INTRODUCTION: THE ORIGINS AND DEVELOPMENT OF SPONTANEOUS CANCERS

Cellular targets of the genetic alterations of cancers: tissue-specific stem cells (TSSCs)

A large collection of evidence indicates that every organ in our body contains TSSCs, which are capable of replicating themselves or committing to differentiate into many or all of the cell types that make up the organ. Among the first clear demonstrations of these TSSCs were hematopoietic stem cells (CD34+ cells), which can produce many of the blood cells and some tissue-associated cells in several organs [1]. These stem cells begin their existence in the embryonic yolk sac, move to the fetal liver, and then reside in the bone marrow throughout the rest of life. The colon stem cell is termed the slim cell. It resides in the crypt of the colon [2] and replicates to produce a large number of cell types that make up the colon. The epidermal stem cells are responsible for producing skin and reside near the hair follicles [3]. These and other TSSCs have a number of properties in common. Each type of TSSC replicates symmetrically (reproducing more stem cells) increasing or replenishing the pool of TSSCs. These stem cells are duplicated in response to signals from the Wnt pathway, employing frizzled receptors that regulate the levels of beta-catenin, which in turn regulate both cell adhesion molecules (Cadherins) and act in the nucleus as a heterodimer with TCF-4 (T-cell factor 4) to transcribe a series of Wnt-regulated genes. The expression of this pathway is further enhanced by extracellular factors termed the R-spondins, which act upon G-protein coupled receptors, increasing Wnt induced transcriptional activity. The G-linked receptors, LGR-4, 5, 6 (leucine-rich repeat) are found in ectodermal and endodermal derived TSSCs (skin and colon) as well as some mesenchymal stem cells (blood cells) [4]. After a pool of TSSCs are formed in an organ, some of them are selected to differentiate, altering their epigenetic marks by producing progenitor cells or intermediates in the developmental pathway forming the tissue. The body of a normal human will produce billions of new cells each day. Some of this comes from tissue repair, but the majority contributes to the varying turnover of tissues; the half-life of colon cells is about four days [5]; the half-life of skin tissue is about twenty-eight days [6], and blood cells turn over continuously [1].

This large amount of replication of many TSSCs in their niches or compartments of the body over a lifetime results in the accumulation of mutations in the TSSCs, which then compete with each other for reproductive fitness, or propagation, of clones of stem cells. Weissman [1] has called these stem cells “units of natural selection for tissue formation, for germ line development and cancer development”. The mutation rate in these tissue-specific stem cells has been estimated at about 20–60 mutations per year of life using whole-genome sequencing of cloned organoids produced from single stem cells in culture. The stem cells were obtained from individuals aged 3–87 years old and from several different organs [7] and single-cell DNA sequencing was also performed [8]. When these mutations impact genes of TSSCs in pathways for cell cycle reproduction, cell death, or DNA damage repair the stem cells can be selected for reproductive fitness and they can compete to take over the pool of TSSCs in an organ. These experiments predict that over a lifetime TSSCs will accumulate clones with mutations that are precursors to a series of mutations that lead to the development of cancers. Deep DNA sequencing from normal tissues of individuals has identified such clones of cells with mutations in oncogenes or tumor suppressor genes [9, 10]. Mutations that give rise to a clonal expansion of a TSSC in a tissue have been termed the initial, or truncal mutation, in a pathway leading to the development of clonal expansion and the start of development of cancer.

Selectivity of expression or function of oncogene or tumor suppressor gene mutations in TSSCs

There are hundreds of oncogene mutations that can contribute to cancer development and tens of hundreds of tumor suppressor gene mutations that can permit cancer development. Based upon the frequency of each oncogene or tumor suppressor gene mutation in
different spontaneous cancerous tissues, and therefore different TSSCs; there is a clear preference for one or another gene mutation in different tissues (see the TCGA as an example [11]). This is even more striking with inherited tumor suppressor gene mutations such as RB1, TP53, PTEN, BRCA1 or BRCA2, etc. [9, 10]. BRCA-1 and 2 are expressed and employed for homologous DNA repair in many cell and tissue types of the body, yet mutations in these genes give rise to ovarian and breast cancers at a higher incidence than many other tissues. Elledge and his colleagues [12] have pointed this out and explored the reasons why the same mutated oncoproteins expressed in different tissue lineages not only can give rise to cancers in one tissue type but not the other but also the same mutant protein (B-RAFV600E) present in colon cancers and melanomas respond to inhibitors in the latter but not in former cancer. They have pointed out that the different epigenetic states in different tissues and TSSCs have an impact upon mutational selection, and the signal transduction pathways that contain these mutated genes are structured differently, with or without feedback loops or additional components.

The order of mutations in oncogenes and tumor suppressor genes in TSSC determines the properties and the age of onset of cancers

