripathi et al. 2012; Calona et al. 2014; Corteza et al. 2014; De Wevera et al. 2014; Escoté and Fajas 2014; Martinez-Outschoorna et al. 2014). Two key properties of ECM are particularly important in the current context: (1) a large number of growth factors tend to be stored with or linked to ECMs and (2) hyaluronic acid, a component of the ECM, has essential roles in all key transitions during carcinogenesis (see Chaps. 6 and 10).

ECM serves as a magnet and storage for a variety of growth factors released into the extracellular space, possibly as a way for their protection against degradation and to maintain them in close proximity to cells. ECM retains growth factors, e.g., bone morphogenetic protein (*BMP*), epidermal growth factor (*EGF*), fibroblast growth factor (*FGF*), hepatocyte growth factor (*HGF*), transforming growth factor β (*TGFβ*) and vascular endothelial growth factor (*VEGF*), by direct binding with ECM proteins such as fibronectin, collagens and proteoglycans (Schultz and Wysocki 2009). Biologically this seems logical since ECM serves as the basis for tissue cells; when a tissue is injured, the damaged ECM will lease its stored growth factors, thus facilitating tissue regeneration and repair.

Constituting a complex network, the ECM contains two main classes of extracellular macromolecules: proteoglycans and fibrous proteins. Several fibrous proteins constitute the non-proteoglycan portion, including the glycoproteins collagen, elastin, fibronectin and laminin. Proteoglycans are formed by the covalent attachment of glycosaminoglycans to proteins. The one exception is that of hyaluronic acid which is not attached to protein. Figure 1.6 shows a schematic illustration of the organization of the ECM.

Glycosaminoglycans refer to unbranched polysaccharide chains comprised of repeating disaccharide units, one of which is an amino sugar. The major amino sugars in glycosaminoglycans are *N*-acetyl-D-glucosamine and *N*-acetylgalactosamine, and the adjoining non-amino sugar is generally D-glucuronic acid or L-iduronic acid. Hyaluronic acid is a glycosaminoglycan like heparin, chondroitin-4-sulfate, chondroitin-6-sulfate, keratin sulfate and dermatan sulfate, and in general, is polydisperse and can contain up to some 250,000 units of the disaccharide D-glucuronic acid and *N*-acetyl-D-glucosamine connected in a $\beta(1 \rightarrow 3)$ linkage (Fig. 1.7).

