Cell

Lessons from the Cancer Genome

Levi A. Garraway[1,2,4] and Eric S. Lander[3,4,5,*]
[1]Department of Medical Oncology and Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[2]Department of Medicine, Brigham and Women's Hospital
[3]Department of Systems Biology
Harvard Medical School, Boston, MA 02115, USA
[4]The Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[5]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: lander@broadinstitute.org
http://dx.doi.org/10.1016/j.cell.2013.03.002

Systematic studies of the cancer genome have exploded in recent years. These studies have revealed scores of new cancer genes, including many in processes not previously known to be causal targets in cancer. The genes affect cell signaling, chromatin, and epigenomic regulation; RNA splicing; protein homeostasis; metabolism; and lineage maturation. Still, cancer genomics is in its infancy. Much work remains to complete the mutational catalog in primary tumors and across the natural history of cancer, to connect recurrent genomic alterations to altered pathways and acquired cellular vulnerabilities, and to use this information to guide the development and application of therapies.

Introduction

More than a century ago, Theodor Boveri proposed that cancer is caused by chromosomal derangements that cause cells to divide uncontrollably (Boveri, 2008)—that, in modern terms, cancer is a "disease of the genome." It took 70 years for molecular biologists to prove this concept by showing the existence of mutated cancer-causing genes (Stehelin et al., 1976; Tabin et al., 1982). By the mid-1980s, researchers had established two main types of cancer-causing genes (oncogenes and tumor suppressor genes) and had defined the genomic alterations that give rise to them (e.g., nucleotide substitutions, chromosomal copy number alterations, and DNA rearrangements; reviewed in Macconaill and Garraway [2010]). These studies also began to suggest considerable complexity in the mutational origins of cancer, with cancer-causing genes varying across and within tumor types and with multiple genes contributing to tumorigenesis.

In an influential commentary in 1986, Renato Dulbecco argued that the complete sequence of the human genome would be an essential tool for systematically discovering the genes that drive cancer (Dulbecco, 1986). "If we wish to learn more about cancer, we must now concentrate on the cellular genome," he wrote. "We have two options: either to try to discover the genes important in malignancy by a piecemeal approach, or to sequence the whole genome…it will be far more useful to begin by sequencing the cellular genome." Responding to this and other calls, the Human Genome Project (HGP) was launched in 1990. A "draft" sequence was completed by 2000 (Lander et al., 2001; Venter et al., 2001) and a near-complete sequence by 2003 (IHGSC, 2004).

With the availability of the genome sequence, cancer researchers rapidly began to develop a new field of "cancer genomics." Cancer genomics involves systematic studies of (some or all of) the genome to find sites of recurrent derangement in specific cancer types. Pioneering genomic studies at the Sanger Institute and Johns Hopkins uncovered genes mutated frequently in melanoma and colon cancer, respectively (Davies et al., 2002; Samuels et al., 2004). Studies by several groups in Boston and New York then discovered frequent activating mutations in lung cancer, which largely explained patient response to a drug (Lynch et al., 2004; Paez et al., 2004; Pao et al., 2004). Soon thereafter, a working group of the U.S. National Cancer Institute proposed a "Human Cancer Genome Project" (see http://www.genome.gov/Pages/About/NACHGR/May2005NACHGRAgenda/ReportoftheWorkingGrouponBiomedicalTechnology.pdf), which came to be called The Cancer Genome Atlas (TCGA). The NCI launched a pilot project for TCGA in 2006 and a full project in 2009. In parallel, an International Cancer Genome Consortium was launched and has grown to involve researchers in more than 15 countries (Hudson et al., 2010).

The notion of taking a genomic approach to characterizing cancer was not universally endorsed, as reflected in the title of one commentary: "Human Cancer Genome Project: Another Misstep in the War on Cancer" (Gabor Miklos, 2005). Some thoughtful critics felt that hypothesis-driven research was the best way to study cancer and worried that systematic studies were so expensive that they would drive out focused investigation (Weinberg, 2010). Proponents argued that science requires investment in both hypothesis generation and hypothesis testing and that unbiased genomic studies were an excellent way to find surprises. They also expected the cost of genomic studies to plummet, with new technologies just over the horizon. Some scientists were skeptical because they believed that there were few cancer-related genes left to discover, whereas others thought that cancers were too hopelessly complicated to yield

to systematic analysis. This open debate helped to shape the design of cancer genome projects. In the end, however, the questions could only be answered with data.

With cancer genome projects now underway for several years (Table 1), the time is right to assess the early returns and to consider next steps for the field. (As a complement to this Review, we recommend an earlier review by Stratton et al. [2009], which describes many foundational aspects of cancer genomics.) Here, we describe the remarkable tapestry of biological, evolutionary, and therapeutic insights that have emerged from systematic cancer genome characterization. At the end, we suggest the next steps for cancer genomics.

Technology Revolution

Initial cancer genome projects had to be carried out with what today seem like primitive technologies. Mutations were identified by traditional capillary-based sequencing in which each exon to be studied was amplified and sequenced individually, and chromosome copy number alterations were surveyed with DNA microarrays. DNA rearrangements could hardly be cataloged at all. The high cost and extensive infrastructure needed for large-scale DNA sequencing placed tight constraints on the amount of data that could be collected. Exome-scale projects could only be carried out on small numbers of samples (Sjöblom et al., 2006; Wood et al., 2007); thus, much effort was spent developing lists of candidate gene for sequencing based on a priori notions of cancer mechanisms or therapeutic targets (Ding et al., 2008; Greenman et al., 2007; CGARN, 2008).

The emergence of massively parallel sequencing (MPS) revolutionized the entire enterprise (Bentley et al., 2008; Margulies et al., 2005). Initially, MPS made it possible to sequence nearly 1 billion bases (1 gigabase [Gb]) in a single run; this number grew to >600 Gb/run by 2012. In parallel, methods were developed that employ hybridization to oligonucleotide "baits" in aqueous solution to capture specific portions of the genome—most importantly, the ~2% of genomic DNA that contains known exons (the "exome") (Gnirke et al., 2009; Hodges et al., 2007). MPS also made it possible to use a single technology platform for all categories of genome analysis (discovering point mutations, assessing copy number alterations and translocations, measuring transcript levels, identifying alternative splicing, detecting DNA methylation, and mapping chromatin structure).

The first whole cancer genome sequenced by MPS was reported in 2008 (Ley et al., 2008). Whereas initial studies were confined to single samples (Pleasance et al., 2010a, 2010b), studies of hundreds of samples have quickly become the norm. Plummeting costs have propelled an unprecedented explosion of sequence data. For example, >16,000 cancer samples had been subjected to genome or exome sequencing by late 2012 just at our institution alone (Broad Institute).

With the MPS data came a need for completely new analytic tools. The first challenge was to accurately determine the sequence in individual tumor and normal samples from the "raw" sequence data. Each type of alteration in the DNA and RNA required a specialized detection method, including for single nucleotide variants, small insertions/deletions, chromosomal rearrangements, gene fusions, alternatively spliced transcripts, chromosomal copy number alterations, and detection of foreign DNA (such as from viruses) (Beroukhim et al., 2007; Chen et al., 2009; Cibulskis et al., 2013; Dees et al., 2012; Kim and Salzberg, 2011; Kostic et al., 2011; Trapnell et al., 2009). Algorithms employ probabilistic methods to identify mutations

Table 1. Current Large-Scale Cancer Genome Projects[a]

Anatomic Site	Tumor Type
Brain/Central nervous system	glioblastoma multiforme
	low-grade glioma
	pediatric: medulloblastoma
	pediatric: pilocytic astrocytoma
Head and neck	head/neck squamous cell cancer
	thyroid carcinoma
Thoracic	lung adenocarcinoma
	lung squamous cell carcinoma
Breast	breast lobular carcinoma
	breast ductal carcinoma
	breast triple-negative
	breast HER-2 positive
	breast ER positive vs. negative
Gastrointestinal	esophageal adenocarcinoma
	esophageal squamous carcinoma
	gastric adenocarcinoma
	gastric (intestinal/diffuse)
	hepatocellular (alcohol/adiposity)
	hepatocellular (virus)
	hepatocellular (general)
	pancreatic adenocarcinoma
	colorectal adenocarcinoma
	colon cancer (non-Western)
Gynecologic	ovarian serous cystadenocarcinoma
	endometrial carcinoma
	cervical cancer (squamous + adeno)
Urologic	renal: clear cell carcinoma
	renal: papillary carcinoma
	renal: chromophobe carcinoma
	bladder cancer
	prostate adenocarcinoma
	prostate adenocarcinoma, early onset
Skin	melanoma, cutaneous
Soft tissue (Sarcoma)	solitary fibrous tumors
	desmoid tumors
	angiorsarcomas
	leiomyosarcomas
	extraskeletal myxoid chondrosarcomas
Hematologic	acute myeloid leukemia
	lymphoma: chronic lymphocytic leuk.
	lymphoma: germinal B cell
	lymphoma: diffuse large B cell
	chronic myeloid disorders

[a]In conjunction with The Cancer Genome Atlas, International Cancer Genome Consortium, and Slim Initiative for Genomic Medicine.

or rearrangements based on their presence in multiple tumor sequence reads and absence in the paired normal DNA sequence (Meyerson et al., 2010).

Detecting mutations with high accuracy turned out to be surprisingly tricky. Because somatic point mutations in cancer are so infrequent (~1/Mb), the background error rate must be an order of magnitude lower to ensure that most apparent events are true positives. Many false positives initially arose from sequencing errors and inaccurate alignment of reads to the human genome. In addition, false negatives may arise from admixture of noncancer cells (tumor purity), copy number variations inherent in cancer genomes (ploidy), and the presence of variant subclones within the cancer cell population (heterogeneity) (Carter et al., 2012). Increasing the sequencing depth (average number of reads per base) was found to improve both the specificity and sensitivity of mutation calling. Currently, tumor sequencing is performed with 100- to 150-fold coverage for whole-exome analysis and 30- to 60-fold coverage for whole-genome analysis. (The whole-genome sequence should ideally be deeper but is currently limited by cost.)

Obtaining accurate mutation calls for a collection of individual samples is only the first step. The harder challenge is to distinguish between "driver" events that are causally related to the development of cancer and random "passenger" events that have simply accumulated over the course of development and cell growth. This requires determining which genes show significantly more mutations than random expectation. Sophisticated mathematical methods are needed to ensure that the "random expectation" properly accounts for (1) variation in background mutation rates across the genome, (2) variation across tumors, and (3) variation in purity and heterogeneity (Chapman et al., 2011a; Dees et al., 2012; Hodis et al., 2012; M.S. Lawrence, personal communication). Without such corrections, genes may be spuriously declared to be drivers, with the problem growing worse as sample size grows (because even modest deviations from expectation will appear to be significant). Recent studies have highlighted some likely spurious results and have developed solutions to eliminate them (G. Getz, personal communication). Perfecting these algorithms remains an area of active research. Other algorithms have been developed to study the structure of amplifications and deletions to detect the possibility of multiple target genes within a given locus (Beroukhim et al., 2007; Beroukhim et al., 2010).

Mutational Mechanisms

The explosion of genomic data quickly shed light on the mutational processes of cancer, revealing an unexpected richness of mechanisms. These insights may propel a deeper understanding of factors governing genome integrity and tumor evolution.

Mutation Rates

Initial plans for cancer genome analysis assumed a single uniform background mutation rate (~1/Mb). In fact, cancer mutation rates turned out to be much more variable, ranging from as low as one base substitution per exome (<0.1/Mb) in some pediatric cancers to thousands of mutations per exome (~100/Mb) in certain mutagen-induced malignancies (such as lung cancer and melanoma). Moreover, mutation rates were found to vary substantially across the genome, governed by processes such as transcription-coupled repair and replication timing.

Mutational Spectra

Cancer genome sequencing has also revealed a wide array of mutational patterns both across and within individual tumor types. Their distinctive characteristics may reflect extrinsic factors (e.g., UV light or tobacco smoke) or intrinsic patterns such as DNA repair deficiencies. For example, a recent study (Nik-Zainal et al., 2012a) pointed to at least five distinct nucleotide substitution patterns, most of which occur by as-yet unknown mechanisms. One such process, which produces C > A, C > G, or C > T substitutions at TpCpX trinucleotides, appeared to underpin most nucleotide substitutions in ~10% of ER-positive breast tumors. These studies also discovered a new regional hypermutation mechanism characterized by multiple base mutations that occur in *cis* near rearrangement breakpoints. Termed "kataegis" (Greek, "kataegis" = "shower" or "thunderstom"), this process likely involves the activation-induced deaminase (AID) and apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) protein families. New mutation patterns (in this case, A > C transversions at "AA" dinucleotides) have also been discovered in esophageal cancers by large-scale sequencing (A.M. Dulak, personal communication).

Chromosomal Gains and Losses

Although tumor cell aneuploidy has long been recognized, global cancer genome studies have yielded a systematic assessment of large-scale (whole chromosome or chromosome arm) and focal copy number aberrations. The typical cancer cell exhibits large-scale gains or losses involving a quarter of its genome and carries focal events affecting 10% (Beroukhim et al., 2010). Based on current sample collections, many focal amplifications and deletions have been localized to "peak" regions containing a median of 6–7 genes (although the number is 150–200 in some cases). For the majority of focal events, the driver gene(s) still cannot be assigned definitively.

Chromosomal Shattering

One of the most striking mutational patterns unveiled by whole-genome sequencing studies consists of a catastrophic phenomenon that produces dozens or even hundreds of rearrangements. The resulting disarray is distinctive for two reasons: it is typically localized within one or a few chromosomes, and it usually involves only two distinct copy number states (Stephens et al., 2011). Termed "chromothripsis" (Greek, "thripsis" = "shattering"), this process occurs in ~2%–3% of human cancers, with an elevated prevalence in bone cancers, pediatric medulloblastoma, and neuroblastoma (Molenaar et al., 2012; Rausch et al., 2012). Genomic shattering appears to develop as a result of erroneous chromosome segregation during mitosis and the subsequent entrapment of individual chromosomes within "micronuclei" (Crasta et al., 2012). Micronuclei have a tendency toward premature chromosome condensation, which may result in pulverization of chromosomal segments. Chromosomes that survive this process, having undergone aberrant reassembly through nonhomologous end-joining, emerge with dense rearrangements that may sometimes dysregulate cancer genes (Stephens et al., 2011).

Chromosomal Chains

A whole-genome sequencing study of primary human prostate cancer (Berger et al., 2011) uncovered a distinct category of complex chromosomal rearrangements. Prostate cancer genomes often exhibit "chains" of copy-neutral rearrangements that consist of ~4–12 distinct breakpoint junctions distributed across multiple chromosomes, with the breakpoints forming a "closed chain" (A to B, B to C, C to D, and finally back to A) that distinguishes the process from chromothripsis or other complex rearrangements. Closed chain rearrangement breakpoints tend to occur near "open" chromatin (that is, transcriptionally active chromatin) in prostate cancer genomes harboring ETS transcription factor rearrangements but near "closed" chromatin in certain "ETS-negative" prostate cancers. These chains have recently been termed "chromoplexy" (Greek, "plexy" = "weave" or "braid") (Baca et al., 2013).