Based upon the conclusions derived from studies such as those discussed above, the initial, or truncal mutation, spontaneously occurring or inherited in a TSSC exerts a positive selection for reproductive fitness in a cell population of TSSCs [9, 10]. This results in a clonal expansion of that stem cell, and so a second mutation is more likely to occur in that clone of TSSCs containing the first mutation. If the second mutation has no further impact upon the fitness of the cell, then it will not contribute to the development of cancer, but if it adds to the fitness of the TSSC, then it expands the cell number and target size for additional mutations. Over the past several years a large literature has accumulated demonstrating the presence of such mutations in selected genes in normal or benign tissue that has clonally expanded in a local area of tissue over a lifetime [13], in skin and esophagus for example [14, 15], where a calculated reproductive fitness with NOTCH1 mutations and TP53 mutations can be quantified [16]. Similarly, the expansion of myeloid precursors in the blood over time, and treatment with mutagens gives rise to CHIP (clonal hematopoiesis of indeterminate potential), a benign expansion of myeloid cell numbers with selected mutations at reproducible frequencies (DNMT3A, TET-2, ASXL-1, JAK-2, SF3B1, and TP53) [17–19]. Clevers and colleagues [20, 21] have obtained a similar result, as the selection of a mutation, so as to confer an added reproductive fitness upon the cell and its progeny, inserts an order to the mutational progression from benign to malignant cancers. This was first shown in three different and independent ways employing colorectal cancers. Vogelstein and his colleagues obtained benign and malignant tumors of the colon from colonoscopies employing a number of human patients. The smallest polyps that were obtained contained a mutation in the adenomatous polyposis coli gene (APC). Slightly larger but still benign polyps had the APC mutation and a KRAS mutation. Even larger benign polyps obtained from patients had the APC mutations, a KRAS mutation, and a SMAD mutation (in the TGF-beta pathway). Finally, malignant colorectal cancers had all three of the previous mutations plus mutations in both alleles of the TP53 gene [20, 21]. Vogelstein inferred an order in these mutations from the sizes of the tumors and the benign to malignant transition. In a second study, Sato and Clevers started with normal human colorectal tissue-specific stem cells in organoid cultures and introduced selected mutations in the APC, KRAS, SMAD-4, and TP53 genes in different orders and combinations, employing CRISPR-Cas9 [22]. These experiments demonstrated that the most efficient way to produce a malignant tumor by adding mutations was in the same order observed by Vogelstein. In the third set of experiments, this time carried out in mice, Jenkins and Copeland initiated colon cancers by turning on a transposon that either activated oncogenes or inactivated tumor suppressor genes by random insertion into the genome. They observed that it took a long time (about 100 days) to obtain these tumors in all the mice under study. When theTp53 gene was inherited in the germline and then the transposon was activated it took 80 days; when Smad4 was inherited it took 60 days; when KRAS was inherited it took 50 days, and when APC was inherited it took 25 days to produce colorectal cancers [23]. Rather clearly, all three experiments agree that colorectal cancers are formed, most commonly, by an ordered selection of random mutations, and individuals will develop colorectal cancer earlier in their lifetime if they inherit a truncal, or initial mutation, in this case APC, first in a specific gene order in a TSSC. The efficiency of tumor formation is determined by an ordered selection for reproductive fitness, resulting in a series of benign clonal expansions, until the last mutation creates a malignant tumor. The random nature of mutations gives rise to a tremendous diversity of mutations in a tumor so as to obscure the ordered progression of the mutated genes essential for cancer formation.

INHERITED TP53 MUTATIONS AND THE LI-FRAUMENI SYNDROME

How can we determine what mutations in tumor suppressor genes are the initial or truncal mutations giving rise to an expanded clonal number of mutant TSSCs for each tissue-specific cancer? By definition, germline mutations in tumor suppressor genes are the truncal or initial mutations in a TSSC that will give rise to cancer at a young age [9, 10]. Because the inherited mutation occurs in every TSSC type in the body, the preferential tissue-specific phenotypes of the cancers formed to provide evidence for selection and clonal expansion of that mutation in a TSSC shortly after birth. Three phenotypes of the tumors arising in these patients determine the nature of the TSSC that has a functional initial or truncal TP53 mutation that initiates tumor formation: (1) A very early age of onset of the tumor formation, (2) the tumor tissue type that is produced at this very early age indicates the TSSC and its clonal expansion, and (3) the excess risk of the tumor tissue type in Li-Fraumeni Syndrome patients compared to the general population. Figure 1 presents these phenotypes for tumor tissue types most commonly produced over the lifetime of a Li-Fraumeni patient with a germline TP53 mutation. Rather clearly, the ages at which specific tumor tissue types are produced fall into three categories: six months to 20 years, 20–45 years, 45–70 years. At the youngest ages are medullary blastomas, choroid plexus carcinomas and papillomas, adenocortical tumors, rhabdomyosarcomas, and osteogenic sarcomas. In the

Tumor Tissue Types, Ages of Onset and Excess Risk of Cancers

| Ages       | 0 months–20 yrs | 20–40 yrs | 40–75 yrs | > 75 yr |
|------------|-----------------|----------|----------|--------|
| Excess Risk| 100X            | 20-40X   | 2-4X     | 0.5X   |
| Tumor Type | Medullary blastoma, Choroid Plexus Glioma, Adrenocortical Carcinoma, Osteosarcoma, Rhabdomyosarcoma | Breast Carcinoma, Liposarcoma, Leiomysarcoma, Soft Tissue Sarcomas | Lung, Colon, Pancreatic | Multiple Types |

Fig. 1 Tumor tissue types, ages of onset, and excess risk of cancers. Li Fraumeni patients inherit a mutant TP53 allele that acts at an early age of onset (six months to 20 years) indicating that the mutant p53 protein functions as an initial or truncal mutation promoting benign cell replication and clonal expansion and producing a high excess risk for cancer. TP53 mutations that give rise to cancers later in life more likely function as a later mutation in an ordered series of mutant genes that drive cancerous growth. By 50–70 years of life, cancers arise with low excess risk and the mutant p53 protein acts in the benign to malignant transformation as the last gene in the ordered series of cancer forming mutations [9, 10, 27].
producing a tetramer that binds to 20 base pairs or two turns of the DNA helix [36]. Some mutations in this domain can also cause familial and spontaneous cancers, but these tend to have lower penetrance and a weaker cancer-causing phenotype [37]. The fifth domain can regulate transcriptional activity. It contains seven lysines whose epsilon amino group can be acetylated to enhance transcription or methylated to inhibit transcription in some stem cells [38, 39]. Deletion of this domain, or phosphorylation of it, enhances DNA binding. These properties are reviewed in Fig. 2.