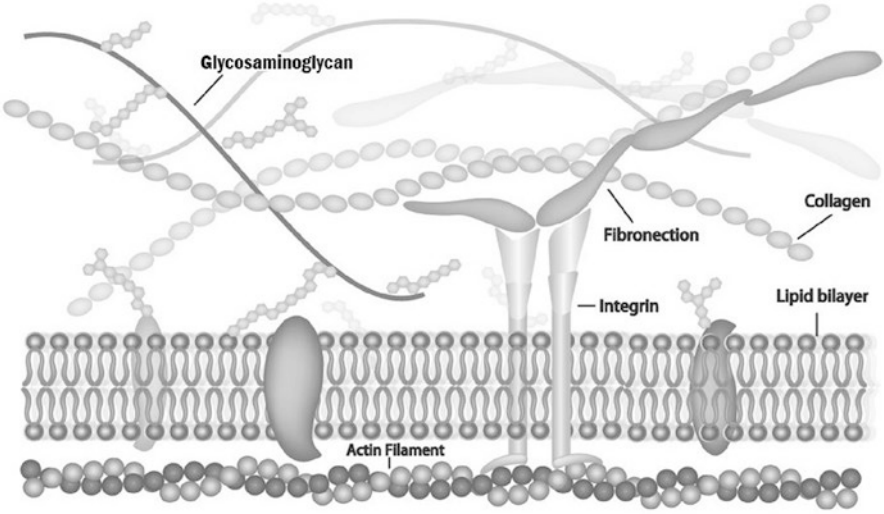

Fig. 1.6 A schematic of extracellular matrix

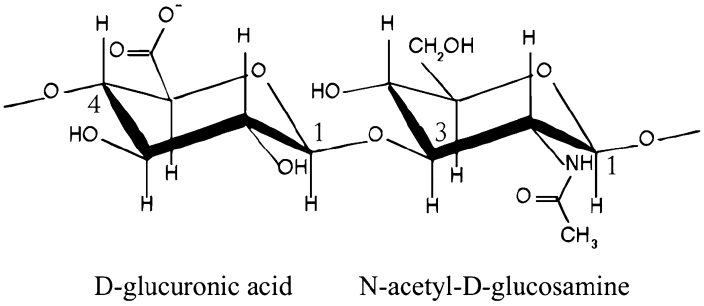

Fig. 1.7 The structure of the repeating disaccharide, D-glucuronic acid and *N*-acetyl-D-glucosamine, that forms hyaluronic acid

The disaccharides are connected to each other in a $\beta(1 \rightarrow 4)$ linkage. Negatively charged due to the COO^- group on D-glucuronic acid, hyaluronic acid is a high molecular weight polyanion that binds several cations such as K^+, Na^+ and Ca^{2+}. It forms a left-handed helix (single strand), one turn of which contains three disaccharides. Pertinent to our interest in this volume, fragments of hyaluronic acid, cleaved by hyaluronidase, have emerged as important structures in cancer, as will be discussed in Chap. 6.

The combination of these complex macromolecular structures yields special biochemical, biomechanical and biophysical properties to the ECM. Although highly complex in nature, the ECM is nonetheless highly regulated during development and tissue homeostasis. Such tight regulation implies well-controlled

transcription, translation and post-translational modifications, and of course remodeling that, for example, may alter the synthesis of one or more components. In addition to the regulation of bioactive ECM macromolecules functioning as structural and cell interaction components, the expression and/or activation of one or more of the ECM degrading enzymes, e.g., matrix metalloproteinases, disintegrin and others, may be changed.

Interaction of the ECM with cells requires a host of macromolecular constituents. One interaction involves members of the class of proteins denoted as integrins. These are heterodimeric (one each of an α and a β subunit) cell surface receptors of which 24 are known, formed from 18 α and 8 β subunits. The integrin ectodomains exhibit binding specificity for a host of ECM macromolecular ligands, including members of the collagen family, fibronectin, laminin, vitronectin and elastin. The ECM ligand-integrin complex mediates its action intracellularly via focal adhesions in which the ligand-bound integrins form clusters, followed by interaction of the integrin cytoplasmic components with a number of cytoskeletal-associated proteins, actin, vinculin, talin and others. Such 'outside-in' signaling of ECM-cell interactions can lead to activation of a number of intracellular signaling pathways involving tyrosine kinases, e.g. *SRC*, focal adhesion kinase (*FAK*), integrin-linked kinase (*ILK*), extracellular-signal-regulated kinase (*ERK*) and others, and tyrosine phosphatases. A particular integrin, β1, is responsible for interacting with the ECM to regulate cell polarity, an important aspect of epithelial cells, particularly relevant to their division. Of interest, the complex just described in which integrin serves as a link between the ECM and the cell interior can function not only for the transmission of information from the extracellular milieu to the cell interior, but the various intracellular interactions with integrin can affect the type of ECM interaction, i.e., signaling from the cell interior to the cell exterior ('inside-out' signaling).

Throughout embryonic development and normal tissue differentiation and homeostasis, there is close interaction between the epithelial cells and the stromal cells. In addition to the important role of epithelial-stromal interaction in normal tissue function, such a cooperation functions in pathological states, e.g., wound healing and cancer, with elaboration on the latter below.

With this abbreviated introduction, the question arises as to how the ECM and stroma affect the initiation, development and metastasis of cancer. For one, the various ECM-cell interactions impact on processes critical to cancer such as proliferation, survival, invasion and migration. An important aspect of tumor initiation and progression involves a change in the integrin expression pattern (Jinka et al. 2012). Higher levels of expression of several integrins correlate with a host of cellular processes conducive to cancer growth and survival: cell proliferation, survival, tissue invasion, migration and new blood vessel formation (angiogenesis). The various integrins preferentially recognize different components of the ECM, e.g., collagen, laminin and fibronectin. These interactions, in turn, lead to activation of a variety of signaling cascades, including *RAS*, *SRC* and others. Several oncogenes, *MYC*, *SRC* and *RAS*, appear to be responsible for the transformation of anchorage-dependent cell growth (normal cells) to anchorage-independent cell growth (cancer cells), which is discussed in detail in Chap. 6.