Additional Processes

Other complex DNA rearrangements seem to arise through errors in DNA replication (Liu et al., 2011). These may include fork-stalling and template-switching events that trigger microhomology-dependent DNA priming, duplications, and DNA template insertions (for a recent review, see Holland and Cleveland [2012]). Interestingly, these replication-dependent rearrangements show a strong correlation with *TP53* mutations in subtypes of medulloblastoma (Rausch et al., 2012). Thus, somatic alterations in DNA-damage-sensing pathways may render tumor progenitor cells vulnerable to ensuing catastrophic genomic events.

Insights into mutational patterns may bring a deeper understanding of tumor evolution. In contrast to a simple gradualist notion in which somatic mutations accumulate steadily, tumor evolution can be punctuated by various types of catastrophic events (Baca et al., 2013). A fuller knowledge of mutational processes—particularly those that preferentially enact cancer genes—may help to identify driver mechanisms in tumors.

New Cancer Genes

A key question is whether cancer genomics has led to the discovery of new genes and, ideally, to new *classes* of genes not previously known to play a causal role in cancer. Although much work still lies ahead, the answer is clear. The trickle of biological discoveries from early studies has become a wave, implicating a wide range of cellular processes in cancer. Whereas some of the new cancer genes encode classical signaling proteins, most populate new and sometimes surprising categories, such as metabolism, epigenetics, chromatin biology, splicing, protein homeostasis, and cell differentiation (Table 2). The insights from these studies are already guiding hypothesis-driven cancer research ranging from basic cell and molecular biology to novel therapeutics.

Signal Transduction Pathways

Studies from the 1980s and 1990s revealed that signaling pathways linked to proliferation and survival played a crucial role in many cancers. Mutations were discovered in key genes that encode members (or regulators) of receptor tyrosine kinase (RTK)-signaling pathways (HER-2, c-KIT, ABL, RAS, NF1, NF2, MET, PTEN), the WNT/β-catenin pathway (APC), and the TGF β pathway (SMAD2 and SMAD4), among others. Moreover, the pharmaceutical industry showed that drugs could be developed to inhibit protein kinases. The poster child was imatinib (Gleevec) against the ABL and KIT kinases, which proved remarkably effective in treating malignancies driven by activating mutations in these oncoproteins (chronic myelogenous leukemia [CML] and gastrointestinal stromal tumors [GIST]) (Demetri et al., 2002; Druker et al., 2001).

Recognition of the importance and druggability of RTKs motivated the first unbiased sequencing surveys in the early 2000s, which employed Sanger sequencing to examine dozens of genes in dozens of patients. The studies quickly hit pay dirt, with the finding of mutations in *BRAF* in 50% of melanomas, *PIK3CA* in ~25%–30% of breast and colorectal cancers, *EFGR* in 10%–15% of non-small cell lung cancers, *FGFR2* in 15%–20% of endometrial cancers, and *JAK2* in myeloproliferative diseases (Davies et al., 2002; Dutt et al., 2008; Kralovics et al., 2005; Levine et al., 2005; Lynch et al., 2004; Paez et al., 2004; Pao et al., 2004; Pollock et al., 2007; Samuels et al., 2004). (Some of the findings came to have a major impact on drug development and clinical treatment, including the development of selective RAF and MEK inhibitors that have produced dramatic remissions in melanoma and the ability to target the use of EGFR inhibitors to the subset of lung cancer patients who derive benefit.)

In turn, these successes led researchers to scale up sequencing surveys to discover additional candidate genes in signaling pathways and eventually to all genes. Recurrent mutations were found in genes involved in several—sometimes surprising—pathways not previously suspected to drive cancer. These included the *MAP3K1* and *MAP2K4* genes in breast cancer (encoding serine/threonine kinases involved in the P38-JNK signaling pathway) (Banerji et al., 2012; Ellis et al., 2012; Stephens et al., 2012; CGAN, 2012), *RAC1* in melanoma (a GTPase involved in the RAC/PAK-signaling module involved in focal adhesion) (Hodis et al., 2012; Krauthammer et al., 2012), *ELMO1* and *DOCK2* in esophageal cancer (two genes that activate RAC/PAK signaling) (A.M. Dulak, personal communication), *MYD88* in diffuse large B cell lymphoma (activates NF-κB signaling) (Ngo et al., 2011), and *PREX2* in melanoma (a guanine nucleotide exchange factor that controls RAC/PAK and PI3K signaling) (Berger et al., 2012). Remarkably, a pathway involved in axon guidance in neurons (the ROBO/SLIT pathway) turned out to be a target of mutations in ~20% of pancreatic adenocarcinomas (Biankin et al., 2012). And, a pathway that governs the oxidative stress response in all cells (the KEAP1/NRF2-signaling pathway) is activated by mutation in >30% of squamous lung cancers (Hammerman et al., 2012; Shibata et al., 2008). Many of these results would have eluded hypothesis-based investigation.

In addition, genome-wide studies of copy number alterations based on DNA microarrays revealed recurrently amplified genes in signaling and cell survival pathways. *MCL1* and *BCL2L1*, which encode anti-apoptotic proteins that are critical regulators of tumor cell survival, were found to be amplified in a wide range of cancers, including breast, lung, colorectal, melanoma, and glioblastoma (Beroukhim et al., 2010). *FGFR1* was found to be amplified in >20% of lung squamous cancer (Weiss et al., 2010) and in ~10% of breast cancers (Chin et al., 2006).

Table 2. Discoveries from Cancer Genome Characterization

Cellular Process Altered by Genomic Alterations	Examples of Cancer Genes Discovered (or Extended to New Cancers*) by Genomics
RTK signaling	*EGFR*,[a] *ERBB2*,[*,a] *MET*,[*,a] *ALK*,[*,a] *JAK2*,[a] *RET*,[*,a] *ROS*,[*,a] *FGFR1*,[*,a] *FGFR2*,[a] *PDGFRA*,[*,a] and *CRKL*[a]
MAPK signaling (oncogenes)	*KRAS*,[*,a] *NRAS*,[*,a] *BRAF*, [a] and *MAP2K1*[a]
MAPK signaling (TSG)	*NF1*[*,b]
PI3K signaling (oncogenes)	*PIK3CA*,[a] *AKT1*,[a] and *AKT3*[a]
PI3K signaling (TSG)	*PTEN*[*,b] and *PIK3R1*[b]
Notch signaling (oncogene or TSG)	*NOTCH1*,[c] *NOTCH2*,[c] and *NOTCH3*[b]
TOR signaling (TSG)	*STK11*,[*,b] *TSC1*,[*,b] and *TSC2*[*,b]
Wnt/β-catenin signaling (TSG)	*APC*[*,b] and *CTNNB1*[*,a]
TGF-β signaling (TSG)	*SMAD2*,[*,b] *SMAD4*,[*,b] and *TGFBR2*[b]
NF-κB signaling (oncogene)	*MYD88*[a]
Other signaling	*RAC1*,[a] *RAC2*,[a] *CDC42*,[a] *KEAP1*,[b] *MAP3K1*,[b] *MAP2K4*,[b] *ROBO1*,[b] *ROBO2*,[b] *SLIT2*,[b] *SEMA3A*,[b] *SEMA3E*,[b] *ELMO1*,[d] and *DOCK2*[d]
Epigenetics DNA methylation	*DNMT3A*[b]
Epigenetics DNA hydroxymethylation	*TET2*[b]
Chromatin histone methyltransferases	*MLL*,[*,b] *MLL2*,[b] *MLL3*,[b] *EZH2*,[c] *NSD1*,[b] and *NSD3*[b]
Chromatin histone demethylases	*JARID1A*,[b] *UTX*,[b] *KDM5A*,[b] and *KDM5C*[b]
Chromatin histone acetyltransferases	*CREBP* [b] and *EP300*[b]
Chromatin SWI/SNF complex	*SMARCA1*,[*,b] *SMARCA4*,[b] *ARID1A*,[b] *ARID2*,[b] *ARID1B*,[b] and *PBRM1*[b]
Chromatin other	*CHD1*,[b] *CHD2*,[b] and *CHD4*[b]
Transcription factor lineage dependency or oncogene	*MITF*,[a] *NKX2-1*,[a] *SOX-2*,[a] *ERG*,[a] *ETV1*,[a] and *CDX2*[a]
Transcription factor other	*MYC*,[*,a] *RUNX1*,[b] *GATA3*,[b] *FOXA1*,[b] *NKX3.1*,[b] *SOX9*,[a] *NFE2L2*,[a] and *MED12*[d]
Splicing	*SF3B1*,[d] *U2AF1*,[d] *SFRS1*,[d] *SFRS7*,[d] *SF3A1*,[d] *ZRSR2*,[b] *SRSF2*,[d] *U2AF2*,[d] and *PRPF40B*[d]
RNA abundance	*DIS3*[d]
Translation/protein homeostasis/ubiquitination	*SPOP*,[d] *FBXW7*,[*,b] *WWP1*,[*,b] *FAM46C*,[d] and *XBP1*[d]
Metabolism	*IDH1*[a] and *IDH2* [a]
Genome integrity	*TP53*,[*,b] *MDM2*,[a] *MSH*,[*,b] *MLH*,[*,b] and *ATM*[*,b]
Telomere stability	*TERT* promoter mutations[a]
Cell cycle (oncogene)	*CCND1*[*,a] and *CCNE1*[*,a]
Cell cycle (TSG)	*CDKN2A*,[*,b] *CDKN2B*,[*,b] and *CDKN1B*[b]
Apoptosis regulation	*MCL1*,[a] *BCL2A1*,[a] and *BCL2L1*[a]

[a]Activating mutation or amplification.
[b]Inactivating mutation or deletion.
[c]Both activating and inactivating genomic events observed.
[d]Effect of mutations on protein function unknown.

CRKL, which encodes a signaling adaptor protein, was found amplified in a subset of lung cancers (Kim et al., 2010).

Despite their successes, these studies were sobering in revealing that the early promise of using a single kinase inhibitor (imatinib) to treat prevalent oncoprotein mutations (as in CML and GIST) was not going to be widely generalizable: most cancers lacked a highly recurrent mutation in genes encoding kinases (or other readily druggable targets) (Greenman et al., 2007). This underscored the importance of more deeply probing the cancer genome.

Metabolism

If there were any doubt that genomic approaches would reveal surprises, they should have been put to rest by a pioneering study in 2008. In this paper (which predated the maturation of MPS technology), Vogelstein and colleagues employed an impressive "brute force" approach to PCR amplify and sequence 175,471 exons from 20,661 genes (Parsons et al., 2008). They were rewarded with the discovery of highly recurrent mutations in the *IDH1* gene, which encodes the cytoplasmic metabolic enzyme isocitrate dehydrogenase, a seemingly unlikely candidate for a cancer gene (Figure 1); the mutations affected a single amino acid in the active site. Subsequent studies found that specific mutations in *IDH1* and *IDH2* (*IDH1*'s mitochondrial homolog) occurred in >70% of secondary glioblastomas, oligodendrogliomas, and high-grade astrocytomas (Parsons et al., 2008; Yan et al., 2009) and in ~15%–30% of acute myelogenous leukemias (AML) (Mardis et al., 2009). Because isocitrate dehydrogenases convert isocitrate to

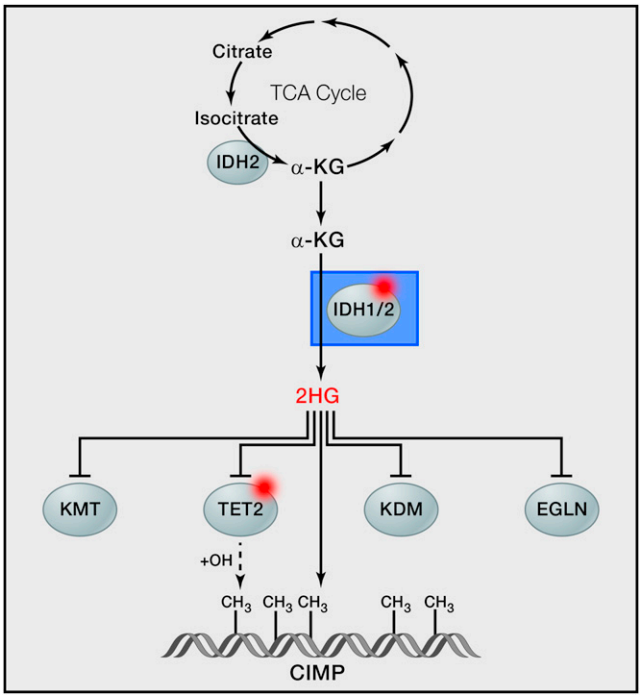

Figure 1. Somatic IDH1/2 Mutations Produce the Oncometabolite 2HG

Oncogenic effects of 2HG include generation of a CIMP-like phenotype and inhibition of α-ketoglutarate-dependent enzymes such as histone methyltransferases (KMT), histone demethylases (KDM), and prolyl hydroxylases (EGLN). *TET2* mutations are mutually exclusive with *IDH1/2* mutations in leukemias and may exert common downstream effects on DNA methylation. Mutant IDH1/2 proteins are the targets of emerging drug discovery efforts (boxed).

α-ketoglutarate (α-KG) in the tricarboxylic acid (TCA) cycle, the observation suggested a previously unrecognized link between cell metabolism and cancer. It soon became clear that the mutations caused a gain-of-function (or "neomorphic") activity, whereby isocitrate was converted to a distinct metabolite: the R-enantiomer of 2-hydroxyglutarate (2HG; Figure 1) (Dang et al., 2009; Ward et al., 2010). How this "oncometabolite" might drive cancer, however, remained a mystery.

The answer emerged from a different type of genomic analysis: genome-wide surveys of DNA methylation. The methylation studies revealed that a subset of glioblastomas (the "proneural" subtype) showed a DNA methylation pattern that strongly resembled the CpG island methylator phenotype (CIMP) originally described in colorectal cancer (Noushmehr et al., 2010). Remarkably, the CIMP-like phenotype was tightly correlated with the presence of *IDH1* mutations (Figure 1). A follow-up study confirmed that introduction of mutant IDH actually caused the CIMP-like phenotype (Turcan et al., 2012).

Unexpectedly, the mechanism was clarified by yet another genomic survey, this time involving acute myelogenous leukemia (AML). This large-scale study showed that *IDH1*/*IDH2* mutations were mutually exclusive with inactivating *TET2* mutations (Figueroa et al., 2010), suggesting that the two types of mutations had similar effects and were thus functionally redundant. The TET2 protein catalyzes 5-methylcytosine hydroxylation in a α-KG-dependent manner, and loss of TET2 produces a CIMP-like phenotype. Studies then showed that 2HG appears to inhibit several α-KG-dependent enzymes (Xu et al., 2011), including Jumonji-C domain histone demethylases that affect gene expression (Lu et al., 2012) and prolyl-4-hydroxylases (EGLN1/2/3) that regulate hypoxia inducible factor (HIF), which is involved in certain cancers (Koivunen et al., 2012) (Figure 1).

The surprising discoveries about IDH1/IDH2 have helped to spark enormous interest in cancer metabolism. They have also spawned new areas for cancer drug discovery that had little precedent prior to these cancer genome studies.

Lineage Survival Oncogene Transcription Factors

Another important discovery concerned "master" lineage-specific transcription factors (TFs). Because such TFs are typically involved in terminal differentiation of cell types, the prevailing hypothesis was that overexpression would suppress cancer by promoting lineage maturation and cell-cycle arrest. Surprisingly, however, an integrative analysis of genome-wide copy number and transcription showed that *MITF*, which encodes the master transcription factor that regulates melanocyte survival and differentiation, underwent gene amplification in a subset of metastatic melanomas (Garraway et al., 2005). *MITF* thus served as a prototype for a new category of cancer genes termed "lineage survival" oncogenes.