The function of the p53 protein is to respond to a wide variety of intrinsic and extrinsic cellular stresses. Stress is defined here as any interference with an orderly progression of cellular functions or division. In response to many different types of stresses (Fig. 3 highlights a DNA damage stress), epigenetic modifications or signals are sent from a stress detector (ATM) through a stress mediator (CHEK2) to the p53 (serine -15) protein and MDM2 protein, which is the E3 ubiquitin ligase regulating the instability of transcriptional gene functions, and even possible gain of function mutations [33]. The frequencies with which each of the ten most common TP53 mutations (33% of cancers), which have very little or no transcriptional activity. More than 350 additional TP53 missense mutations have a weak transcriptional activity, a much lower frequency of occurrence in producing cancers over a four log distribution, and cause 67% of cancers. The hypothesis that pro-apoptotic XAF-1 binding to a weak transcriptional p53 protein promotes apoptosis so that the weak allele fails to form cancer. An XAF-1 gene transcriptionally silenced by epigenetic marks or a polymorphism that inactivates the protein would then permit weak, minor TP53 alleles to produce cancer [37].

**THE DOMAIN STRUCTURE AND FUNCTIONS OF THE P53 PROTEIN**

The human p53 protein is a transcription factor that regulates a series of genes forming a large interconnected pathway [10]. The human p53 protein is composed of 393 amino acids commonly numbered from the N-terminus –1, to the C-terminus –393. The protein is composed of five structural and functional domains (Fig. 2). Amino acids 1–55 include two transactivation regions (amino acids 21–28 and amino acids 47–55) [29, 30] each of which regulates some similar and some different sets of genes. These amino acid sequences attract and assemble RNA polymerases and chromatin-modifying enzymes that open the chromatin for transcription and add protein modifications to the polymerase and p53 protein, helping to direct which genes are transcribed. Amino acids 55–100 form a proline-rich domain with repeated sequences of PXXP (P representing proline and X any amino acid). This is a protein-protein interaction domain that regulates cell growth and apoptotic efficiency [31, 32]. Amino acids 100–300 form a sequence-specific DNA binding domain that imparts upon the p53 protein the specificity to identify genes it regulates. There are hundreds of different missense mutations in this domain and they differ from each other in their DNA binding efficiencies, loss of transcriptional gene functions, and even possible gain of function mutations [33–35]. The frequencies with which each of these hundreds of different mutations occur in all cancer types differs by up to four orders of magnitude. The loss of DNA binding, transcriptional efficiency, and frequency of each of these mutations are strongly correlated [33–35]. The fourth domain at amino acids 320–345 is the tetramerization domain of the p53 protein. The p53 transcription factor forms from two dimers producing a tetramer that binds to 20 base pairs or two turns of the DNA helix [36]. Some mutations in this domain can also cause familial and spontaneous cancers, but these tend to have lower penetrance and a weaker cancer-causing phenotype [37]. The fifth domain can regulate transcriptional activity. It contains seven lysines whose epsilon amino group can be acetylated to enhance transcription or methylated to inhibit transcription in some stem cells [38, 39]. Deletion of this domain, or phosphorylation of it, enhances DNA binding. These properties are reviewed in Fig. 2.

The function of the p53 protein is to respond to a wide variety of intrinsic and extrinsic cellular stresses. Stress is defined here as any interference with an orderly progression of cellular functions or division. In response to many different types of stresses (Fig. 3 highlights a DNA damage stress), epigenetic modifications or signals are sent from a stress detector (ATM) through a stress mediator (CHEK2) to the p53 (serine -15) protein and MDM2 protein, which is the E3 ubiquitin ligase regulating the instability of transcriptional gene functions, and even possible gain of function mutations [33]. The frequencies with which each of the ten most common TP53 mutations (33% of cancers), which have very little or no transcriptional activity. More than 350 additional TP53 missense mutations have a weak transcriptional activity, a much lower frequency of occurrence in producing cancers over a four log distribution, and cause 67% of cancers. The hypothesis that pro-apoptotic XAF-1 binding to a weak transcriptional p53 protein promotes apoptosis so that the weak allele fails to form cancer. An XAF-1 gene transcriptionally silenced by epigenetic marks or a polymorphism that inactivates the protein would then permit weak, minor TP53 alleles to produce cancer [37].