Changing our focus to the stromal fibroblasts, it has been known for some time that cancer-associated fibroblasts differ from normal fibroblasts (Tripathi et al. 2012). For example, cancer-associated fibroblasts can respond to transformed epithelial cells with increased production of proteases, growth factors and collagen; moreover, the loss of transforming growth factor-β (*TGFβ*) on fibroblasts can serve as initiators of tumorigenesis. The stroma also responds to secretion of *VEGF* by cancer cells, a necessary event in the promotion of new blood vessels to provide blood-borne nutrients for a growing tumor and for colonization to distant sites.

In summary, the interactions of epithelial cells with the ECM and stroma contribute to the formation and growth of epithelial cell cancer, as further discussed in multiple chapters of the book. A better understanding of the various players and mechanisms could lead to new therapeutic modalities.

1.11 Challenging Questions in Classifying and Diagnosing Cancer

With the plethora of potential causes of cancer, ranging from metabolic alterations, hypoxia, inflammation, genomic changes and other changes, coupled with the known heterogeneity of this disease, it should be no surprise that attempts to consistently classify the extent and severity of cancer are challenging. In large part, the identification is based on the site of origin, the appearance of the cells, again compromised to some extent by the heterogeneity of cancer, and its spread to distant sites (often not known). This section provides a synopsis of the current methods in use for cancer diagnosis, grading and staging; a more detailed discussion and the introduction of emerging *omic* contributions are presented in Chap. 3.

Complementing the physical examination, there are a number of techniques in current use to aid in the identification of cancer. These include mammography, positron emission tomography (PET scanning), magnetic resonance (MR) imaging, and in some instances radiographic analysis and measurement of biomarkers, e.g. concentration of circulating prostate specific antigen (*PSA*). The final diagnosis is, however, based on pathological examination of tissue sections from biopsy or resection.

A specimen is judged to be benign or malignant and is then graded. The purpose of cancer grading is to provide an assessment of how abnormal the cells appear and indicate possible treatment modalities. In addition to visual inspection of the section, immunocytochemistry is often used to identify the presence of certain markers that impact on treatment and prognosis, e.g. estrogen receptor in breast cancer. Grading of most solid tumors is done using one of four possibilities, although prostate cancer grading is based on a different scale. Aside from GX which indicates that the grade cannot be assessed, grading will lead to one of the following, where high grade tumors require more aggressive treatment than low

grade tumors: G1: well differentiated (low grade); G2: moderately differentiated (intermediate grade); G3: poorly differentiated (high grade); and G4: undifferentiated (high grade).

Cancer staging, on the other hand, is an assessment of the severity of an individual's cancer, and the stage assigned influences the choice of treatment and provides some information of the prognosis. The following components impact of the staging: tumor size and location, lymph node involvement, cell type and metastasis. The TNM staging system refers to the following three elements: T: extent of the tumor; N: whether or not the cancer cells are present in close proximity in lymph nodes; and M: whether metastasis has occurred.

The extent of the tumor, aside from TX (the primary tumor for whatever reason cannot be evaluated) or T0 (there is no evidence of a primary tumor), is given as Tis, referring to carcinoma *in situ* where the abnormal cells are localized and have not spread to other sites, or one of four designations: T1, T2, T3 and T4, reflecting the size and extent of the primary tumor. The regional lymph nodes, i.e., in close proximity to the primary tumor, can be designated as NX in which the neighboring lymph nodes cannot be evaluated; N0 which specifies that lymph nodes in the immediate vicinity are not involved; and N1, N2, N3, indicating the number of lymph nodes showing involvement. Distant metastasis is represented by MX, M0 or M1, referring to metastasis that cannot be evaluated, no metastasis, or the presence of metastasis, respectively.