Systematic genomic studies subsequently uncovered several additional lineage survival oncogene TFs. Examples include *NKX2.1* in lung adenocarcinoma, *SOX2* in esophageal cancer, and *CDX2* in colorectal cancer (Bass et al., 2009; Salari et al., 2012; Weir et al., 2007). In hindsight, these TFs are analogous to the androgen receptor (AR), a nuclear hormone TF that plays crucial roles in proliferation and survival of normal and malignant prostate epithelia and is frequently amplified or mutated during tumor progression (Taplin et al., 1995). Exome-sequencing studies of castration-resistant prostate cancer have recently identified somatic mutations in both AR and several key coregulators (Grasso et al., 2012).

Epigenomics

One of the most far-reaching discoveries from genomic studies has been the critical role of epigenomic changes in tumorigenesis, which in turn has unleashed a torrent of hypothesis-driven studies and drug discovery efforts. Abnormal DNA methylation and chromatin structure were known to be common in cancers, but it was unclear whether these epigenomic changes played a *causal* role in cancer or were simply a noncausal correlate of the cancerous state. The question was settled with the recognition that ~40 genes encoding epigenomic regulators show highly recurrent somatic alterations across a wide range of cancer types (reviewed in Dawson and Kouzarides, 2012). Mutations that affect the epigenome would seem like a highly efficient mechanism to rewire cellular circuitry because they provide a way to affect multiple target genes simultaneously. The next several sections discuss various epigenomic processes related to chromatin and DNA methylation that are affected by mutations in cancer (Figure 2).

Chromatin: Histone Modification

Genomic studies have provided clear genetic evidence that dysregulation of chromatin modifiers drives many types of cancer. Recurrent mutations were found in genes encoding enzymes

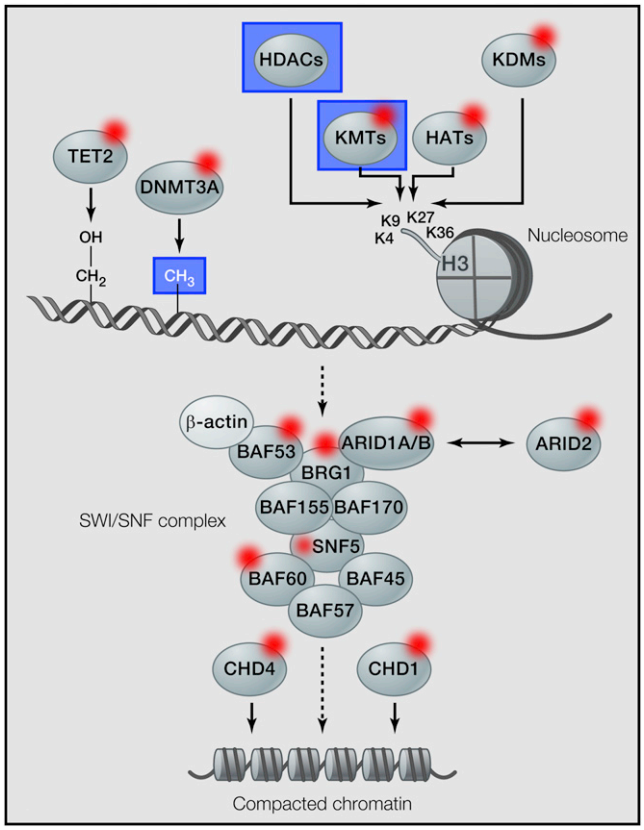

Figure 2. Genes Encoding Epigenetic and Chromatin Regulators Are Frequent Targets of Mutations in Cancer

The enzymes DNMT3A and TET2 regulate 5-methylcytosine and 5-hydroxymethylcytosine production in genomic DNA; the genes encoding these enzymes are frequently mutated in leukemias. The histone H3 component of the nucleosome undergoes extensive modifications involving its lysine (K)-rich tail. Genes encoding enzymes that read, produce, or interpret these modifications are frequently mutated in cancer. Examples include histone lysine methyltransferases (KMTs), histone lysine demethylases (KDMs), and histone acetyltransferases (HATs). Genes encoding components of the SWI/SNF chromatin-remodeling complex are also recurrently mutated in cancer. Novel therapeutics targeting chromatin and epigenetic mechanisms have entered clinical use or are in development (boxed).

that add, subtract, or interpret posttranslational modifications to histone H3. These enzymes include histone (lysine) methyltransferases (KMTs) and histone (lysine) demethylases (KDM), which activate or repress genes by modifying specific lysine residues; histone acetyltransferases (HATs), which regulate transcription by adding acetyl groups to the histone H3 tail; and histone readers, which bind various histone modifications and recruit additional protein complexes to carry out specific effector functions (Figure 2). Among the KMTs, mutations affect the MLL subfamily, which acts on lysine 4 of H3 (e.g., H3K4); the NSD subfamily, which acts on H3K36 (Dolnik et al., 2012); and EZH2, which methylates H3K27 (Morin et al., 2010). Among the KDMs, mutations affect JARID1A and UTX, which demethylate H3K4 and H3K27, respectively. Among the HATs, mutations affect *CREBP* and *EP300* (Gui et al., 2011; Morin et al., 2011; Peifer et al., 2012).

The genes encoding histone-modifying enzymes typically exhibit lineage-restricted mutational patterns. For example, the NSD1 and NSD3 KMTs have so far been found rearranged only in AML (Jaju et al., 2001; Rosati et al., 2002), and mutations affecting the histone demethylases (HDMs) KDM5A and KDM5C appear to occur exclusively in AML and renal cell cancer, respectively. However, some genes show a much broader mutational distribution. Although MLL is named for its association with a particular leukemia subtype (*m*ixed *l*ineage *l*eukemia), cancer sequencing studies have found recurrent *MLL* gene mutations in a variety of hematologic and solid tumors, including small-cell lung cancer (Peifer et al., 2012), lung squamous cancer (Hammerman et al., 2012), gastric cancer (Zang et al., 2012), head and neck cancer (Stransky et al., 2011), and prostate cancer (Barbieri et al., 2012; Grasso et al., 2012). Interestingly, the DOT1L KMT, which methylates H3K79, is not itself mutated but becomes essential in MLL-translocated leukemias (Bernt et al., 2011). HAT mutations are found in B cell lymphomas (Pasqualucci et al., 2011a), small-cell lung cancers (Peifer et al., 2012), and medulloblastoma (Robinson et al., 2012). These distinctive patterns suggest that mutations affecting chromatin-modifying enzymes contribute to cancer by disrupting expression of specific target genes that play critical roles in particular cell types. However, the identities of these target genes remain unknown, and we lack systematic methods for identifying them.

Affected tumors are typically heterozygous for apparent loss-of-function alleles, indicating that haploinsufficiency for these chromatin-modifying enzymes propels cancer and that complete loss is cell lethal. (An exception is EZH2, in which gain-of-function mutants are observed in follicular lymphoma [Morin et al., 2010], whereas loss-of-function events are seen in myeloid cancers [Jankowska et al., 2011; Makishima et al., 2010]). This makes chromatin-modifying enzymes attractive targets for anticancer drugs because cancer cells carrying only one functional gene could be uniquely sensitive to inhibitors (a synergy termed "synthetic lethality"). Vigorous drug discovery efforts are currently underway against many of the enzymes. So far, the only drugs targeting histone-modifying enzymes in clinical use are the histone deacetylase (HDAC) inhibitors. Ironically, the HDAC inhibitor vorinostat is approved for treatment of myelodysplastic syndromes and cutaneous T cell lymphomas, although HDAC genes have not been found mutated in these (or any other) malignancies.

Chromatin: Nucleosome Remodeling

Another major mutational target affecting chromatin biology is the SWI/SNF complexes (Figure 2), which regulate chromatin structure through ATP-dependent nucleosome remodeling (for a recent review, see Wilson and Roberts [2011]). The importance of these complexes in tumor biology was initially suggested by the discovery of biallelic deletions involving *SNF5* (a core SWI/SNF protein) in malignant rhabdoid tumors (an aggressive pediatric cancer). Multiple cancer sequencing surveys then revealed that the class of genes encoding SWI/SNF factors is one of the most commonly mutated targets in cancer. In renal cell cancer, 41% of tumors harbor mutations in *PBRM1*, which encodes BAF180, a histone acetylation reader and integral component of the so-called "BAF" SWI/SNF complex) (Varela et al., 2011);

only the VHL tumor suppressor is mutated more commonly in this malignancy. Similarly, >50% of ovarian clear cell carcinomas carry inactivating mutations in *ARID1A*, which encodes another BAF protein (Jones et al., 2010; Wiegand et al., 2010). Frequent *ARID1A* mutations have since been observed in many other cancer types, including up to 30% of hepatocellular carcinomas (Fujimoto et al., 2012; Huang et al., 2012a), 34% of bladder cancers, and 21% of endometrioid cancers. Its homologs *ARID1B* or *ARID2* (a component of the "PBAF" SWI/SNF complex) harbor recurrent mutations in melanoma (Hodis et al., 2012; Krauthammer et al., 2012), hepatocellular (Fujimoto et al., 2012; Li et al., 2011), and pancreatic cancers (Biankin et al., 2012). As with other histone-modifying proteins, the SWI/SNF gene mutations are typically loss-of-function alleles; they often exhibit biallelic inactivation or loss of protein expression, consistent with a tumor suppressor mechanism.

Chromatin: Compaction

Another unexpected chromatin-related target is the chromodomain-helicase-DNA-binding (CHD) gene family. CHD proteins regulate chromatin compaction during stem cell differentiation and may also promote genome stability (Ho and Crabtree, 2010) (Figure 2). Inactivating *CHD1* mutations and deletions comprise likely founder events (together with *SPOP* mutations) in a newly recognized "ETS-negative" genetic subtype of prostate cancer (Barbieri et al., 2012), where they appear to confer distinct patterns of genome derangement (Huang et al., 2012b). The homolog *CHD4* is frequently deleted in endometrial cancers (Le Gallo et al., 2012). Histone H3.3 itself contains highly recurrent hot spot mutations in pediatric astrocytoma and a subtype of medulloblastoma (Robinson et al., 2012; Schwartzentruber et al., 2012). Overall, the discovery of extensive chromatin and epigenetic mutations by unbiased cancer genome characterization has opened up vast new areas of basic and clinical discovery.

DNA Methylation

Systematic surveys have revealed that DNA methylation also plays a critical role in shaping the cancer genome (Figure 2). In particular, some cancers show a clear CpG island methylator phenotype (CIMP). The notion that DNA hypermethylation might define a biologically important cancer subtype in colorectal cancer (CRC) originated from focused studies of individual genes (Toyota et al., 1999), but other reports challenged its existence—or at least its biological relevance. Systematic interrogation of all available methylation markers (at that time) across >100 CRC samples provided definitive evidence for CIMP in CRC ("CRC-CIMP") (Weisenberger et al., 2006). Most CRC-CIMP tumors show high microsatellite instability (MSI) (Ogino et al., 2006; Weisenberger et al., 2006); this is likely due to the fact that such tumors typically have hypermethylation (and hence repression) of the MLH locus, whose loss of expression results in MSI. The etiology of CIMP in CRC remains mysterious, with these tumors showing few mutations in the DNA methylation machinery. Subsets of glioblastoma and AML were also found to have CIMP-like patterns (as described above). In these cases, the phenomenon is likely due, in part, to 2HG generated from mutant IDH1/2 proteins (Noushmehr et al., 2010), as described above.

DNA hypomethylation also plays an important role in some cancers. A whole-genome sequencing survey revealed that ~25% of AMLs carry inactivating mutations in *DNMT3A* (Ley et al., 2010), an enzyme that catalyzes the addition of methyl groups to CpG dinucleotides. AML cells with *DNMT3A* mutations show reduced DNA methylation at the promoter of many genes involved in cancer (Hájková et al., 2012); these mutations correlate with poorer overall survival (Ley et al., 2010). Subsequently, recurrent *DNMT3A* mutations were also found in the myelodysplastic syndrome (MDS) (Walter et al., 2011), a neoplastic condition that often progresses to AML.

The recognition of the key role of DNA methylation has galvanized interest in drugs that inhibit this process, such as 5-azacitidine and decitabine. Conceivably, these drugs may act in a synthetic lethal manner against tumors carrying mutations in *DNMT3A* and other genes affecting DNA methylation. Azacitidine has proved especially intriguing: it is the first drug to improve the survival of patients with myelodysplastic syndrome (MDS) and has also shown promising efficacy in AML (reviewed in Estey [2007]). *DNMT3A* mutations or other altered methylation phenotypes may define leukemic patient subpopulations that are more likely to benefit from these drugs (Marcucci et al., 2012). As with chromatin dysregulation, the critical genes affected by aberrant DNA methylation remain unclear.

DNA Hydroxyl Methylation

Genomic studies have uncovered a link between a novel epigenetic modification and cancer. In 2009, biochemical studies identified a new type of DNA modification: the conversion of 5-methylcytosine (5mC) at CpG islands to a hydroxylated variant called 5-hydroxymethylcytosine (5hmC) by the ten/eleven translocation (TET) family of DNA hydroxylases (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009) (Figure 2). Soon thereafter, genomic surveys found that a family member TET2 shows recurrent inactivating mutations in AML, MDS, and other myeloproliferative disorders (Delhommeau et al., 2009; Langemeijer et al., 2009). As noted above, the TET enzymes require α-ketoglutarate for their activity and are inhibited by the 2HG oncometabolite product of mutant IDH1/2. TET2 and IDH1/2 mutations thus act, at least in part, through a common mechanism; as would be expected, these mutations rarely co-occur in AML. Interestingly, however, *TET2* and *DNMT3A* mutations frequently co-occur in MDS, pointing to an as-yet unexplained cooperativity between dysregulation of 5mC and 5hmC in leukemogenesis.

RNA Splicing

Complementing the targets above affecting RNA transcription, cancer sequencing uncovered other important targets involved in RNA splicing (Figure 3). Though it had long been known that cancers showed aberrant splicing patterns, it was impossible to know whether these events played a causal role in cancer or were simply an effect of cancer.

The answer became clear with exome-sequencing studies in chronic myelogenous leukemia (CLL) and myelodysplastic syndromes (MDS). In CLL, the spliceosome gene *SF3B1* is mutated in 10%–15% of cases, and other spliceosomal genes, such as *SFRS1*, *SFRS7*, and *U2AF2*, are also mutated at lower frequencies (Puente et al., 2011; Quesada et al., 2012; Wang et al., 2011a). In MDS, the spectrum is even more striking: 45%–85% of cases harbor mutations in a spliceosome gene, with *SF3B1* and *U2AF1* being the most common and other

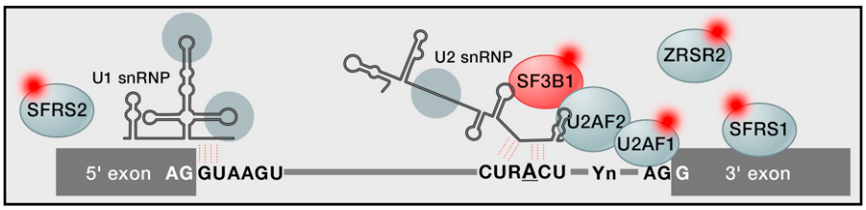

Figure 3. Cancer-Associated Mutations in the RNA-Splicing Machinery
Genes encoding spliceosomal components are recurrently mutated in both hematologic malignancies and solid tumors. Drugs that target SF3B1 have entered clinical trials.

genes (such as *SF3A1*, *ZRSR2*, *SRSF2*, and *U2AF2*) occurring at lower frequencies (Papaemmanuil et al., 2011; Yoshida et al., 2011). Spliceosomal genes have also been found significantly mutated in solid tumors—most notably, *U2AF1*, *SF3B1*, *U2AF2*, and *PRPF40B* mutations in lung adenocarcinomas (Imielinski et al., 2012). *SF3B1* is also recurrently mutated in breast cancer (Ellis et al., 2012) and pancreatic cancer (Biankin et al., 2012).