**THE DOMAIN STRUCTURE AND FUNCTIONS OF THE P53 PROTEIN**

The human p53 protein is a transcription factor that regulates a series of genes forming a large interconnected pathway [10]. The human p53 protein is composed of 393 amino acids commonly numbered from the N-terminus –1, to the C-terminus –393. The protein is composed of five structural and functional domains (Fig. 2). Amino acids 1–55 include two transactivation regions (amino acids 21–28 and amino acids 47–55) [29, 30] each of which regulates some similar and some different sets of genes. These amino acid sequences attract and assemble RNA polymerases and chromatin-modifying enzymes that open the chromatin for transcription and add protein modifications to the polymerase and p53 protein, helping to direct which genes are transcribed. Amino acids 55–100 form a proline-rich domain with repeated sequences of PXXP (P representing proline and X any amino acid). This is a protein-protein interaction domain that regulates cell growth and apoptotic efficiency [31, 32]. Amino acids 100–300 form a sequence-specific DNA binding domain that imparts upon the p53 protein the specificity to identify genes it regulates. There are hundreds of different missense mutations in this domain and they differ from each other in their DNA binding efficiencies, loss of transcriptional gene functions, and even possible gain of function mutations [33–35]. The frequencies with which each of these hundreds of different mutations occur in all cancer types differs by up to four orders of magnitude. The loss of DNA binding, transcriptional efficiency, and frequency of each of these mutations are strongly correlated [33–35]. The fourth domain at amino acids 320–345 is the tetramerization domain of the p53 protein. The p53 transcription factor forms from two dimers producing a tetramer that binds to 20 base pairs or two turns of the DNA helix [36]. Some mutations in this domain can also cause familial and spontaneous cancers, but these tend to have lower penetrance and a weaker cancer-causing phenotype [37]. The fifth domain can regulate transcriptional activity. It contains seven lysines whose epsilon amino group can be acetylated to enhance transcription or methylated to inhibit transcription in some stem cells [38, 39]. Deletion of this domain, or phosphorylation of it, enhances DNA binding. These properties are reviewed in Fig. 2.

The function of the p53 protein is to respond to a wide variety of intrinsic and extrinsic cellular stresses. Stress is defined here as any interference with an orderly progression of cellular functions or division. In response to many different types of stresses (Fig. 3 highlights a DNA damage stress), epigenetic modifications or signals are sent from a stress detector (ATM) through a stress mediator (CHEK2) to the p53 (serine -15) protein and MDM2 protein, which is the E3 ubiquitin ligase regulating the instability of transcriptional gene functions, and even possible gain of function mutations [33]. The frequencies with which each of the ten most common TP53 mutations (33% of cancers), which have very little or no transcriptional activity. More than 350 additional TP53 missense mutations have a weak transcriptional activity, a much lower frequency of occurrence in producing cancers over a four log distribution, and cause 67% of cancers. The hypothesis that pro-apoptotic XAF-1 binding to a weak transcriptional p53 protein promotes apoptosis so that the weak allele fails to form cancer. An XAF-1 gene transcriptionally silenced by epigenetic marks or a polymorphism that inactivates the protein would then permit weak, minor TP53 alleles to produce cancer [37].
of the p53 protein by poly-ubiquitination. The inhibition of the MDM2 activity (along with other proteins that function with MDM2) increases the half-life of the p53 protein within minutes of the stress occurring. The half-life of the p53 protein increases from minutes to hours and the epigenetic modifications plus the increased concentrations of the p53 protein activate transcription of a selected set of genes (Fig. 3). The epigenetic modifications provide information about the nature and intensity of the stress, which determines the response (transcriptional program) that may be either cell cycle arrest and repair of the damage or cell death.

The properties of the p53 pathway include:
1. The gaining of epigenetic information about the nature and intensity of stress.
2. The making of a decision about which transcriptional program will be employed to respond to the stress.
3. The p53 pathway is very redundant in how it mediates cell cycle arrest at G-1 and at G-2.
4. The pathway can make any of five choices of the types of cell death it employs to kill cells.
5. The pathway is extensively connected to many other pathways (metabolic, immunological, the cell cycle, DNA damage repair, ribosome biogenesis, cell death and differentiation, etc.) [10].

These five properties of the p53 pathway give rise to the observation that mutations in the TP53 gene are the most common genetic alteration in human cancers. Understanding the loss and/or the gains of functions of the p53 mutant protein is essential to an understanding of the origins and properties of diverse cancers.

THE FREQUENCIES OF SPONTANEOUS AND INHERITED TP53 MUTATIONS IN HUMANS WITH AND WITHOUT CANCERS
As previously discussed, there are hundreds of different TP53 missense, nonsense, deletion, frameshift, and insertion mutations in the TP53 genes sequenced from individuals with spontaneous and inherited cancers [33–35]. Some of these mutations tend to be tissue-specific because they are brought about by different known mutagens in the environment (34,35, see Fig. 2). About 20–25% of them are true loss of function mutations (nonsense, deletions, frameshifts), but the missense mutations, which represent 75–80% of the mutations, occur at different frequencies (over four orders of magnitude) in all tissue types of human cancers [33–35]. These different missense mutations often have different phenotypes that are based upon diverse DNA binding efficiencies, transcription of different genes, temperature-sensitive properties, and possible gain of functional phenotypes. The ten most frequent TP53 mutations in cancers are present in 33% of the cancers (Fig. 2). The remaining hundreds of other TP53 missense mutations are found (more than ten independent times) in 67% of cancers. The top ten most frequent TP53 mutant alleles (Fig. 2) have common features. They are very poor at transcribing p53 specific genes [33–35] and some mutations are at the DNA binding contacts of the protein while others change the structure and melting temperatures of the p53 protein [35]. Attempts are now underway to classify different mutant alleles that are inherited and to explore the penetrance of different TP53 mutations, their ability to show any tissue specificity (expression or function or loss of function in TSCCs), or other phenotypes.

The great majority of individuals who inherit TP53 genetic alterations are identified by virtue of developing specific tumor types at a very young age and then having their TP53 gene sequenced, or because they are related to patients from identified Li-Fraumeni families. These individuals and families are estimated to occur in the general population at about 1/20,000 people, which is not too different from the frequency of mutations in other tumor suppressor genes occurring with fairly high penetrance.