It is mentioned only in passing that this staging method is not used for all cancers, but it covers the majority of solid tumors. Yet, the current grading and staging systems are quite subjective in many aspects and woefully inadequate in fully characterizing the important genetic changes leading to the particular molecular and cellular alterations in transformed cells; moreover, they lack discriminatory power when making choices for adjuvant treatment and for predicting likely outcomes with any degree of confidence.

The landscape of cancer characterization is rapidly changing with individual genome sequencing and the use of many of the *omics* techniques in this volume (Cowin et al. 2010). For example, a comprehensive study of breast cancer from 510 tumors obtained from 507 patients was conducted using a variety of methods: exome sequencing, microRNA sequencing, DNA methylation, genomic DNA copy number arrays, mRNA arrays and reverse-phase protein arrays (The Cancer Genome Atlas Network 2012c). Upon combining data from five platforms, they were able to classify four major classes of breast cancer in their starting population.

Studies such as this, now in the experimental stages, will surely emerge in time to offer a more meaningful and systematic classification of all cancers. Such detailed characterization should also prove very useful in deciding on treatment options and providing better prognoses for likely outcomes and disease recurrence. Detailed data of this type will also prove useful in distinguishing driver from passenger mutations and hopefully will provide specificity in seeking specific biomarkers in serum or urine.

1.12 Concluding Remarks

Cancer is a multi-faceted disease, a full understanding of which requires knowledge and information that span a number of scientific disciplines including biochemistry, genetics, and molecular, cellular and developmental biology. The material covered in this chapter on biochemistry and molecular and cellular biology provides the basic knowledge for the reader to follow the discussions in later chapters and to critically assess and utilize the material presented throughout the book for the reader's own research. It is worth emphasizing that cancer is a rapidly evolving system, so that the knowledge learned here, such as biochemical reactions or molecular interactions, is applicable to individual snapshots of an evolving system. Specifically, the environments where the biochemical reactions and molecular interactions take place continue to change. As the environment changes, the catalysts of these reactions and interactions will be altered according to the instructions encoded in the genome and the epigenome in response to the intra- and extracellular conditions, such as the oxygen level, the oxidative stress, the pH and a few others, which are determined by invading endogenous factors, immune responses, cellular metabolism, the genome and epigenomes of the relevant cells. Basically attention is drawn to the study of a dynamic biochemical reaction system. Superimposed upon this evolving cellular reaction system for individual cells, changes also occur at the cancer tissue level, which selects certain cells, and hence their reaction systems that best fit the current environment, and eliminate the others, i.e., Darwin's natural selection theory at work. Specifically, a cancer tissue is constantly changing its cell population by amplifying one sub-clone and inducing the demise of the other subclones as the disease evolves. The knowledge learned here is applicable to each snapshot as a static reaction system, and the information presented in Chaps. 3 through 13 will guide the reader to connect the snapshots along the possible evolution trajectories from multiple perspectives.

References

Alexandrov LB, Stratton MR (2014) Mutational signatures: The patterns of somatic mutations hidden in cancer genomes. Curr Opin Gen Devel 24: 52–60.

Alexandrov LB, Nik-Zainal S, Wedge DC et al. (2013) Signatures of mutational processes in human cancer. Nature 500: 415–421.

Anastasiou D, Poulogiannis G, Asara JM et al. (2011) Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. Science 334: 1278–1283.

Ashworth A, Lord, CJ, Reis-Filho JS (2011) Genetic interactions in cancer progression and treatment. Cell 145: 30–38.

Beck S et al. for the AACR Cancer Epigenome Task Force (2012) A blueprint for an international cancer epigenome consortium. A report from the AACR Cancer Epigenome Task Force. Cancer Res 72: 6319–6324.

Bald T, Quast T, Landsberg J, et al. (2014) Ultraviolet-radiation-induced inflammation promotes angiotropism and metastasis in melanoma. Nature 507: 109–113.