The pattern of mutations in the spliceosomal genes contains important clues about their function. First, the mutations tend to occur in a mutually exclusive fashion in all tumor types examined, suggesting that they play similar roles and are thus functionally redundant with respect to causing cancer (for a recent review, see Lindsley and Ebert [2013]). Second, several of the genes carry heterozygous missense mutations affecting specific protein domains, suggesting that they confer a gain of function. *SF3B1* (encoding a member of the splicing factor 3b complex, which interacts with SF3A proteins and a snRNA species to form the U2 small nuclear ribonucleoprotein [snRNP]) has mutations affecting the carboxy-terminal HEAT domains. *U2AF1* (encoding a member of the U2 snRNP auxiliary factor, a spliceosomal component that binds the 3′ splice acceptor site within target pre-mRNAs) has mutations affecting conserved zinc finger domains. SRSF2 (a serine-arginine-rich protein that mediates U2 snRNP assembly through binding of exon-splicing enhancer elements within pre-mRNA species) also has distinct codon localizations (Yoshida et al., 2011). In contrast, the *ZRSR2* gene (encoding a spliceosomal adaptor protein) has mutations distributed throughout its open reading frame and has frequent nonsense mutations; the pattern is indicative of loss-of-function mutations. What is missing, of course, is knowledge of the specific aberrant cancer-related splicing events caused by these mutations.

Genotype-phenotype connections offer some additional clues. In MDS, *SF3B1* mutations occur primarily in subtypes associated with ring sideroblasts (Papaemmanuil et al., 2011; Yoshida et al., 2011), whose presence signifies defective erythrocyte maturation. This observation raises the possibility that mutated SF3B1 may cause ring sideroblast formation, at least in some MDS subtypes, by governing splicing of a key erythroid lineage differentiation factor.

Mutations in several splicing factors carry prognostic information that might influence clinical management. For example, *U2AF1* mutations have been linked to increased progression from MDS to AML, and *SRSF2* mutations correlate with the so-called chronic myelomonocytic leukemia (CMML) subtype of MDS. In CLL, SF3B1 mutations correlate with more rapid disease progression and lower overall survival (Quesada et al., 2012; Wang et al., 2011a). *U2AF1* mutations were associated with poor progression-free survival in lung adenocarcinoma (Imielinski et al., 2012).

Splicing factors were not previously considered attractive targets for anticancer therapies, but that assessment is changing (Figure 3). Indeed, several small molecules and natural products known to target the spliceosome have been reported, including spliceostatin A (SSA), a metabolite derived from *Pseudomonas* that inhibits the SF3b complex and suppresses splicing in vitro, and pladienolide, a compound produced by *Streptomyces platensis* that inhibits the SF3B1 protein directly (Kaida et al., 2007). A derivative of pladienolide called E7107 has entered phase I clinical trials and shows moderate activity in thyroid cancer (Folco et al., 2011).

Protein Homeostasis

Genome-wide and exome-wide sequencing in multiple myeloma (MM) suggested an unexpected (and still unexplained) role of protein synthesis and degradation. In MM, mutations were found at high frequencies in *DIS3*, *FAM46C*, and *XBP1* (Chapman et al., 2011a). DIS3 is an RNA exonuclease that regulates RNA abundance through the exosome complex. FAM46C is a protein whose function remains unknown but whose expression pattern is nearly perfectly correlated with that of genes encoding ribosomal proteins, eukaryotic initiation factors, and translation elongation factors (Chapman et al., 2011a). *XBP1* encodes a factor involved in the unfolded protein response; mutations in the mouse homolog cause a myeloma-like condition.

Unbiased genomic studies have also uncovered unexpected roles for the ubiquitination machinery in cancer. In prostate and endometrial cancers, mutations in *SPOP* have been observed in 8%–14% of cases (Barbieri et al., 2012; Kan et al., 2010; Le Gallo et al., 2012). *SPOP* encodes the substrate recognition component of an E3 ubiquitin ligase complex. In prostate cancer, the SPOP mutations affect highly conserved amino acid residues situated within the substrate-binding motif (MATH domain), suggesting that they abrogate normal ligase/substrate interactions. These mutations are mutually exclusive with ETS rearrangements, thereby defining a distinct genetic subtype of prostate cancer (Barbieri et al., 2012). In endometrial cancer, SPOP mutations also occur in the MATH domain but involve different amino acid residues than those seen in prostate cancer (Le Gallo et al., 2012). The distinct pattern of mutations in these two cancers suggests loss of recognition for distinct substrates, leading to their accumulation.

The ubiquitin ligase gene *FBXW7* shows recurrent mutations in endometrial, head/neck, bladder, and GI cancers but only rarely shows recurrent mutations in prostate cancer. In contrast to *SPOP*, the mutations appear to be simple loss-of-function events. The ubiquitin ligase gene *WWP1* thus far only shows recurrent mutations in liver cancer (Fujimoto et al., 2012). The

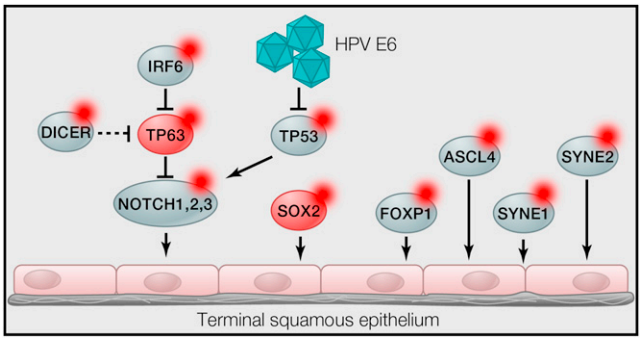

Figure 4. Genetic Alterations Affecting Lineage Specification Are Common in Squamous Tumors

NOTCH and several other lineage regulatory factors are disrupted by genomic alterations in lung, cervical, head/neck, and cutaneous squamous carcinomas. SOX2 is also a lineage survival oncogene that regulates squamous maturation. Genes that encode proteins shaded in red undergo mutational activation or amplification; those shaded gray undergo mutational inactivation or deletion.

distinct spectra of cancers seem likely to result from insufficient degradation of different proteins that are critical for different cell types. Finding the protein targets is a high priority.

Squamous Differentiation

Exome-sequencing studies in head and neck squamous cell carcinoma (HNSCC) revealed unexpected roles for pathways involved in squamous cell differentiation (Agrawal et al., 2011; Stransky et al., 2011). The studies found mutations in *NOTCH1* in ~15% of cases, as well as mutations and focal copy number alterations of *NOTCH2* and *NOTCH3* in an additional 11% (Stransky et al., 2011) (Figure 4). Whereas *activating* NOTCH1/2 mutations had been reported in various blood cancers (Lohr et al., 2012; Pasqualucci et al., 2011b; Puente et al., 2011; Weng et al., 2004), the NOTCH mutations in HNSCC were clearly *loss-of-function* events. Parallel studies in myeloid leukemia also identified recurrent loss-of-function NOTCH mutations (Klinakis et al., 2011).

The NOTCH mutations turned out to be just a part of the story. A more sophisticated analysis (of gene *sets* with recurrent mutations) pointed to genes known to be involved in epidermal development and squamous differentiation in HNSCC (Stransky et al., 2011) (Figure 4). Additional genes mutated in HNSCC (such as *RIPK2*, *EZH2*, and *DICER1*) were linked to the squamous differentiation program based on results from genetically engineered mice. Two further genes (*SYNE1* and *SYNE2*, mutated in 20% and 8% of cases, respectively) were also implicated; these genes encode proteins that control nuclear polarity and spindle orientation, which stand upstream of NOTCH signaling in squamous lineage development (Williams et al., 2011). In all, nearly one-third of HNSCC tumors appeared to harbor at least one mutation predicted to affect squamous differentiation.

Comprehensive genomic studies soon demonstrated the importance of dysregulated squamous differentiation in other tumor types. For example, inactivating NOTCH1/2 mutations occur in >75% of cutaneous squamous cell carcinomas (Wang et al., 2011b). Moreover, a study of squamous lung cancer revealed that 44% of cases harbored mutations in genes that regulate squamous differentiation (Hammerman et al., 2012). The loss of function in squamous differentiation contrasts with the *SOX2* lineage survival TF oncogene, which undergoes frequent amplification in squamous lung cancer, HNSCC, and cervical squamous cancers (http://www.cbioportal.org/public-portal/index.do).

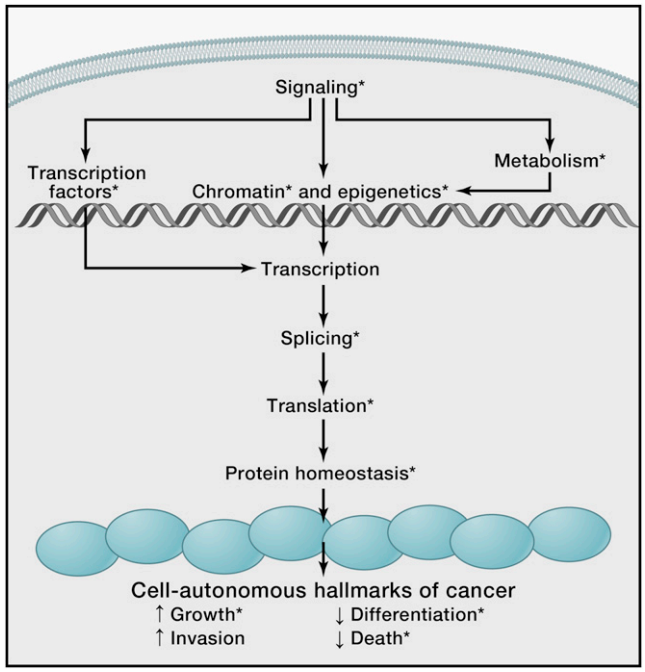

Figure 5. Genetic Alterations Disrupt Multiple Cellular Processes

Alterations in a range of cellular processes presumably contribute to cancer through their action on one or more target genes, mRNAs, or proteins, although the precise targets remain unknown in many cases (illustrated by shaded ovals). Even in advance of such knowledge, many mutations suggest potential targets for therapeutic development and allow stratification for clinical trials of targeted drugs.

Connecting the Dots: From Cancer Genes to Cancer Processes

Hanahan and Weinberg have proposed "hallmark" processes that must become dysregulated in tumorigenesis and metastasis (Hanahan and Weinberg, 2000, 2011). These processes include genome instability, unlimited cell division, sustained proliferative signaling, evasion of growth suppression, cellular energetics, and resisting apoptosis. Many classical cancer genes encode proteins that mediate or control such processes: for example, mutations in receptor tyrosine kinases or cell-cycle inhibitors can be directly understood in terms of "jamming the accelerator pedal" or "eliminating the brakes" on cell growth.

By contrast, many of the newly discovered cancer genes affect global processes whose precise connection to cancer remains obscure. These cancer genes act by deranging gene expression (through changes to chromatin and DNA methylation), RNA splicing, protein synthesis and degradation, and cellular metabolism (Figure 5). Presumably, these global changes propel cancer by affecting one or more specific targets involved in cancer processes—activating or repressing specific

genes, altering the isoforms of specific mRNAs, and increasing or decreasing steady-state levels of specific proteins. The key targets are likely cell type specific, accounting for the presence of specific subsets of driver genes in particular cancer types.

For the most part, we are ignorant of the precise targets—or whether we are looking for single targets or multiple targets. Indeed, mutations affecting global processes seemingly provide an efficient mechanism by which multiple coregulated targets might be affected. In some respects, the situation may be analogous to amplification and deletion of chromosome arms, which may provide a similarly efficient means to dysregulate multiple targets. In each case, identifying the full range of target genes will likely require unbiased genomic surveys at the DNA, RNA, and protein levels to generate hypotheses, as well as focused experiments to prove them.

Connecting the new cancer genes to known (or as-yet unknown) cancer processes will surely accelerate efforts to understand and treat cancer. Of course, therapeutic progress can be made even without a full understanding of their action. For example, inhibitors of the neomorphic IDH1/2 enzymes or perturbed splicing factors may prove valuable even without understanding the full range of enzymes affected by 2HG or SF3B1. Moreover, the set of cancer genes mutated in a tumor provides a powerful classification tool, identifying natural subtypes that can be studied in both preclinical and clinical investigation to detect distinct vulnerabilities and correlate outcomes.

Completing the Picture: Long Tails, Dark Matter, Heterogeneity, and Heredity

Genomic studies have definitely shown that our previous inventory of cancer genes was far from complete. The question now is do we finally have a near-comprehensive catalog? The honest answer: we don't know.

Long Tails

For many cancer types, a handful of cancer genes are mutated at high frequency, but many more cancer-related genes are found mutated at much lower frequencies. For example, a recent genomic study of breast cancer reported 40 loci that were mutated at statistically significant rates (Stephens et al., 2012); of these, 53% of the apparent driver mutations or focal copy number alterations were concentrated in six genes (*TP53*, *PIK3CA*, *ERBB2*, *FGFR1/ZNF703*, and *GATA3*), and the remainder were dispersed across 34 genes. Only eight of the genes were mutated in at least 10% of breast cancers. Many tumor types exhibit similar "long tail" distributions.

Some of the genes found mutated at low frequencies in some cancers are more commonly (and significantly) mutated in other cancers. In the breast cancer example mentioned above, "long tail" genes that are significantly mutated in other cancers include the SWI/SNF complex genes *ARID1A* and *ARID1B*, the KMT-encoding genes *MLL2* and *MLL3*, and *KRAS*. This finding might suggest that the discovery of new driver genes is approaching a plateau. On the other hand, the fact that so many driver genes occur at lower frequencies raises the possibility that many such genes may yet remain undiscovered. Moreover, some tumors (e.g., some primary prostate cancers) appear to lack even a single mutation in a proven driver gene.

The problem is due, in part, to the fact that most studies to date have been insufficiently powered—lacking adequate sample size to detect low-frequency events and/or adequate depth of sequence coverage to overcome impurity due to stromal contamination. Fortunately, it should be feasible to enumerate all genes carrying nonsynonymous coding mutations in at least 2% of tumors of every cancer type by sequencing a sufficiently large number of tumor-normal pairs. (Roughly 950 pairs will be needed per tumor type if the background mutation rate in the cancer is 2 mutations per Mb and 2,500 pairs if the rate is 10/Mb.) This scale seems readily achievable for many tumor types over the next several years.

Dark Matter

In contrast to point mutations in coding regions, our ability to discover and understand other types of driver mutations is still distressingly limited. Many more important cancer drivers may be lurking in the places that we cannot currently interpret. These include copy number alterations, chromosomal rearrangements, and noncoding regions.