K. de Andrade, P. Hainaut, M. Achatz, and S. Savage [40, 41] and others have explored sequencing databases composed of unrelated individuals not selected for cancer history. Then they ask what the frequencies of TP53 germine mutations in the remaining group are, so as to eliminate the bias of having been diagnosed with specific cancers at a very young age. Surprisingly, in these databases the prevalence of carriers of potentially pathogenic germine TP53 mutations varies from 1/500–1/5,000 persons compared to the previous estimated incidence of 1/20,000 who are first identified because of cancer. Although these prevalence estimates are dynamically altered based on variable mutation classifications and highly impacted by some specific controvertible spontaneous TP53 mutations, the databases employed could introduce other possible biases, such as CHIP, or enrichment of mutations potentially associated with lower cancer penetrance or with phenotypes not typically associated with LFS [17, 18]. With some uncertainty, these initial numbers bring up the possibility that there may be environmental or genetic suppressors of TP53 mutations that lower cancer penetrance or weaken the phenotype. A weak TP53 mutant allele, defined as being at very low frequencies in individuals with either spontaneous or inherited cancers, could be at that low frequency because of genetic and/or environmental factors. As a larger number of these individuals with TP53 mutations, but no cancerous phenotypes are identified, the frequency, nature, and properties of each allele will be recorded, and possible genetic suppressors or environmental history could help to explain the phenotype. The identification of such genetic or environmental suppressors is valuable, in that the identified genes, or environmental activity, can be useful in developing therapeutic approaches to treatments.

GENETIC MODIFIERS THAT COULD BE OF INTEREST
There is an excellent and detailed review of TP53 polymorphisms acting as genetic modifiers of p53 related phenotypes that demonstrate the complexities of multi-genic or even multi-allelic variations upon cancer phenotypes [42]. What follows is a brief overview of these polymorphisms and some additional observations that reinforce the importance of this topic, both at the basic and clinical levels of understanding tp53-related cancers.

XAF-1 (XIAF associated factor-1) gene and protein
The XAF-1 protein is a 33.1 Kda protein whose gene is located on human Chromosome 17, just two mega-bases away from the TP53 gene. The protein has seven zinc fingers, and was first shown to bind to the linked inhibitor of apoptosis (XIAF) and inactivate it, promoting apoptosis [43]. The XAF1 gene is positively regulated for transcription by IRF-1 (interferon regulatory factor-1) promoting TNF mediated cell death; XAF1 transcription is negatively regulated by HSF-1 (heat shock factor-1) [44]. The gene has been classified as a tumor suppressor that is commonly shut down at the level of transcription by heavily methylated chromatin [45]. In addition to these regulatory functions, the XAF1 protein acts at three different p53 intersecting pathways to increase p53 levels in the cell, resulting in the transcription of several p53 regulated genes whose proteins promote apoptosis (Fig. 4). First, the XAF1 protein strongly promotes p53-mediated apoptosis by the binding of XAF1 to the proline domain of p53, which in turn competes off the binding of the MDM-2 protein, which is the E3-ubiquitin ligase for p53. This stabilizes the p53 protein, increasing its half-life and concentration [46]. In addition, the proline-rich domain is known to impart growth-inhibiting and pro-apoptotic activity to the p53 protein [29, 30]. Second, the XAF1 protein binds to and inhibits the ubiquitin ligase SIAH-2, which in turn, stabilizes the HIPK2 (homeobox interacting protein kinase -2) that now phosphorylates serine 46 of the p53 protein, promoting the transcription of several p53-regulated pro-apoptotic genes (APAF-1, BAX, NOXA,
TP53 Pro 47 Ser, (rs1800371) polymorphism

The proline-47 to serine amino acid polymorphism in the TP53 gene arose in an African haplotype, is found in 6% of Africans, and is present in 1% of African Americans [47]. Given the discussion above, where phosphorylation of serine-46 is required for an efficient p53 apoptotic pathway, it is understandable that serine-47 in the p53 protein impairs apoptosis after most, but not all, DNA damage [47]. The Serine-47 polymorphism results in reduced levels of phosphorylation of serine-46 after many types of DNA damage, chemotherapy, and radiation. This variant of p53 is defective in the ability to regulate genes involved in the sensitivity of cells to ferroptosis, a process of iron- and lipid peroxide-mediated cell death. This ferroptotic defect is caused by increased anti-oxidant accumulation in serine-47 cells [48]. This leads to iron accumulation in mouse and human cells containing the serine-47 variant, and indeed this variant is associated with a disease called “iron overload” in individuals of African descent [49]. Serine-47 is also associated with increased cancer risk in mice and humans and with decreased sensitivity of tumor cells to many chemotherapeutic drugs [47]. However, the cisplatin class of drugs and BET inhibitors (Bromo-domain and extra terminal motif) do induce apoptosis in tumors with serine-47 [50, 51], demonstrating remarkable specificity between the use of therapies and the genotype of the TP53 gene and its protein. Interestingly, the increased glutathione accumulation in serine-47 mice causes increased activity of mTOR, a major regulator of metabolism [52]. Consequently, serine-47 mice are larger and more physically fit than control littermates, suggesting that this variant may have been selected for at one time [52].
Fraumeni families [28].

It is possible that this polymorphism is protective, with increased levels of prostate cancers (OR = 1.44), gliomas (OR = 2.35), and colorectal adenocarcinomas (OR = 1.39), and the SNP is strongly associated with cutaneous basal cell carcinoma [58]. Several of the SNPs (polymorphisms) discussed above are in strong linkage disequilibrium when measured in selected populations (African descent and Caucasians) and environments: SNPs in introns 2 and 3 along with codon 47 (proline to serine) and codon 72 (proline to arginine) demonstrate positive linkage disequilibrium. This suggests that these SNPs have co-evolved and define complex haplotypes that are found in populations that evolved in different geographic localizations. It is likely, therefore, that these sets of SNPs (haplotypes) are composed of genetic variants that interact or cooperate to express specific functional properties that either optimize stress responses and/or tumor suppression in response to environmental variations, as first suggested by P. Hainaut [27].