Bensinger SJ, Christofk HR (2012) New aspects of the Warburg effect in cancer cell biology. Sem Cell Devel Biol 23: 352–361.
Bhowmick NA, Neilson EG, Moses HL (2004) Stromal fibroblasts in cancer initiation and progression. Nature 432: 332–337.
Burrell RA, Swanton C (2014) The evolution of the unstable cancer genome. Curr Opin Gen Devel 24: 61–67.
Burrell RA, McGranahan N, Bartek J, Swanton C (2013) The causes and consequences of genetic heterogeneity in cancer evolution. Nature 501: 338–345.
Cairns RA, Harris IS, Mak TW (2011) Regulation of cancer cell metabolism. Nature Rev Cancer 11: 85–95.
Calona A, Taurielloa DVF, Batlle E (2014) TGF-beta in CAF-mediated tumor growth and metastasis. Sem Cancer Biol 25: 15–22.
Catalano V, Turdo A, Di Franco S, Dieli F, Todaro M, Stassi G (2013) Tumor and its microenvironment: A synergistic interplay. Sem Cancer Biol 23P: 522–532.
Chen J-Q, Russo J (2012) Dysregulation of glucose transport, glycolysis, TCA cycle and glutaminolysis by oncogenes and tumor suppressors in cancer cells. Biochim Biophys Acta 1826: 370–384.
Corteza E, Roswallb P, Pietras K (2014) Functional subsets of mesenchymal cell types in the tumor microenvironment. Sem Cancer Biol 25: 3–9.
Costa A, Scholer-Dahirel A, Mechta-Grigoriou F (2014) The role of reactive oxygen species and metabolism on cancer cells and their microenvironment. Sem Cancer Biol 25: 23–32.
Cowin PA, Anglesio M, Etemadmoghadam D, Bowtell DL (2010) Profiling the cancer genome. Ann Rev Genomics Human Gen 11: 133–159.
Dang CV (2012) Links between metabolism and cancer. Genes Devel 26: 877–890.
De Wevera O, Van Bockstalb M, Mareela M, Hendrixa A, Brackea M (2014) Carcinoma-associated fibroblasts provide operational flexibility in metastasis. Sem Cancer Biol 25: 33–46.
Duesberg PH (1987) Cancer genes: Rare recombinants instead of activated oncogenes. Proc Natl Acad Sci 84: 2117–2124.
Eifert C, Powers RS (2012) From cancer genomes to oncogenic drivers, tumour dependencies and therapeutic targets. Nature Rev Cancer 12: 572–578.
Elmore, S (2007) Apotosis: A review of programmed cell death. Toxic Path 35: 495–516.
Eltzchig HK, Carmeliet P (2011) Hypoxia and inflammation. New Engl J Med 364: 656–665.
Escoté X and Fajas L (2014) Metabolic adaptation to cancer growth: From the cell to the organism. Cancer Lett: in press. doi: http://dx.doi.org/10.1016/j.canlet.2014.03.034
Faubert B, Vincent EE, Poffenberger MC, Jones RG (2014) The AMP-activated protein kinase (AMPK) and cancer: Many faces of a metabolic regulator. Cancer Lett: in press. http://dx.doi.org/10.1016/j.canlet.2014.01.018
Fearson ER, Vogelstein B (1990) A genetic model for colorectal tumorigenesis. Cell 61: 759–767.
Ferreira LMR (2010) Cancer metabolism: The Warburg effect today. Exper Mol Path 89: 372–380.
Friedl P, Alexander S (2011) Cancer invasion and the microenvironment: Plasticity and reciprosity. Cell 147: 992–1009.
Fullgrabe J, Kavanagh E, Joseph B (2011) Histone oncomodifications. Oncogene 30: 3391–3403.
Garraway LA, Lander ES (2013) Lessons from the cancer genome. Cell 153: 17–37.
Gorlach A (2014) Hypoxia and reactive oxygen species. In: Melillo G (ed) Hypoxia and cancer, Humana Press/Springer, New York, pp. 65–90.
Grivennikov SI, Karin M (2010) Inflammation and oncogenesis: A vicious connection. Curr Opin Genet Dev. 20: 65. doi:10.1016/j.gde.2009.11.004
Grivennikov SI, Greten FR, Karin M (2010) Immunity, inflammation, and cancer. Cell 140: 883–899.
Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100: 57–70.
Hanahan D, Weinberg RA (2011) Hallmarks of cancer: The next generation. Cell 144: 646–674.