As noted above, gains and losses spanning whole chromosome arms occur commonly in most types of cancer, but it is difficult to pinpoint the key genes for which the presence of a few extra copies contributes to cancer. Even for focal amplifications or deletions, finding the target genes can be difficult. A study of copy number alterations across cancer types found that proven cancer genes were known for less than half of recurrent focal amplifications and an even smaller proportion of recurrent focal deletions (Beroukhim et al., 2010). Incorporating sample-matched data sets can help to suggest candidates for functional validation. For example, a study in glioblastoma showed that gain of extra copies of chromosome 7 was associated with dysregulation of the HGF-MET axis (Beroukhim et al., 2007); pharmacologic experiments showed that cell lines carrying nonfocal chromosome 7 gains together with HGF and MET overexpression were preferentially dependent on MET signaling.

Chromosomal rearrangements are also pervasive in many cancers, but our ability to characterize and interpret their impact has been limited. Whereas basic cancer genome analyses can be accomplished by mapping short DNA sequences to a fixed reference sequence, comprehensive study of rearrangements requires obtaining larger-scale "linking" information to reconstruct unexpected genomic junctions and performing transcriptome sequencing to detect expressed fusion genes. These efforts have been aided by recent computational advances, such as algorithms that reconstruct transcriptomes without the need for an underlying reference genome (Grabherr et al., 2011). Most rearrangements may be random passenger events, but some clearly disrupt cancer genes by creating fusion proteins or by subjecting a gene to new regulation. Genome analysis has identified several new fusions involving known cancer genes, including RAS, RAF, *ERG*, and *PTEN*, in prostate cancer and in other malignancies (L.A.G. and E.S.L, unpublished data; Palanisamy et al., 2010; Wang et al., 2011c) and NOTCH genes in breast cancers (Robinson et al., 2011). Although relatively few instances of recurrent rearrangements implicating new cancer genes have emerged (possibly owing to limited sample sizes and the challenge of interpreting these events),

those that have been discovered may implicate new biological processes. Examples include MAST kinases in breast cancer (Robinson et al., 2011) and R-spondin family members in 10% of colon cancers (Seshagiri et al., 2012).

The great uncharted frontier is the >98% of the human genome that does not encode proteins. Our ignorance is due to two factors. First, we have lacked adequate data because cancer genome studies to date have largely focused on the exome rather than on the whole genome for reasons of cost. Second, we lack adequate analytical techniques to recognize recurrent mutations in nongenic territory. To detect a cancer-associated target, one must aggregate mutations across a defined region to test whether the rate is sufficiently elevated above background. This is straightforward for protein-coding regions, where one can aggregate nonsynonymous mutations across thousands of bases. But it is more challenging for the rest of the genome. Unbiased searches require scanning millions of small regions across the genome to find those with an unusually high mutation rate. If one searches with a small window, the mutational signal will be weak (unless the mutation frequency is very high) and detection will require large sample numbers. If one uses a large window size, the signal may be obscured by random noise in the surroundings. At present, the best approach may be to focus on regions defined by features corresponding to known biological functions, such as promoters, evolutionary conservation, and epigenomic modification.

A recent study of regulatory regions in melanoma has confirmed that important mutations may be lurking in noncoding regions (Huang et al., 2013). Whole-genome sequencing revealed the presence of highly recurrent somatic mutations at two specific nucleotides situated within the promoter of the *TERT* gene, which encodes a reverse transcriptase component of the telomerase enzyme. Both of these mutations are cytidine-to-thymidine transitions that generate a de novo binding site for the ETS transcription factor. These sites increase expression from the *TERT* promoter in reporter assays. The mutations occur in >70% of melanomas and ~16% of other tumor types examined, including bladder and hepatocellular carcinomas.

Heterogeneity

Cancer genome analyses have largely focused on tumors as a whole. Yet it has been clear for decades that tumors show extensive cellular and molecular heterogeneity. Indeed, heterogeneity was inherent in Nowell's original clonal model for tumor evolution (Nowell, 1976). Some early genomic studies have begun to come to grips with tumor heterogeneity. Initial forays have documented subclonal variation across distinct geographic regions of a primary tumor (Gerlinger et al., 2012) and within hematopoietic malignant populations (Ding et al., 2012).

Studies of heterogeneity are beginning to provide fascinating glimpses into paths of tumor evolution. For example, a study of 21 breast cancers showed that the most recent common ancestral tumor cell—which contains the full complement of mutations common to all tumor cells—arose remarkably early in "molecular time" (Nik-Zainal et al., 2012b). The precursor cell typically gives rise to a dominant subclone that represents at least 50% of all cells in the primary tumor. Knowledge of intratumoral heterogeneity has also revealed instances of convergent evolution. In one study of renal cancer, only 30%–35% of somatic mutations were concordant across multiple primary and metastatic sites sampled (Gerlinger et al., 2012); however, several cancer genes contained distinct genomic alterations that had arisen in geographically disparate regions of the primary tumor. This observation thus revealed a remarkable mutational consolidation that engaged critical pathways linked to chromatin regulation (SETD2, KDM5C) or signal transduction (PTEN, mTOR).

Tumor heterogeneity could have important implications for "precision" cancer medicine. Some subclones may contain pre-existing mutations that confer drug resistance or accelerate tumor relapse in cancers that show poor clinical responses to targeted inhibitors. Studies that seek to stratify patients for clinical trials of targeted agents based on specific "actionable" mutations may be confounded if a biopsy sample is not representative of the whole tumor. On the other hand, the ability to identify driver or resistance mutations within subclonal populations may allow improved prediction of clinical outcomes (Landau et al., 2013). The growing understanding of intratumoral heterogeneity may inform the design of clinical studies that account for this process (e.g., by following the therapeutic response of the biopsied lesion in addition to the overall tumor burden) and circumvent its subversive effects (e.g., by developing therapeutic combinations directed against major and minor subclones).

Studies of cancer heterogeneity will be accelerated by recent genomic advances enabling single-cell sequencing (Navin et al., 2011). Whole-exome sequences have been produced from single cells in both hematologic neoplasms and solid tumors (Hou et al., 2012; Xu et al., 2012). Moreover, new protocols that yield more uniform and accurate whole-genome amplification have been developed (Zong et al., 2012). Single-cell analyses have already provided new insights into the evolutionary history of tumors within individual patients and have revealed functional differences across individual tumor cells (Kreso et al., 2013). In the future, these advances may enable detailed genomic studies of circulating tumor cells, thereby providing high-resolution monitoring of therapeutic responses or emerging resistance mechanisms and facilitating detection of aggressive tumor subclones.

Heredity

Although many of the genetic factors that drive a cancer are acquired through somatic mutation, some are inherited at birth. Epidemiological studies have long noted an increased risk of cancer in relatives of affected individuals (Pomerantz and Freedman, 2011). Genomics has revealed many genes that influence predisposition to cancer, although the picture remains far from complete. Our focus in this Review is on somatic mutations, but we briefly summarize the current state of progress for inherited variation (see recent reviews by Hindorff et al. [2011] and Chung and Chanock [2011]).

One method to identify genes that confer predisposition to cancer is to study rare, highly penetrant Mendelian cancer syndromes. These syndromes arise when mutant alleles confer such a high increased risk (>10-fold) that it is straightforward to trace their transmission in families by linkage analysis. More than 100 genes underlying such cancer syndromes have been

identified, including those underlying retinoblastoma (*RB1*), breast cancer (*BRCA1*, *BRCA2*), and colon cancer (*APC*, *MUTYH* and the mismatch repair genes *MLH1*, *MSH1*, *MSH6*, and *PMS2*). Such genes have been deeply informative about cancer biology but together account for <5% of the estimated heritability of cancer (Cazier and Tomlinson, 2010).

To identify cancer genes that confer more modest risks, it is necessary to use population-based association studies rather than family-based linkage studies. The methodology for association studies depends on whether one wishes to study "common" (>1%) or "rare" (<1%) variants. Common variants are frequent enough that they can be tested for their individual effects on cancer risk by genotyping of millions of variants in cases and controls in "genome-wide association studies" (GWAS) (Altshuler et al., 2008). Rare variants must be combined together for analysis: studies examine the *aggregate* frequency of rare coding variants in each gene to look for an elevated frequency in cases versus controls. More than 150 cancer risk loci have been identified thus far, with most having been found through GWAS (Chung and Chanock, 2011; Hindorff et al., 2011). The common alleles appear to include many regulatory variants and to confer a lower increased risk (<30%), whereas the rare alleles affect coding regions of known cancer genes (such as *ATM*, *BRIP1*, *CHEK2*, *PALB2*, and *RAD51C* in breast cancer) and tend to have higher risk (2- to 3-fold). The relative roles of the two classes vary among cancer types. Importantly, the risk factors identified to date explain only a fraction of the heritability of cancer (Hindorff et al., 2011). Genomic studies with much larger samples will be needed to obtain a fuller picture of the inherited basis of cancer risk.

Understanding the mechanisms by which common inherited genetic variants predispose to cancer will require integrative genomic analysis, which will likely yield important biological insights. One instructive case is found in a 500 kb "gene desert" in chromosome 8q24. Whereas most cancer-associated loci are tumor-type specific, this region contains variants that affect risk of prostate, colon, esophagus, head/neck, breast, and pancreas cancers (reviewed in Hindorff et al. [2011]). Epigenetic and chromosome conformation studies in human and genetic engineering studies in mouse suggest that the variants alter distal regulatory sequences controlling the *MYC* locus, which lies telomeric to the region (Ahmadiyeh et al., 2010; Pomerantz et al., 2009; Sur et al., 2012; Tuupanen et al., 2009). A similar situation occurs in a 500 kb region in 9p21, where different variants affect multiple types of cancer (including breast cancer, melanoma, glioma, and leukemia) as well as noncancer-related diseases such as type 2 diabetes and myocardial infarction; these variants likely alter regulation of the cell-cycle genes *CDKN2A/CDKN2B*. The observation that a number of additional cancer-related loci also affect diabetes (or insulin dysregulation) suggests an important role for metabolic processes in cancer (Dupuis et al., 2010; Pal et al., 2012).

Finally, understanding the inherited factors may help to explain some disparities among ethnic groups. For example, a proportion of the higher risk of prostate cancer in African Americans and other men of African descent may be due, in part, to allele frequency differences at chromosome 8q24 (Haiman et al., 2011; Murphy et al., 2012).

Applying the Knowledge: Diagnostics and Therapeutics

The ultimate test of cancer genomics will be its ability to improve diagnostics and therapeutics. Academic centers are already beginning to adopt "first-generation" genome profiling platforms to guide cancer treatment (Dias-Santagata et al., 2010; MacConaill et al., 2009; Thomas et al., 2007; Wagle et al., 2012) These platforms involve testing a few hundred specific cancer-associated mutations or performing full sequencing of a limited set of cancer-associated genes (Lipson et al., 2012; Wagle et al., 2012). The early returns suggest that, in ~40%–60% of cases for many common solid tumors, the information points to at least one alteration that might influence therapeutic decision-making or might suggest enrollment in a particular clinical trial (Beltran et al., 2012; Hammerman et al., 2012; CGAN, 2012). As sequencing costs fall, diagnostics may move to whole-exome or whole-genome sequencing. The challenge will be to filter and annotate the results for oncologists, based on a constantly changing landscape of scientific knowledge. Eventually, genomic analysis will likely become part of the standard of care for cancer patients.

Cancer genomics will also become a key component in the design, execution, and interpretation of clinical trials. Investigators are already using genomic information for retrospective clinical analyses that correlate treatment response with specific genomic features. There is growing interest in using deep genomic characterization of "exceptional cases," such as rare tumors that show a complete clinical response to a particular anticancer regimen. For example, a recent tumor genome-sequencing study of a bladder cancer patient who experienced a complete response to a TOR inhibitor (everolimus) identified two distinct cancer gene mutations (*TSC1* and *NF2*) predicted to affect oncogenic TOR signaling (Iyer et al., 2012). Sequencing of additional everolimus-treated tumors confirmed that *TSC1* mutations correlate with clinical response.

The prospective use of genomic information may substantially transform trial design. Cancer trials have traditionally selected patients based on histologic tumor subtypes. However, it makes more sense to test targeted therapeutics on the subset of patients carrying the relevant genetic lesions; by selecting the patients most likely to benefit, one decreases sample size, cost, and unjustified harm. In some cases, it will make sense to enroll patients carrying the same genetic alteration across a wide range of tumor types (for example, a trial led by investigators at Memorial Sloan-Kettering Cancer Center in which $BRAF^{V600}$ mutant tumors from colon, thyroid, lung, and other organ sites are treated with a selective RAF or MEK inhibitor). Moreover, novel designs are becoming possible in which one simultaneously tests multiple drugs or drug combinations. In these "basket trials," patients are assigned to different therapeutic regimens based on the specific genetic profiles in their tumor (Kim et al., 2011). Basket trials may employ an "adaptive design," allowing "real-time" adjustments if hints of specific genotype-driven responses are detected (Berry, 2011). Trials can further be shaped by genomic analysis from serial biopsies to assess pharmacodynamics response and to characterize the presence of resistance mechanisms. It may be useful to create a worldwide "clearinghouse" mechanism that connects patients to trials based on their genotype, especially to obtain

a large enough sample to evaluate responses in tumors with rarer genetic features.

Cancer drug discovery efforts are already being shaped by the findings from genomic studies. In some cases, the product of the mutated gene may be an appropriate drug target. In many other cases, mutations may confer specific vulnerabilities on the cancer cell that can be discovered through functional genomic studies, such as comprehensive gene inhibition screens with RNA interference across large numbers of cancer cell lines with varying genotypes.

Beyond the development of specific drugs, knowledge of the cancer genome will be critical to design combination therapies, which will be essential for conquering cancer. Most tumors eventually develop resistance to single-agent therapeutics (reviewed in Garraway and Jänne [2012]). For example, the use of RAF and MEK inhibitors in BRAF mutant melanomas leads to spectacular responses, but tumors reappear within a year (Chapman et al., 2011b; Flaherty et al., 2012a, 2012b; Sosman et al., 2012). Multiple genetic mechanisms of resistance have been described (oncogenic *NRAS* mutations, COT/*MAP3K8* gains, *BRAF* amplification, activating *MEK1* mutations, and *NF1* loss), each of which produces sustained MAP kinase (ERK) activity in the presence of drug (Emery et al., 2009; Johannessen et al., 2010; Nazarian et al., 2010; Poulikakos et al., 2011; Wagle et al., 2011; Whittaker et al., 2013). These findings raise the possibility that adding an ERK inhibitor to existing RAF/MEK inhibitor regimens could provide an additional clinical benefit (Whittaker et al., 2013).

Systematic preclinical studies may make it possible to anticipate the mechanisms of resistance, allowing therapeutic scientists to plan for resistance long before it arises in the clinic. For example, a recent study performed large-scale screens (using RNAi knockdown and ORF overexpression) to identify genes whose loss or amplification can confer resistance to RAF inhibition in a melanoma cell line; the results were confirmed by clinical observations in patients' tumors (Johannessen et al., 2010; Whittaker et al., 2013). Another group systematically screened stromal cell lines to identify those that secrete factors that confer resistance on adjacent cancer cells; the screen revealed that hepatocyte growth factor confers resistance to RAF inhibition (Straussman et al., 2012). Such approaches may make it possible to formulate rational combination therapies even before the results of single-agent clinical trials are known.