SOME QUESTIONS REMAIN TO BE EXPLORED
What are the functions of the TP53 gene and protein that are important to prevent cancers?
The p53 gene and protein are only present in multicellular animals, not in plants, bacteria or yeast. The origins of this gene have been traced back to a common ancestor of choanoflagellates and humans, some 600–800 million years ago [59]. In all invertebrates the TP53 protein DNA binding domain binds to the same DNA sequence it binds to in humans. Furthermore, the vertebrate and invertebrate orthologues of the p53 protein regulate similar genes for cell death in response to genomic damage [59]. This conservation of protein sequences, DNA binding sequences and regulated gene functions suggest a real importance for multicellular animal life processes. The second reason why TP53 is important to study is that genetic alterations in the TP53 gene are the most common mutations observed in cancers of humans [10, 11, 33, 34].

These two observations suggest that the answer to the question, “what p53 regulated gene protects us against cancers” would be useful to ask experimentally. Attempts to test this question have not led to the uncovering of a single important gene for tumor suppression. Rather, it seems more likely that the answer is in the structure and functions of the entire p53 pathway and its properties discussed in this article. The redundancies in the pathway allow it to function in spite of destructive mutations eliminating multiple pathways for enforcing cell death. In addition, the method or process of cell death chosen by the p53 protein sensing a stress (Figure 3) could impact the nature and types of immune response to antigen presentation by dying cells. If redundancy is such an important feature, why do simple single missense mutations in the TP53 gene itself, disrupt the entire pathway? Why didn’t evolution select for redundancies in TP53 like functions producing many TP53 like genes to back it up? Some answers to this and related questions are explored elsewhere [10].

What is the mutational profile, or signature, of genetic alterations in cells with TP53 mutations? Is this a clue to why it is central to cancer production?
The p53 protein regulates the number of centrosomes that duplicate at each cell division and permit accurate segregation of chromosomes [60]. The loss of this p53 function results in a failure to synthesize p21 and thus cyclin E is no longer under p53 control. Both p21 and cyclin E bind to centrosomes, and the subsequent overexpression of cyclin E gives contributes to multiple centrosomes and subsequent aneuploidy. 1. There are losses and gains in chromosomes and dramatic changes in gene copy number. 2. The loss of p53 functions results in deletions and gene amplification, both by loss of entire chromosomes and by local DNA loss and gene amplifications, 3. chromothripsis, the fragmentation of chromosomes and the reassembly of the parts in a novel order, and the formation of circular DNA and double minute chromosomes.

These types of mutations result in high levels of cell death, selection for cells that replicate with abnormal gene copy numbers, divide and are selected for fitness, metastasis, gene amplification, and hyper- and hypodiploid genotypes. Chromosomal and gene copy number abnormalities are commonly lethal in normal cells, perhaps because the cells are killed by TP53 functions that sense the abnormality, but the loss of TP53 seems to permit the toleration of copy number abnormalities. The mechanisms behind the tolerance of abnormal gene copy numbers are not well understood. These phenotypes form the signature of TP53 loss of function. Perhaps the restoration of p53 functions in a cancer cell will restore the lethality of abnormal gene copy numbers and eliminate cancer.

Are there genetic, epigenetic, and environmental modifiers of TP53 mutations?
An abundance of genetic and epigenetic modifiers of p53 function, are highlighted in this review. There are several interesting conclusions. Epigenetic modifiers can change the age of onset of tumors, the elimination of tumors by weak TP53 mutant alleles by apoptosis, and even the frequency of tumor types and tissue types. There is a strong developmental impact upon TP53 gene expression or function in different TSSCs derived from ectoderm, mesoderm, and endoderm. Different environmental carcinogens (cigarette smoke, lung cancers and codon 157 mutations; aflatoxin, liver cancers and codon 249 mutations) have been shown to act upon the TP53 gene mutating specific bases and codons in specific tissues that result in tissue-specific tumor types (Fig. 2). There are clearly genetic, epigenetic and environmental modifiers of the TP53 gene and p53 protein functions.

What is the relationship between epigenetic alterations and the p53 protein functions and responses?
The first line of evidence that large epigenetic changes were one of the intrinsic stresses that are sensed by and activate p53 was demonstrated by Jaenisch and his colleagues [61]. They created a conditional knockout of the DNA methyl-transferase gene-1 (Dnmt1) in a CRE-Lox mouse. The Dnmt1 enzyme is a cytosine methyltransferase that adds a methyl group to cytosines in new strands of DNA at GpC locations opposite methyl-CpG residues in spectrum in mice harboring a p53 hot spot mutation compared to those with the SNP309 T allele [56].
its protein appear to be both the guardian of the genome (Fig. 3) that interact with the genome remains unclear. The chromatin that packages the genome, or the transcription factors that are 80% and the kinetics of producing IPSC occurred in days, not weeks or months, all in a temperature-sensitive fashion, proving it was p53 that regulated these processes [63, 64]. These experiments demonstrate that the p53 protein can sense a dramatic change in the levels of genomic methylated CpG residues, and in response, kill these cells by apoptosis. One of the stresses that activate p53 is a change in genomic epigenetic marks. Azacytidine and decitabine are incorporated in place of cytidine in DNA but do not permit methylation of the cytosine residues. Curiously, these two drugs are much more efficient in killing cells withTp53 mutations than they are in killing cells with wild-type TP53 [65]. This is also true for cells with mutant and wild-type TP53 that are tumorigenic in animals [66] and humans [67].