Icard P, Lincet H (2012) A global view of the biochemical pathways involved in the regulation of the metabolism of cancer cells. Biochim Biophys Acta 1826: 423–433.
Icard P, Poulain L, Lincet H (2012) Understanding the central role of citrate in the metabolism of cancer cells. Biochim Biophys Acta 1825: 111–116.
Irmisch A, Huelsken J (2013) Metastasis: New insights into organ-specific extravasation and metastatic niches. Exp Cell Res 319: 1604–1610.
Jang M, Kim SS, Lee J (2013) Cancer cell metabolism: Implications for therapeutic targets. Exper Mol Med 45: e45. doi:10.1038/emm.2013.85
Ji R-C (2014) Hypoxia and lymphangiogenesis in tumor microenvironment and metastasis. Cancer Lett 346: 6–16.
Jinka R, Kapoor R, Pavuluri S, Raj AT, Kumar MJ, Rao L, Pande G (2010) Differential gene expression and clonal selection during cellular transformation induced by adhesion deprivation. BMC Cell Biol 11:93 doi: 10.1186/1471-2121-11-93
Jinka R, Kapoor R, Sistla PG, Raj TA, Pande G (2012) Alterations in cell-extracellular matrix interactions during progression of cancers. Internat J Cell Biology 2012: ID 219196. doi:10.1155/2012/219196
Johnson C, Warmoes MO, Shen X, Locasale JW (2014) Epigenetics and cancer metabolism. Cancer Lett: in press. doi.org/10.1016/j.canlet.2013.09.043.
Kandoth C, McLellan MD, Vandin F et al. (2013) Mutational landscapes and significance across 12 major cancer types. Nature 502: 333–339.
Klein CA (2013) Selection and adaptation during metastatic cancer progression. Nature 501: 365–372.
Koppenol WH, Bounds PL, Dang CV (2011) Otto Warburg's contributions to current concepts of cancer metabolism. Nature Rev Cancer 11: 325–337.
Lawrence MS, Stojanov P, Polak P et al. (2013) Mutational Heterogeneity in cancer and the search for new cancer-associated genes. Nature 499: 214–218.
Liang Y, Liu J, Feng Z (2013) The regulation of cellular metabolism by tumor suppressor p53. Cell Biosci 3: 9. doi:10.1186/2045-3701-3-9
Liou G-Y and Storz P (2010) Reactive oxygen species in cancer. Free Rad Res 44: 479–496.
Loeb LA (1989) Endogenous carcinogenesis: Molecular oncology into the twenty-first century-presidential address Cancer Res 49: 5489–5496.
Lu P, Weaver VM, Werb Z (2012) The extracellular matrix: A dynamic niche in cancer progression. J Cell Biol 196: 395–406.
Lu W, Ogasawara MA, Huang P (2007) Models of reactive oxygen species in cancer. Drug Disc Today Dis Models 4: 67–73.
Mack SC, Witt H, Piro RM et al. (2014)Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. Nature 506: 445–550.
Martinez-Outschoorna UE, Lisanti MP, Sotgiab F (2014) Catabolic cancer-associated fibroblasts transfer energy and biomass to anabolic cancer cells, fueling tumor growth. Sem Cancer Biol 25: 47–60.
Meacham CE, Morrison SJ (2013) Tumour heterogeneity and cancer cell plasticity. Nature 501: 328–337.
Nakajima EC, Van Houten B (2013) Metabolic symbiosis in cancer: Refocusing the Warburg lens. Mol Carcinogen 52: 329–337.
Nowell PC (1976) The clonal evolution of tumor cell populations. Science 194: 23–28.
Oermann EK, Wu J, Guan K-L, Xiong Y (2012) Alterations in metabolic genes and metabolites in cancer. Sem Cell Devel Biol 23: 370–380.
Parker M, Mohankumar KM, Punchihewa C et al. (2014) C11orf95-RELA fusions drive oncogenic NF-κB signalling in ependymoma. Nature 506: 451–455.
Parsons DW, Jones S, Zhang X et al. (2008) An integrated genome analysis of human glioblastoma multiforme. Science 321: 1807–1812.
Pleasance ED, Cheetham RK, Stephens PJ et al. (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 463: 191–197.