In the end, combination therapy depends on shifting the odds of resistance. There is cause for optimism: mathematical modeling suggests that resistance may often be due to pre-existing mutations in the tumor cell population (Michor et al., 2005). If so, it should be possible to prevent recurrence by treating simultaneously with drugs directed against several independent targets so that the chance of a single cell carrying all the necessary resistance mutations is vanishingly small. This is, of course, the basis for the successful triple-drug combinations against HIV. Ultimately, cancer genomics should aim to provide a comprehensive roadmap for selecting rational, multidrug combinations for anticancer therapy.

Next Steps for Cancer Genomics

The early fruits of cancer genome studies have confirmed Renato Dulbecco's prediction about the value of complementing "piecemeal approaches" with systematic genome-wide studies. The results have already opened new frontiers in basic, translational, and clinical investigation. Still, current studies have only scratched the surface of what can be learned from comprehensive study of the cancer genome. Cancer genomics has largely focused on documenting the mutations in primary tumors. Over the coming years, the field should expand its focus to gather systematic information to inform a wider range of biological and clinical questions. Below, we suggest four important components for the next phase of cancer genomics.

Complete the Mutational Atlas of Primary Tumors

A straightforward but critical component is to finish compiling the catalog of significantly mutated genes in primary tumors of every feasible cancer type. Given the long tail of cancer genes and the variable background mutation rates, such studies will require thousands of tumor-normal pairs. Why bother to press for completeness? Scientifically, because the low-frequency drivers may in aggregate make a substantial contribution and because they are likely to harbor further surprises. Medically, because physicians will want to be able to recognize *all* driver mutations in each patient to optimize therapy. Fortunately, these efforts should become increasingly feasible and affordable given the decreasing costs of sequencing and the increasing ability to analyze small amounts of starting material from formalin-fixed, paraffin-embedded archival samples. The analysis must expand beyond the exome to include the whole genome (including long-range links to detect translocations), the transcriptome, and the epigenome (at least the methylome and key chromatin modifications). Improved laboratory and analytical methods will be needed to discern the targets of nonfocal chromosome copy number aberrations, epigenomic modifications, and nongenic translocations. In addition, the genomic information should be thoroughly mined to identify germline variants that contribute to cancer risk.

Expand the Mutational Atlas beyond Primary Tumors

The second component is to systematically expand the atlas beyond primary tumors to include the natural history of human cancer, as well as the homology to cancer in key model systems. A mutational atlas of the natural history of cancer would involve comprehensive genomic analysis of preneoplastic lesions, metastases from various organ sites, and tumors that show different types of responses to therapies, including extreme response, intrinsic resistance, and acquired resistance. Ideally, all clinical trials in oncology would be subject to such analysis. Genomic characterization should also be applied to animal models of cancer so that we can better connect these to human cancers based on mechanism. In addition to genetically engineered mouse models, intensive studies of naturally occurring cancers in large animals, especially dogs (Karlsson and Lindblad-Toh, 2008), may provide both insights and important preclinical models for drug testing.

Create a Functional Encyclopedia of Altered Pathways and Acquired Vulnerabilities

Though a mutational catalog will provide a comprehensive picture of cancer genomes, this catalog alone is not enough. We need to produce a functional encyclopedia of altered cellular pathways and acquired vulnerabilities that correspond to each

cancer genome. Genomic approaches can propel systematic functional studies, just as they have propelled comprehensive structural studies. Building a functional encyclopedia will involve (1) creating tractable models representing the full range of cancer genotypes and (2) characterizing these models with respect to their genomic alterations, essential pathways, and therapeutic vulnerabilities. Already, ongoing projects are assembling large collections of cancer cell lines; defining their genomic changes; characterizing their cellular states at the RNA, protein, and posttranslational levels; and determining their sensitivities to anticancer drugs, RNAi-based inhibition of every gene, and microenvironmental interactions (Barretina et al., 2012; Garnett et al., 2012). With a sufficiently large collection of cell models, one can correlate pathways and vulnerabilities with specific genetic lesions, providing invaluable insights into cancer biology, markers for patient selection in clinical trials, and potential new targets for cancer drug development.

One limitation has been that current cancer cell lines represent a biased sampling of cancer and cancer genotypes, owing to differences in the ability to derive cell lines. However, new methods (such as Rho kinase inhibitor-treated feeder layers and ''organoid'' culture systems) appear poised to greatly expand the repertoire of available cancer models (Huch et al., 2013; Liu et al., 2012). Patient-derived xenografts can also play a key role in preclinical studies of new therapeutics.

Enable and Promote Sharing of Cancer Genomic Information

Finally, there is one critical component that is an essential foundation for the others: widespread information sharing. Cancer genome information will grow exponentially in the years ahead as genome analysis moves from the research lab to routine clinical care for millions of patients around the world. If it were possible to share and analyze this torrent of genomic information together with associated clinical outcome data, it could significantly accelerate the understanding and treatment of cancer. The information would speed not only the identification of cancer genes, but also the correlation of therapeutic responses to specific tumor genotype, including dramatic responses to new targeted agents seen in some patients and more modest responses to different regimens. In effect, it would connect cancer care around the world into a laboratory for continuous improvement.

Making this world a reality will require coordinated efforts by researchers, hospitals, and patient groups to accomplish two goals: (1) creating the computational infrastructure to enable sharing and (2) promoting a culture of sharing. It is easy to imagine an alternative future in which cancer genomic information cannot be aggregated because it is stored in inaccessible sites and incompatible formats, much as is the case with electronic medical records in the U.S. To avoid this outcome, it will be necessary to have common or interoperable standards for data and analysis, cloud-based storage solutions to ensure data security, and rigorous systems to enforce patients' instructions concerning their data. But technology platforms alone will not suffice. Clinicians, hospitals, and healthcare networks will need to become engaged in collecting and sharing clinical outcome data. Pharmaceutical companies and others will need to share data from completed clinical trials. Ultimately, patient advocacy groups may provide the impetus for cultural change, as happened with AIDS. Though it must be up to each patient to decide whether to share his or her data, we suspect that most cancer patients will actively want to allow their information to be appropriately aggregated and shared (with appropriate rules and technology to protect privacy) to accelerate progress for this and future generations of patients. We must ensure that patients have the right and ability to contribute their information to a global fight against cancer.

Conclusions

Genomics has become a powerful tool for cancer research, yielding important biological surprises and enabling systematic classification based on cellular mechanism. Cancer genomics is just now emerging from its first phase, which has been largely focused on creating basic mutational catalogs in primary tumors. To fulfill its full promise, the field will need to deepen the structural characterization of cancer genomes, complement it with comprehensive functional characterization of cancer cells, and enable and promote information sharing across the world. Ultimately, cancer genomics is about fully knowing the enemy. While not alone a guarantee of victory, it is an essential part of any overall plan of attack.

ACKNOWLEDGMENTS

We thank Todd Golub, Matthew Meyerson, and Michael Stratton for reviewing the manuscript and for valuable comments. We thank Gad Getz, Charles Roberts, and Matthew Freedman for helpful discussions. The design of figures was aided by materials from ScienceSlides (http://www.visiscience.com). L.A.G. and E.S.L. are supported by the U.S. National Human Genome Research Institute, the U.S. National Cancer Institute, the Slim Initiative for Genomic Medicine, and the Broad Institute. L.A.G. also receives support from the Department of Defense, the Starr Cancer Consortium, the Prostate Cancer Foundation, the Melanoma Research Alliance, and the Adelson Medical Research Foundation. L.A.G. and E.S.L. are consultants and equity holders in Foundation Medicine, Inc.

REFERENCES

Agrawal, N., Frederick, M.J., Pickering, C.R., Bettegowda, C., Chang, K., Li, R.J., Fakhry, C., Xie, T.X., Zhang, J., Wang, J., et al. (2011). Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science *333*, 1154–1157.

Ahmadiyeh, N., Pomerantz, M.M., Grisanzio, C., Herman, P., Jia, L., Almendro, V., He, H.H., Brown, M., Liu, X.S., Davis, M., et al. (2010). 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. Proc. Natl. Acad. Sci. USA *107*, 9742–9746.

Altshuler, D., Daly, M.J., and Lander, E.S. (2008). Genetic mapping in human disease. Science *322*, 881–888.

Baca, S.C., Prandi, D., Lawrence, M.S., Mosquera, J.M., Romanel, A., Drier, Y., Park, K., Kitabayashi, N., MacDonald, T.Y., Ghandi, M., et al. (2013). Punctuated evolution of prostate cancer genomes. Cell *153*. Published online April 25, 2013. http://dx.doi.org/10.1016/j.cell.2013.03.021.

Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K.K., Carter, S.L., Frederick, A.M., Lawrence, M.S., Sivachenko, A.Y., Sougnez, C., Zou, L., et al. (2012). Sequence analysis of mutations and translocations across breast cancer subtypes. Nature *486*, 405–409.

Barbieri, C.E., Baca, S.C., Lawrence, M.S., Demichelis, F., Blattner, M., Theurillat, J.P., White, T.A., Stojanov, P., Van Allen, E., Stransky, N., et al. (2012). Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat. Genet. *44*, 685–689.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

Bass, A.J., Watanabe, H., Mermel, C.H., Yu, S., Perner, S., Verhaak, R.G., Kim, S.Y., Wardwell, L., Tamayo, P., Gat-Viks, I., et al. (2009). SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat. Genet. *41*, 1238–1242.

Beltran, H., Yelensky, R., Frampton, G.M., Park, K., Downing, S.R., Macdonald, T.Y., Jarosz, M., Lipson, D., Tagawa, S.T., Nanus, D.M., et al. (2012). Targeted next-generation sequencing of advanced prostate cancer identifies potential therapeutic targets and disease heterogeneity. Eur. Urol. Published online September 5, 2012. http://dx.doi.org/10.1016/j.eururo.2012.08.053.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. Nature *456*, 53–59.

Berger, M.F., Lawrence, M.S., Demichelis, F., Drier, Y., Cibulskis, K., Sivachenko, A.Y., Sboner, A., Esgueva, R., Pflueger, D., Sougnez, C., et al. (2011). The genomic complexity of primary human prostate cancer. Nature *470*, 214–220.

Berger, M.F., Hodis, E., Heffernan, T.P., Deribe, Y.L., Lawrence, M.S., Protopopov, A., Ivanova, E., Watson, I.R., Nickerson, E., Ghosh, P., et al. (2012). Melanoma genome sequencing reveals frequent PREX2 mutations. Nature *485*, 502–506.

Bernt, K.M., Zhu, N., Sinha, A.U., Vempati, S., Faber, J., Krivtsov, A.V., Feng, Z., Punt, N., Daigle, A., Bullinger, L., et al. (2011). MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. Cancer Cell *20*, 66–78.

Beroukhim, R., Getz, G., Nghiemphu, L., Barretina, J., Hsueh, T., Linhart, D., Vivanco, I., Lee, J.C., Huang, J.H., Alexander, S., et al. (2007). Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proc. Natl. Acad. Sci. USA *104*, 20007–20012.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. Nature *463*, 899–905.

Berry, D.A. (2011). Adaptive clinical trials in oncology. Nat. Rev. Clin. Oncol. *9*, 199–207.

Biankin, A.V., Waddell, N., Kassahn, K.S., Gingras, M.C., Muthuswamy, L.B., Johns, A.L., Miller, D.K., Wilson, P.J., Patch, A.M., Wu, J., et al.; Australian Pancreatic Cancer Genome Initiative. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature *491*, 399–405.

Boveri, T. (2008). Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. J. Cell Sci. *121*(*Suppl 1*), 1–84.

CGAN (Cancer Genome Atlas Network). (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

CGARN (Cancer Genome Atlas Research Network). (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061–1068.

Carter, S.L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P.W., Onofrio, R.C., Winckler, W., Weir, B.A., et al. (2012). Absolute quantification of somatic DNA alterations in human cancer. Nat. Biotechnol. *30*, 413–421.

Cazier, J.B., and Tomlinson, I. (2010). General lessons from large-scale studies to identify human cancer predisposition genes. J. Pathol. *220*, 255–262.

Chapman, M.A., Lawrence, M.S., Keats, J.J., Cibulskis, K., Sougnez, C., Schinzel, A.C., Harview, C.L., Brunet, J.P., Ahmann, G.J., Adli, M., et al. (2011a). Initial genome sequencing and analysis of multiple myeloma. Nature *471*, 467–472.

Chapman, P.B., Hauschild, A., Robert, C., Haanen, J.B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., et al.; BRIM-3 Study Group. (2011b). Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N. Engl. J. Med. *364*, 2507–2516.

Chen, K., Wallis, J.W., McLellan, M.D., Larson, D.E., Kalicki, J.M., Pohl, C.S., McGrath, S.D., Wendl, M.C., Zhang, Q., Locke, D.P., et al. (2009). BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat. Methods *6*, 677–681.

Chin, K., DeVries, S., Fridlyand, J., Spellman, P.T., Roydasgupta, R., Kuo, W.L., Lapuk, A., Neve, R.M., Qian, Z., Ryder, T., et al. (2006). Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell *10*, 529–541.

Chung, C.C., and Chanock, S.J. (2011). Current status of genome-wide association studies in cancer. Hum. Genet. *130*, 59–78.

Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. Published online February 20, 2013. http://dx.doi.org/10.1038/nbt.2514.

Crasta, K., Ganem, N.J., Dagher, R., Lantermann, A.B., Ivanova, E.V., Pan, Y., Nezi, L., Protopopov, A., Chowdhury, D., and Pellman, D. (2012). DNA breaks and chromosome pulverization from errors in mitosis. Nature *482*, 53–58.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature *462*, 739–744.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., et al. (2002). Mutations of the BRAF gene in human cancer. Nature *417*, 949–954.

Dawson, M.A., and Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. Cell *150*, 12–27.

Dees, N.D., Zhang, Q., Kandoth, C., Wendl, M.C., Schierding, W., Koboldt, D.C., Mooney, T.B., Callaway, M.B., Dooling, D., Mardis, E.R., et al. (2012). MuSiC: identifying mutational significance in cancer genomes. Genome Res. *22*, 1589–1598.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.P., Robert, F., Alberdi, A., et al. (2009). Mutation in TET2 in myeloid cancers. N. Engl. J. Med. *360*, 2289–2301.

Demetri, G.D., von Mehren, M., Blanke, C.D., Van den Abbeele, A.D., Eisenberg, B., Roberts, P.J., Heinrich, M.C., Tuveson, D.A., Singer, S., Janicek, M., et al. (2002). Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. N. Engl. J. Med. *347*, 472–480.

Dias-Santagata, D., Akhavanfard, S., David, S.S., Vernovsky, K., Kuhlmann, G., Boisvert, S.L., Stubbs, H., McDermott, U., Settleman, J., Kwak, E.L., et al. (2010). Rapid targeted mutational analysis of human tumours: a clinical platform to guide personalized cancer medicine. EMBO Mol. Med. *2*, 146–158.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., et al. (2008). Somatic mutations affect key pathways in lung adenocarcinoma. Nature *455*, 1069–1075.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., et al. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature *481*, 506–510.

Dolnik, A., Engelmann, J.C., Scharfenberger-Schmeer, M., Mauch, J., Kelkenberg-Schade, S., Haldemann, B., Fries, T., Krönke, J., Kühn, M.W., Paschka, P., et al. (2012). Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. Blood *120*, e83–e92.