How p53 senses changes in the epigenome, and perhaps in the chromatin that packages the genome, or the transcription factors that interact with the genome remains unclear. The TP53 gene and its protein appear to be both the guardian of the genome (Fig. 3) and of the epigenome, ensuring fidelity with a penalty of death. That is a regulatory mechanism that a multicellular organism can live with. It seems likely that changes in CpG methylation in the genome change with age, are predictive of a healthy longevity and differ between males and females [68]. There are numerous suggestions in the literature that histone methyltransferases and demethylases play a role in the sexual dimorphism that impacts life span, cancer incidences, autoimmune, immune responses to infectious diseases, metabolic differences and even cognitive differences [69–71]. We are just starting to learn about the mechanisms that regulate this sexual dimorphism [72]. It is clear that p53 interfaces with every one of these pathways and phenotypes, perhaps through epigenetic marks, and a better understanding of these processes is our next challenge.

REFERENCES

1. Weissman I. Stem cells are units of natural selection for tissue formation for germline development and cancer development. Proc Natl Acad Sci USA. 2015;112:8922–8.
2. Sato T, Stange DE, Ferrante M, Vries RG, Van Es JH, Van den Brink S, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett’s epithelium. Gastroenterology. 2011;141:1762–72.
3. Blauman C, Fuchs E. Epidermal stem cells of the skin. Annu Rev Cell Developmental Biol. 2006;22:339–73.
4. Barker N, Tan S, Cleverson H. Lgr proteins in epithelial stem cell biology. Development 2013;140:2484–94. https://doi.org/10.1242/dev.083113.
5. Dubrovsky G, Dunn JC. Mechanisms for intestinal regeneration. Curr Opin Pediatrics. 2018;30:424–9.
6. Takeda H, Wei Z, Koso H, Rust AG, Yew CC, Mann CB, et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. Nat Genet 2015;47:142–50.
7. Oliver M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences and clinical use. Cold Spring Harb Perspect Biol 2010;2:a001008.
8. Hainaut P, Pfeifer GP. Somatic TP53 mutations in the era of genome sequencing. Cold Spring Harb Perspect Med. 2016;6:a026179.
9. Levine AJ, Jenkins NA, Copeland NG. The roles of initiating truncal mutations in human cancers: the order of mutations and tumor cell type matters. Cancer Cell. 2019;35:10–15.
10. Levine AJ. p53: 800 million years of evolution and 40 years of discovery. Nat Rev Cancer. 2020;20:471–80. https://doi.org/10.1038/s41568-020-0262-1.
11. Cancer Genome Atlas Research Network Weinstein JN, Collisson EA, Mills GB, Shmulevich I, Kahn SR, et al. The cancer genome atlas pan-cancer analysis project. Nat Genet 2013;45:1133–20.
12. Haigis KM, Cichowski K, Elledge SJ. Tissue-specificity in cancer: the rule, not the exception. Science. 2019;363:1150–1.
13. Salk JJ, Loubet-Senek A, Markus M, Valentine CC, Williams LN, Higgins JE, et al. Ultra-sensitive TP53 sequencing for cancer detection reveals progressive clonal selection in normal tissue over a century of human lifespan. Cell Rep. 2019;28:132–44.
14. Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, Takeuchi Y, et al. Age-related remodelling of oesophageal epithelia by mutated cancer drivers. Nature. 2019;565:312–7.
15. Martinecora I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. Tumor evolution high burden and pervasive positive selection of somatic mutations in normal human skin. Science. 2015;348:880–6.
16. Williams MJ, Zapata L, Werner B, Barnes CP, Sottovia A, Graham TA. Measuring the distribution of fitness effects in somatic evolution by combining clonal dynamics with dN/dS ratios. Elife. 2020;9:e48714. https://doi.org/10.7554/eLife.48714. PMID: 32223898; PMCID: PMC7105384.
17. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science. 2019;366:aeea6673. https://doi.org/10.1126/science.aeea6673. PMID: 31672865; PMCID: PMC6050831.
18. Bowman RL, Busque L, Levine RL. Clonal hematopoiesis and evolution to hematopoietic malignancies. Cell Stem Cell. 2018;22:157–70.
19. Silver AJ, Jaiswal S. Clonal hematopoiesis: pre-cancer PLUS. Adv Canc Res. 2019;141:85–128.
20. Baker SJ, Markowitz S, Fearon ER, Willson JK, Vogelstein B. Suppression of human colorectal carcinoma cell growth by wild-type p53. Science. 1990;249:912–5.
21. Baker SJ, Preisinger AC, Jessup JM, Parasekova C, Markowitz S, Willson JK, et al. p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigencis. Cancer Res. 1990;50:717–22.
22. Matano M, Date S, Shimokawa M, Takanao A, Fujii M, Ohta Y, et al. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. Nat Med. 2015;21:256–62.
23. Takeda H, Wei Z, Koso H, Rust AG, Yew CC, Mann MB, et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. Nat Genet 2015;47:142–50.
24. Oliver M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences and clinical use. Cold Spring Harb Perspect Biol 2010;2:a001008.
25. Hainaut P, Pfeifer GP. Somatic TP53 mutations in the era of genome sequencing. Cold Spring Harb Perspect Med. 2016;6:a026179.
26. Levine AJ, Chan CS, Dudgeon C, Puzio-Kuter A, Hainaut P. The evolution of colorectal cancer using CRISPR-Cas9 mediated engineering of human intestinal organoids. Nat Med. 2015;21:256–62.
27. Schneider K, Zellek K, Nichols K, Garber J. Li-Fraumeni Syndrome. NCBI Bookshelf, a service of the National Library of Medicine NIH. https://www.ncbi.nlm.nih.gov/bookshelf.
28. Lin J, Chen J, Elenbaas B, Levine AJ. Several hydrophobic amino acids in the p53 amino-terminus domain are required for transcriptional activation, binding to mdm-2 and the adenovirus S1 E1B 55-kD protein. Genes Dev. 1994;8:1235–1246.
29. Mello SS, Attardi LD. Deciphering p53 signaling in tumor suppression. Curr Opin Cell Biol. 2018;51:65–72. https://doi.org/10.1016/jceb.2017.11.005. (e-pub ahead of print) 2017 Nov 28. PMID: 29195118; PMCID: PMC5949255.
30. Walker KK, Levine AJ. Identification of a novel p53 functional domain that is necessary for efficient growth suppression. Proc Natl Acad Sci USA. 1996;93:1333–40. https://doi.org/10.1073/pnas.93.26.13335. PMID: 8986812; PMCID: PMC26405.
31. Baptiste N, Friedlander P, Chen X, Prives C. The p53-rich domain of p53 is required for cooperation with anti-neoplastic agents to promote apoptosis of tumor cells. Oncogene. 2002;21:9–21. https://doi.org/10.1038/sj.onc.1205015. PMID: 11791172.
32. Oliver M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P. The IARC TP53 database: new online mutations analysis and recommendations to users. Hum Mutat. 2002;19:607–14.
33. Oliver M, Hussain SP, Caron de Fromentel C, Hainaut P, Harris CC. TP53 mutation spectra and load: a tool for generating hypotheses on the etiology of cancer. IARC Sci. Public. 2006;247–70.
35. Baugh EH, Ke H, Levine AJ, Bonneau RA, Chan CS. Why are there hotspot mutations in the TP53 gene in human cancers? Cell Death Differ. 2018;15:54–60. https://doi.org/10.1038/cdd.2017.180. e-pub 2017 Nov 3. PMID: 29099487; PMCID: PMC5729536.