Santos CR, Schulze A (2012) Lipid metabolism in cancer. FEBS J 279: 2610–2623.
Schultz GS and Wysocki A (2009) Interactions between extracellular matrix and growth factors. Wound Repair Regen 17: 153–162.
Shay JES, Simon MC (2012) Hypoxia-inducible factors: Crosstalk between inflammation and metabolism. Sem Cell Devel Biol 23: 389–394.
Shen H and Laird PW (2013) Interplay between the cancer genome and epigenome. Cell 153: 38–55.
Soga T (2013) Cancer metabolism: Key players in metabolic reprogramming. Cancer Sci 104: 275–281.
Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. Nature 458: 719–724.
Suva ML, Riggi N, Bernstein BE (2013) Epigenetic reprogramming in cancer. Science 339: 1567–1570.

The Cancer Genome Atlas Research Network [see cancergenome.nih.gov]. Nine of the multi-authored papers are listed below

McLendon R, Friedman A, Bigner D et al. (2011a) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455:1061–1068.
Bell D, Berchuck A, Birrer M et al. (2011b) Integrated genomic analyses of ovarian carcinoma. Nature 474: 609–615.
Muzny DM, Bainbridge MN, Chang K et al. (2012a) Comprehensive molecular characterization of human colon and rectal cancer. Nature 487: 330–337.
Hammerman PS, Lawrence MS, Voet D et al. (2012b) Comprehensive genomic characterization of squamous cell lung cancers. Nature 489: 519–525.
Koboldt DC, Fulton RS, McLellan MD et al. (2012c) Comprehensive molecular portraits of human breast tumors. Nature 490: 61–70.
Getz G, Gabriel SB, Cibulskis K et al. (2013a) Integrated genomic characterization of endometrial carcinoma. Nature 497: 67–73.
Creighton CJ, Morgan M, Gunaratne PH et al. (2013b) Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 499: 43–49.
Ley TJ, Miller C, Ding L et al. (2013c) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. New Engl J Med 368: 2059–2074.
Weinstein JN, Akbani R, Broom BM et al. (2014) Comprehensive molecular characterization of urothelial bladder carcinoma. Nature 507: 315–322.
Timp W and Feinberg AP (2013) Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. Nature Rev Cancer 13: 497–510.
Tomasetti C, Vogelstein B, Parmigiani G (2013) Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. Proc Natl Acad Sci 110: 1999–2004.
Tripathi M, Billet S, Bhowmick NA (2012) Understanding the role of stromal fibroblasts in cancer progression. Call Adhes Migr 6: 231–235.
Tymoczko JL, Berg JM, Stryer L (2013) Biochemistry: A short course, 2nd edn. W.H. Freeman and Company, New York.
Van Dijk M, Goransson SA, Stromblad S (2013) Cell to extracellular matrix interactions and their reciprocal nature in cancer. Exp Cell Res 319: 1663–1670.
Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz, Jr. LA, Kinzler KW (2013) Cancer genome landscapes. Nature 339: 1546–1558.
Vogt PK (2012) Retroviral oncogenes: A historical primer. Nature Rev Cancer 12: 639–648.
Waldman T and Schneider R (2013) Targeting histone modifications-Epigenetics in cancer. Curr Opin Cell Biol 25: 184–189.

Waris G and Ahsan H (2006) Reactive oxygen species: Role in the development of cancer and various chronic conditions. J Carcinogen 5: 14. doi:10.1186/1477-3163-5-14

Webb BA, Chimenti M, Jacobson MP, Barber DL (2011) Dysregulated pH: A perfect storm for cancer progression. Nature Rev Cancer 11: 671–677.

Weinberg RA (2012) The biology of cancer. Garland Science, New York.

Wilson WR, Hay MP (2011) Targeting hypoxia in cancer therapy. Nature Rev Cancer 11: 393–410.

Wu Y, Antony S, Meitzler JL, Doroshow JH (2014) Molecular mechanisms underlying chronic inflammation-associated cancers. Cancer Lett 345: 164–173.

Yoshii Y, Furukawa T, Saga T, Fujibayashi (2014) Acetate/acetyl-CoA metabolism associated with cancer fatty acid synthesis: Overview and application. Cancer Lett 347: 204–211.