Druker, B.J., Talpaz, M., Resta, D.J., Peng, B., Buchdunger, E., Ford, J.M., Lydon, N.B., Kantarjian, H., Capdeville, R., Ohno-Jones, S., and Sawyers, C.L. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N. Engl. J. Med. *344*, 1031–1037.

Dulbecco, R. (1986). A turning point in cancer research: sequencing the human genome. Science *231*, 1055–1056.

Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U., Wheeler, E., Glazer, N.L., Bouatia-Naji, N., Gloyn, A.L., et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten

on behalf of Procardis Consortium; MAGIC investigators. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat. Genet. *42*, 105–116.

Dutt, A., Salvesen, H.B., Chen, T.H., Ramos, A.H., Onofrio, R.C., Hatton, C., Nicoletti, R., Winckler, W., Grewal, R., Hanna, M., et al. (2008). Drug-sensitive FGFR2 mutations in endometrial carcinoma. Proc. Natl. Acad. Sci. USA *105*, 8713–8717.

Ellis, M.J., Ding, L., Shen, D., Luo, J., Suman, V.J., Wallis, J.W., Van Tine, B.A., Hoog, J., Goiffon, R.J., Goldstein, T.C., et al. (2012). Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature *486*, 353–360.

Emery, C.M., Vijayendran, K.G., Zipser, M.C., Sawyer, A.M., Niu, L., Kim, J.J., Hatton, C., Chopra, R., Oberholzer, P.A., Karpova, M.B., et al. (2009). MEK1 mutations confer resistance to MEK and B-RAF inhibition. Proc. Natl. Acad. Sci. USA *106*, 20411–20416.

Estey, E. (2007). Acute myeloid leukemia and myelodysplastic syndromes in older patients. J. Clin. Oncol. *25*, 1908–1915.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., et al. (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell *18*, 553–567.

Flaherty, K.T., Infante, J.R., Daud, A., Gonzalez, R., Kefford, R.F., Sosman, J., Hamid, O., Schuchter, L., Cebon, J., Ibrahim, N., et al. (2012a). Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. N. Engl. J. Med. *367*, 1694–1703.

Flaherty, K.T., Robert, C., Hersey, P., Nathan, P., Garbe, C., Milhem, M., Demidov, L.V., Hassel, J.C., Rutkowski, P., Mohr, P., et al.; METRIC Study Group. (2012b). Improved survival with MEK inhibition in BRAF-mutated melanoma. N. Engl. J. Med. *367*, 107–114.

Folco, E.G., Coil, K.E., and Reed, R. (2011). The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. Genes Dev. *25*, 440–444.

Fujimoto, A., Totoki, Y., Abe, T., Boroevich, K.A., Hosoda, F., Nguyen, H.H., Aoki, M., Hosono, N., Kubo, M., Miya, F., et al. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nat. Genet. *44*, 760–764.

Gabor Miklos, G.L. (2005). The human cancer genome project—one more misstep in the war on cancer. Nat. Biotechnol. *23*, 535–537.

Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A., Lau, K.W., Greninger, P., Thompson, I.R., Luo, X., Soares, J., et al. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature *483*, 570–575.

Garraway, L.A., and Jänne, P.A. (2012). Circumventing cancer drug resistance in the era of personalized medicine. Cancer Discov. *2*, 214–226.

Garraway, L.A., Widlund, H.R., Rubin, M.A., Getz, G., Berger, A.J., Ramaswamy, S., Beroukhim, R., Milner, D.A., Granter, S.R., Du, J., et al. (2005). Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature *436*, 117–122.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N. Engl. J. Med. *366*, 883–892.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., et al. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat. Biotechnol. *27*, 182–189.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Grasso, C.S., Wu, Y.M., Robinson, D.R., Cao, X., Dhanasekaran, S.M., Khan, A.P., Quist, M.J., Jing, X., Lonigro, R.J., Brenner, J.C., et al. (2012). The mutational landscape of lethal castration-resistant prostate cancer. Nature *487*, 239–243.

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Stevens, C., et al. (2007). Patterns of somatic mutation in human cancer genomes. Nature *446*, 153–158.

Gui, Y., Guo, G., Huang, Y., Hu, X., Tang, A., Gao, S., Wu, R., Chen, C., Li, X., Zhou, L., et al. (2011). Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. Nat. Genet. *43*, 875–878.

Haiman, C.A., Chen, G.K., Blot, W.J., Strom, S.S., Berndt, S.I., Kittles, R.A., Rybicki, B.A., Isaacs, W.B., Ingles, S.A., Stanford, J.L., et al. (2011). Characterizing genetic risk at known prostate cancer susceptibility loci in African Americans. PLoS Genet. *7*, e1001387.

Hájková, H., Marková, J., Haškovec, C., Sárová, I., Fuchs, O., Kostečka, A., Cetkovský, P., Michalová, K., and Schwarz, J. (2012). Decreased DNA methylation in acute myeloid leukemia patients with DNMT3A mutations and prognostic implications of DNA methylation. Leuk. Res. *36*, 1128–1133.

Hammerman, P.S., Hayes, D.N., Wilkerson, M.D., Schultz, N., Bose, R., Chu, A., Collisson, E.A., Cope, L., Creighton, C.J., Getz, G., et al.; Cancer Genome Atlas Research Network. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*, 519–525.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hindorff, L.A., Gillanders, E.M., and Manolio, T.A. (2011). Genetic architecture of cancer and other complex diseases: lessons learned and future directions. Carcinogenesis *32*, 945–954.

Ho, L., and Crabtree, G.R. (2010). Chromatin remodelling during development. Nature *463*, 474–484.

Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M.N., Smith, S.W., Middle, C.M., Rodesch, M.J., Albert, T.J., Hannon, G.J., and McCombie, W.R. (2007). Genome-wide in situ exon capture for selective resequencing. Nat. Genet. *39*, 1522–1527.

Hodis, E., Watson, I.R., Kryukov, G.V., Arold, S.T., Imielinski, M., Theurillat, J.P., Nickerson, E., Auclair, D., Li, L., Place, C., et al. (2012). A landscape of driver mutations in melanoma. Cell *150*, 251–263.

Holland, A.J., and Cleveland, D.W. (2012). Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. Nat. Med. *18*, 1630–1638.

Hou, Y., Song, L., Zhu, P., Zhang, B., Tao, Y., Xu, X., Li, F., Wu, K., Liang, J., Shao, D., et al. (2012). Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell *148*, 873–885.

Huang, J., Deng, Q., Wang, Q., Li, K.Y., Dai, J.H., Li, N., Zhu, Z.D., Zhou, B., Liu, X.Y., Liu, R.F., et al. (2012a). Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. Nat. Genet. *44*, 1117–1121.

Huang, S., Gulzar, Z.G., Salari, K., Lapointe, J., Brooks, J.D., and Pollack, J.R. (2012b). Recurrent deletion of CHD1 in prostate cancer with relevance to cell invasiveness. Oncogene *31*, 4164–4170.

Huang, F.W., Hodis, E., Xu, M.J., Kryukov, G.V., Chin, L., and Garraway, L.A. (2013). Highly recurrent TERT promoter mutations in human melanoma. Science *339*, 957–959.

Huch, M., Dorrell, C., Boj, S.F., van Es, J.H., Li, V.S., van de Wetering, M., Sato, T., Hamer, K., Sasaki, N., Finegold, M.J., et al. (2013). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature *494*, 247–250.

Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bell, C., Bernabé, R.R., Bhan, M.K., Calvo, F., Eerola, I., Gerhard, D.S., et al.; International Cancer Genome Consortium. (2010). International network of cancer genome projects. Nature *464*, 993–998.

Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., et al. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell *150*, 1107–1120.

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. Nature *431*, 931–945.

Iyer, G., Hanrahan, A.J., Milowsky, M.I., Al-Ahmadie, H., Scott, S.N., Janakiraman, M., Pirun, M., Sander, C., Socci, N.D., Ostrovnaya, I., et al. (2012). Genome sequencing identifies a basis for everolimus sensitivity. Science *338*, 221.

Jaju, R.J., Fidler, C., Haas, O.A., Strickson, A.J., Watkins, F., Clark, K., Cross, N.C., Cheng, J.F., Aplan, P.D., Kearney, L., et al. (2001). A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. Blood *98*, 1264–1267.

Jankowska, A.M., Makishima, H., Tiu, R.V., Szpurka, H., Huang, Y., Traina, F., Visconte, V., Sugimoto, Y., Prince, C., O'Keefe, C., et al. (2011). Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. Blood *118*, 3932–3941.

Johannessen, C.M., Boehm, J.S., Kim, S.Y., Thomas, S.R., Wardwell, L., Johnson, L.A., Emery, C.M., Stransky, N., Cogdill, A.P., Barretina, J., et al. (2010). COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. Nature *468*, 968–972.

Jones, S., Wang, T.L., Shih, IeM., Mao, T.L., Nakayama, K., Roden, R., Glas, R., Slamon, D., Diaz, L.A., Jr., Vogelstein, B., et al. (2010). Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. Science *330*, 228–231.

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat. Chem. Biol. *3*, 576–583.

Kan, Z., Jaiswal, B.S., Stinson, J., Janakiraman, V., Bhatt, D., Stern, H.M., Yue, P., Haverty, P.M., Bourgon, R., Zheng, J., et al. (2010). Diverse somatic mutation patterns and pathway alterations in human cancers. Nature *466*, 869–873.

Karlsson, E.K., and Lindblad-Toh, K. (2008). Leader of the pack: gene mapping in dogs and other model organisms. Nat. Rev. Genet. *9*, 713–725.

Kim, D., and Salzberg, S.L. (2011). TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. *12*, R72.

Kim, Y.H., Kwei, K.A., Girard, L., Salari, K., Kao, J., Pacyna-Gengelbach, M., Wang, P., Hernandez-Boussard, T., Gazdar, A.F., Petersen, I., et al. (2010). Genomic and functional analysis identifies CRKL as an oncogene amplified in lung cancer. Oncogene *29*, 1421–1430.

Kim, E.S., Herbst, R.S., Wistuba, I.I., Lee, J.J., Blumenschein, G.R., Jr., Tsao, A., Stewart, D.J., Hicks, M.E., Erasmus, J., Jr., Gupta, S., et al. (2011). The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov. *1*, 44–53.

Klinakis, A., Lobry, C., Abdel-Wahab, O., Oh, P., Haeno, H., Buonamici, S., van De Walle, I., Cathelin, S., Trimarchi, T., Araldi, E., et al. (2011). A novel tumour-suppressor function for the Notch pathway in myeloid leukaemia. Nature *473*, 230–233.

Koivunen, P., Lee, S., Duncan, C.G., Lopez, G., Lu, G., Ramkissoon, S., Losman, J.A., Joensuu, P., Bergmann, U., Gross, S., et al. (2012). Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. Nature *483*, 484–488.

Kostic, A.D., Ojesina, A.I., Pedamallu, C.S., Jung, J., Verhaak, R.G., Getz, G., and Meyerson, M. (2011). PathSeq: software to identify or discover microbes by deep sequencing of human tissue. Nat. Biotechnol. *29*, 393–396.

Kralovics, R., Passamonti, F., Buser, A.S., Teo, S.S., Tiedt, R., Passweg, J.R., Tichelli, A., Cazzola, M., and Skoda, R.C. (2005). A gain-of-function mutation of JAK2 in myeloproliferative disorders. N. Engl. J. Med. *352*, 1779–1790.

Krauthammer, M., Kong, Y., Ha, B.H., Evans, P., Bacchiocchi, A., McCusker, J.P., Cheng, E., Davis, M.J., Goh, G., Choi, M., et al. (2012). Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. Nat. Genet. *44*, 1006–1014.

Kreso, A., O'Brien, C.A., van Galen, P., Gan, O.I., Notta, F., Brown, A.M., Ng, K., Ma, J., Wienholds, E., Dunant, C., et al. (2013). Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. Science *339*, 543–548.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929–930.

Landau, D.A., Carter, S.L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M.S., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., et al. (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell *152*, 714–726.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al.; International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Langemeijer, S.M., Kuiper, R.P., Berends, M., Knops, R., Aslanyan, M.G., Massop, M., Stevens-Linders, E., van Hoogen, P., van Kessel, A.G., Raymakers, R.A., et al. (2009). Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat. Genet. *41*, 838–842.

Le Gallo, M., O'Hara, A.J., Rudd, M.L., Urick, M.E., Hansen, N.F., O'Neil, N.J., Price, J.C., Zhang, S., England, B.M., Godwin, A.K., et al.; NIH Intramural Sequencing Center (NISC) Comparative Sequencing Program. (2012). Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. Nat. Genet. *44*, 1310–1315.

Levine, R.L., Wadleigh, M., Cools, J., Ebert, B.L., Wernig, G., Huntly, B.J., Boggon, T.J., Wlodarska, I., Clark, J.J., Moore, S., et al. (2005). Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell *7*, 387–397.

Ley, T.J., Mardis, E.R., Ding, L., Fulton, B., McLellan, M.D., Chen, K., Dooling, D., Dunford-Shore, B.H., McGrath, S., Hickenbotham, M., et al. (2008). DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature *456*, 66–72.

Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E., Kandoth, C., Payton, J.E., Baty, J., Welch, J., et al. (2010). DNMT3A mutations in acute myeloid leukemia. N. Engl. J. Med. *363*, 2424–2433.

Li, M., Zhao, H., Zhang, X., Wood, L.D., Anders, R.A., Choti, M.A., Pawlik, T.M., Daniel, H.D., Kannangai, R., Offerhaus, G.J., et al. (2011). Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. Nat. Genet. *43*, 828–829.

Lindsley, R.C., and Ebert, B.L. (2013). Molecular Pathophysiology of Myelodysplastic Syndromes. Annu. Rev. Pathol. *8*, 21–47. Published online August 28, 2012. http://dx.doi.org/10.1146/annurev-pathol-011811-132436.

Lipson, D., Capelletti, M., Yelensky, R., Otto, G., Parker, A., Jarosz, M., Curran, J.A., Balasubramanian, S., Bloom, T., Brennan, K.W., et al. (2012). Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat. Med. *18*, 382–384.

Liu, P., Erez, A., Nagamani, S.C., Dhar, S.U., Kołodziejska, K.E., Dharmadhikari, A.V., Cooper, M.L., Wiszniewska, J., Zhang, F., Withers, M.A., et al. (2011). Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. Cell *146*, 889–903.

Liu, X., Ory, V., Chapman, S., Yuan, H., Albanese, C., Kallakury, B., Timofeeva, O.A., Nealon, C., Dakic, A., Simic, V., et al. (2012). ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells. Am. J. Pathol. *180*, 599–607.

Lohr, J.G., Stojanov, P., Lawrence, M.S., Auclair, D., Chapuy, B., Sougnez, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y.W., Slager, S.L., et al. (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc. Natl. Acad. Sci. USA *109*, 3879–3884.

Lu, C., Ward, P.S., Kapoor, G.S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A., et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature *483*, 474–478.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating mutations in the epidermal growth factor receptor

underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med. *350*, 2129–2139.

Macconaill, L.E., and Garraway, L.A. (2010). Clinical implications of the cancer genome. J. Clin. Oncol. *28*, 5219–5228.

MacConaill, L.E., Campbell, C.D., Kehoe, S.M., Bass, A.J., Hatton, C., Niu, L., Davis, M., Yao, K., Hanna, M., Mondal, C., et al. (2009). Profiling critical cancer gene mutations in clinical tumor samples. PLoS ONE *4*, e7887.