36. Jeffrey PD, Gorina S, Pavletich NP. Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 anstroms. Science. 1999;285:1498–502. https://doi.org/10.1126/science.285.5431.1498. PMID: 10101119; PMCID: PMC788469; MEDLINE: 10101119.

37. Pinto EM, Figueiredo BC, Chen W, Galvao HR, Formiga MN, Fragoso MCBV, et al. XAF1 as a modifier of p53 function and cancer susceptibility. Sci Adv. 2020;6:eaba2331 https://doi.org/10.1126/sciadv.aba2331. PMID: 32637605; PMCID: PMC7314530.

38. Levine AJ, Berger SL. The interplay between epigenetic changes and the p53 protein in stem cells. Genes Dev. 2017;31:1195–201. https://doi.org/10.1011/gad.20098941.1117. PMID: 28765161; PMCID: PMC5558922.

39. Zhu J, Dou Z, Sammons MA, Levine AJ, Berger SL. Lysine methylation represses p53 activity in teratocarcinoma cancer cells. Proc Natl Acad Sci USA 2016;113:9822–7. https://doi.org/10.1073/pnas.1603871113. PMID: 27539333; PMCID: PMC4205458.

40. de Andrade KC, Mirabello L, Stewart DR, Kung CP, Good M, Leu JI, et al. Tailoring genetic variation in the TP53 pathway and their impact on Cancer. J Biol Mol. 2019;11:578–85. https://doi.org/10.1016/j.jbromol.2019.01.013. PMID: 30793234; PMCID: PMC6682780.

41. de Andrade KC, Frone MN, Wegman-Ostrosky T, Khincha PP, Kim J, Amadou A, et al. Population genetics of the TP53 pathway and somatic cell reprogramming and stem cell differentiation. Cancer Res 2016;76:2023–36. https://doi.org/10.1158/0008-5472.CAN-15-2931. PMID: 27059731; PMCID: PMC52317532.

42. Levine AJ. Targeting therapies for the p53 protein in cancer treatments. Annu Rev Cancer Biol. 2019;3:1.1–1.14. https://doi.org/10.1146/annrevcancerbio-030518-54545.

43. Yi L, Yang Y, Levine AJ. Selected drugs that inhibit DNA methylation can preferentially kill p53 deficient cells. Oncotarget. 2014;5:8924–36. https://doi.org/10.18632/oncotarget.2441. PMID:142546370.

44. Welch JS, O’Malley CM, Frick RH, Lu L, Capparelli C, et al. TP53 and the improvement of this review. The ideas, some of the experiments and results discussed in this review were supported by a grant from the NIH, NCI PO1- to A. J. Levine.

ACKNOWLEDGEMENTS

The author thanks S Christen for help in the preparation of this review and Drs. K.C. de Andrade, M. Murphy and G. Zambetti for their critical reading and suggestions for the improvement of this review. The ideas, some of the experiments and results discussed in this review were supported by a grant from the NIH, NCI PO1- to A. J. Levine.
AUTHOR CONTRIBUTIONS
AJL is the sole author. He wrote the manuscript, reviewed the literature, and added some new ideas.

COMPETING INTERESTS
AJL is a founder, board member and shareholder in PMV Pharmaceuticals that produce small molecule reactivators of p53 mutant proteins. He is on the board and holds stock options in Meira GTX (retinal gene therapy) and Genecentric (RNA diagnostics) and receives fees for advice to Pharmabody (monoclonal antibodies). He is a member of the SAB of InterVenn Biosciences, and Janssen Pharmaceuticals.

ADDITIONAL INFORMATION
Correspondence and requests for materials should be addressed to A.J.L.
Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021