Makishima, H., Jankowska, A.M., Tiu, R.V., Szpurka, H., Sugimoto, Y., Hu, Z., Saunthararajah, Y., Guinta, K., Keddache, M.A., Putnam, P., et al. (2010). Novel homo- and hemizygous mutations in EZH2 in myeloid malignancies. Leukemia: official journal of the Leukemia Society of America. Leukemia *24*, 1799–1804.

Marcucci, G., Metzeler, K.H., Schwind, S., Becker, H., Maharry, K., Mrózek, K., Radmacher, M.D., Kohlschmidt, J., Nicolet, D., Whitman, S.P., et al. (2012). Age-related prognostic impact of different types of DNMT3A mutations in adults with primary cytogenetically normal acute myeloid leukemia. J. Clin. Oncol. *30*, 742–750.

Mardis, E.R., Ding, L., Dooling, D.J., Larson, D.E., McLellan, M.D., Chen, K., Koboldt, D.C., Fulton, R.S., Delehaunty, K.D., McGrath, S.D., et al. (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. N. Engl. J. Med. *361*, 1058–1066.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature *437*, 376–380.

Meyerson, M., Gabriel, S., and Getz, G. (2010). Advances in understanding cancer genomes through second-generation sequencing. Nat. Rev. Genet. *11*, 685–696.

Michor, F., Hughes, T.P., Iwasa, Y., Branford, S., Shah, N.P., Sawyers, C.L., and Nowak, M.A. (2005). Dynamics of chronic myeloid leukaemia. Nature *435*, 1267–1270.

Molenaar, J.J., Koster, J., Zwijnenburg, D.A., van Sluis, P., Valentijn, L.J., van der Ploeg, I., Hamdi, M., van Nes, J., Westerman, B.A., van Arkel, J., et al. (2012). Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. Nature *483*, 589–593.

Morin, R.D., Johnson, N.A., Severson, T.M., Mungall, A.J., An, J., Goya, R., Paul, J.E., Boyle, M., Woolcock, B.W., Kuchenbauer, F., et al. (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat. Genet. *42*, 181–185.

Morin, R.D., Mendez-Lago, M., Mungall, A.J., Goya, R., Mungall, K.L., Corbett, R.D., Johnson, N.A., Severson, T.M., Chiu, R., Field, M., et al. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature *476*, 298–303.

Murphy, A.B., Ukoli, F., Freeman, V., Bennett, F., Aiken, W., Tulloch, T., Coard, K., Angwafo, F., and Kittles, R.A. (2012). 8q24 risk alleles in West African and Caribbean men. Prostate *72*, 1366–1373.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94.

Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R.C., Lee, H., Chen, Z., Lee, M.K., Attar, N., Sazegar, H., et al. (2010). Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature *468*, 973–977.

Ngo, V.N., Young, R.M., Schmitz, R., Jhavar, S., Xiao, W., Lim, K.H., Kohlhammer, H., Xu, W., Yang, Y., Zhao, H., et al. (2011). Oncogenically active MYD88 mutations in human lymphoma. Nature *470*, 115–119.

Nik-Zainal, S., Alexandrov, L.B., Wedge, D.C., Van Loo, P., Greenman, C.D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, L.A., et al.; Breast Cancer Working Group of the International Cancer Genome Consortium. (2012a). Mutational processes molding the genomes of 21 breast cancers. Cell *149*, 979–993.

Nik-Zainal, S., Van Loo, P., Wedge, D.C., Alexandrov, L.B., Greenman, C.D., Lau, K.W., Raine, K., Jones, D., Marshall, J., Ramakrishna, M., et al.; Breast Cancer Working Group of the International Cancer Genome Consortium. (2012b). The life history of 21 breast cancers. Cell *149*, 994–1007.

Noushmehr, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P., Pan, F., Pelloski, C.E., Sulman, E.P., Bhat, K.P., et al.; Cancer Genome Atlas Research Network. (2010). Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell *17*, 510–522.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. Science *194*, 23–28.

Ogino, S., Cantor, M., Kawasaki, T., Brahmandam, M., Kirkner, G.J., Weisenberger, D.J., Campan, M., Laird, P.W., Loda, M., and Fuchs, C.S. (2006). CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. Gut *55*, 1000–1006.

Paez, J.G., Jänne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497–1500.

Pal, A., Barber, T.M., Van de Bunt, M., Rudge, S.A., Zhang, Q., Lachlan, K.L., Cooper, N.S., Linden, H., Levy, J.C., Wakelam, M.J., et al. (2012). PTEN mutations as a cause of constitutive insulin sensitivity and obesity. N. Engl. J. Med. *367*, 1002–1011.

Palanisamy, N., Ateeq, B., Kalyana-Sundaram, S., Pflueger, D., Ramnarayanan, K., Shankar, S., Han, B., Cao, Q., Cao, X., Suleman, K., et al. (2010). Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. Nat. Med. *16*, 793–798.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., et al. (2004). EGF receptor gene mutations are common in lung cancers from ''never smokers'' and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc. Natl. Acad. Sci. USA *101*, 13306–13311.

Papaemmanuil, E., Cazzola, M., Boultwood, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J.S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., et al.; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. (2011). Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N. Engl. J. Med. *365*, 1384–1395.

Parsons, D.W., Jones, S., Zhang, X., Lin, J.C., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.M., Gallia, G.L., et al. (2008). An integrated genomic analysis of human glioblastoma multiforme. Science *321*, 1807–1812.

Pasqualucci, L., Dominguez-Sola, D., Chiarenza, A., Fabbri, G., Grunn, A., Trifonov, V., Kasper, L.H., Lerach, S., Tang, H., Ma, J., et al. (2011a). Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature *471*, 189–195.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., et al. (2011b). Analysis of the coding genome of diffuse large B-cell lymphoma. Nat. Genet. *43*, 830–837.

Peifer, M., Fernández-Cuesta, L., Sos, M.L., George, J., Seidel, D., Kasper, L.H., Plenker, D., Leenders, F., Sun, R., Zander, T., et al. (2012). Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat. Genet. *44*, 1104–1110.

Pleasance, E.D., Cheetham, R.K., Stephens, P.J., McBride, D.J., Humphray, S.J., Greenman, C.D., Varela, I., Lin, M.L., Ordóñez, G.R., Bignell, G.R., et al. (2010a). A comprehensive catalogue of somatic mutations from a human cancer genome. Nature *463*, 191–196. Published online December 16, 2009. http://dx.doi.org/10.1038/nature08658.

Pleasance, E.D., Stephens, P.J., O'Meara, S., McBride, D.J., Meynert, A., Jones, D., Lin, M.L., Beare, D., Lau, K.W., Greenman, C., et al. (2010b). A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature *463*, 184–190. Published online December 16, 2009. http://dx.doi.org/10.1038/nature08629.

Pollock, P.M., Gartside, M.G., Dejeza, L.C., Powell, M.A., Mallon, M.A., Davies, H., Mohammadi, M., Futreal, P.A., Stratton, M.R., Trent, J.M., and

Goodfellow, P.J. (2007). Frequent activating FGFR2 mutations in endometrial carcinomas parallel germline mutations associated with craniosynostosis and skeletal dysplasia syndromes. Oncogene *26*, 7158–7162.

Pomerantz, M.M., and Freedman, M.L. (2011). The genetics of cancer risk. Cancer J. *17*, 416–422.

Pomerantz, M.M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M.P., Doddapaneni, H., Beckwith, C.A., Chan, J.A., Hills, A., Davis, M., et al. (2009). The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat. Genet. *41*, 882–884.

Poulikakos, P.I., Persaud, Y., Janakiraman, M., Kong, X., Ng, C., Moriceau, G., Shi, H., Atefi, M., Titz, B., Gabay, M.T., et al. (2011). RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). Nature *480*, 387–390.

Puente, X.S., Pinyol, M., Quesada, V., Conde, L., Ordóñez, G.R., Villamor, N., Escaramis, G., Jares, P., Beà, S., González-Díaz, M., et al. (2011). Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature *475*, 101–105.

Quesada, V., Conde, L., Villamor, N., Ordóñez, G.R., Jares, P., Bassaganyas, L., Ramsay, A.J., Beà, S., Pinyol, M., Martínez-Trillos, A., et al. (2012). Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat. Genet. *44*, 47–52.

Rausch, T., Jones, D.T., Zapatka, M., Stütz, A.M., Zichner, T., Weischenfeldt, J., Jäger, N., Remke, M., Shih, D., Northcott, P.A., et al. (2012). Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell *148*, 59–71.

Robinson, D.R., Kalyana-Sundaram, S., Wu, Y.M., Shankar, S., Cao, X., Ateeq, B., Asangani, I.A., Iyer, M., Maher, C.A., Grasso, C.S., et al. (2011). Functionally recurrent rearrangements of the MAST kinase and Notch gene families in breast cancer. Nat. Med. *17*, 1646–1651.

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X., et al. (2012). Novel mutations target distinct subgroups of medulloblastoma. Nature *488*, 43–48.

Rosati, R., La Starza, R., Veronese, A., Aventin, A., Schwienbacher, C., Vallespi, T., Negrini, M., Martelli, M.F., and Mecucci, C. (2002). NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). Blood *99*, 3857–3860.

Salari, K., Spulak, M.E., Cuff, J., Forster, A.D., Giacomini, C.P., Huang, S., Ko, M.E., Lin, A.Y., van de Rijn, M., and Pollack, J.R. (2012). CDX2 is an amplified lineage-survival oncogene in colorectal cancer. Proc. Natl. Acad. Sci. USA *109*, E3196–E3205.

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S.M., Riggins, G.J., et al. (2004). High frequency of mutations of the PIK3CA gene in human cancers. Science *304*, 554.

Schwartzentruber, J., Korshunov, A., Liu, X.Y., Jones, D.T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.A., Tönjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231.

Seshagiri, S., Stawiski, E.W., Durinck, S., Modrusan, Z., Storm, E.E., Conboy, C.B., Chaudhuri, S., Guan, Y., Janakiraman, V., Jaiswal, B.S., et al. (2012). Recurrent R-spondin fusions in colon cancer. Nature *488*, 660–664.

Shibata, T., Ohta, T., Tong, K.I., Kokubu, A., Odogawa, R., Tsuta, K., Asamura, H., Yamamoto, M., and Hirohashi, S. (2008). Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. Proc. Natl. Acad. Sci. USA *105*, 13568–13573.

Sjöblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., et al. (2006). The consensus coding sequences of human breast and colorectal cancers. Science *314*, 268–274.

Sosman, J.A., Kim, K.B., Schuchter, L., Gonzalez, R., Pavlick, A.C., Weber, J.S., McArthur, G.A., Hutson, T.E., Moschos, S.J., Flaherty, K.T., et al. (2012). Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. N. Engl. J. Med. *366*, 707–714.

Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976). DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature *260*, 170–173.

Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., et al. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell *144*, 27–40.

Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R., et al.; Oslo Breast Cancer Consortium (OSBREAC). (2012). The landscape of cancer genes and mutational processes in breast cancer. Nature *486*, 400–404.

Stransky, N., Egloff, A.M., Tward, A.D., Kostic, A.D., Cibulskis, K., Sivachenko, A., Kryukov, G.V., Lawrence, M.S., Sougnez, C., McKenna, A., et al. (2011). The mutational landscape of head and neck squamous cell carcinoma. Science *333*, 1157–1160.

Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature *458*, 719–724.

Straussman, R., Morikawa, T., Shee, K., Barzily-Rokni, M., Qian, Z.R., Du, J., Davis, A., Mongare, M.M., Gould, J., Frederick, D.T., et al. (2012). Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. Nature *487*, 500–504.

Sur, I.K., Hallikas, O., Vähärautio, A., Yan, J., Turunen, M., Enge, M., Taipale, M., Karhu, A., Aaltonen, L.A., and Taipale, J. (2012). Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. Science *338*, 1360–1363.

Tabin, C.J., Bradley, S.M., Bargmann, C.I., Weinberg, R.A., Papageorge, A.G., Scolnick, E.M., Dhar, R., Lowy, D.R., and Chang, E.H. (1982). Mechanism of activation of a human oncogene. Nature *300*, 143–149.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Taplin, M.E., Bubley, G.J., Shuster, T.D., Frantz, M.E., Spooner, A.E., Ogata, G.K., Keer, H.N., and Balk, S.P. (1995). Mutation of the androgen-receptor gene in metastatic androgen-independent prostate cancer. N. Engl. J. Med. *332*, 1393–1398.

Thomas, R.K., Baker, A.C., Debiasi, R.M., Winckler, W., Laframboise, T., Lin, W.M., Wang, M., Feng, W., Zander, T., MacConaill, L., et al. (2007). High-throughput oncogene mutation profiling in human cancer. Nat. Genet. *39*, 347–351.

Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J.G., Baylin, S.B., and Issa, J.P. (1999). CpG island methylator phenotype in colorectal cancer. Proc. Natl. Acad. Sci. USA *96*, 8681–8686.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W., Lu, C., Ward, P.S., et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature *483*, 479–483.

Tuupanen, S., Turunen, M., Lehtonen, R., Hallikas, O., Vanharanta, S., Kivioja, T., Björklund, M., Wei, G., Yan, J., Niittymäki, I., et al. (2009). The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat. Genet. *41*, 885–890.

Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P., Davies, H., Jones, D., Lin, M.L., Teague, J., et al. (2011). Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature *469*, 539–542.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

Wagle, N., Emery, C., Berger, M.F., Davis, M.J., Sawyer, A., Pochanard, P., Kehoe, S.M., Johannessen, C.M., Macconaill, L.E., Hahn, W.C., et al. (2011). Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. J. Clin. Oncol. *29*, 3085–3096.

Wagle, N., Berger, M.F., Davis, M.J., Blumenstiel, B., Defelice, M., Pochanard, P., Ducar, M., Van Hummelen, P., Macconaill, L.E., Hahn, W.C., et al. (2012). High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer Discov. *2*, 82–93.

Walter, M.J., Ding, L., Shen, D., Shao, J., Grillot, M., McLellan, M., Fulton, R., Schmidt, H., Kalicki-Veizer, J., O'Laughlin, M., et al. (2011). Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia: official journal of the Leukemia Society of America. Leukemia *25*, 1153–1158.

Wang, L., Lawrence, M.S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D.S., Zhang, L., et al. (2011a). SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N. Engl. J. Med. *365*, 2497–2506.

Wang, N.J., Sanborn, Z., Arnett, K.L., Bayston, L.J., Liao, W., Proby, C.M., Leigh, I.M., Collisson, E.A., Gordon, P.B., Jakkula, L., et al. (2011b). Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. Proc. Natl. Acad. Sci. USA *108*, 17761–17766.

Wang, X.S., Shankar, S., Dhanasekaran, S.M., Ateeq, B., Sasaki, A.T., Jing, X., Robinson, D., Cao, Q., Prensner, J.R., Yocum, A.K., et al. (2011c). Characterization of KRAS rearrangements in metastatic prostate cancer. Cancer Discov. *1*, 35–43.

Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Coller, H.A., Cross, J.R., Fantin, V.R., Hedvat, C.V., Perl, A.E., et al. (2010). The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell *17*, 225–234.

Weinberg, R. (2010). Point: Hypotheses first. Nature *464*, 678.

Weir, B.A., Woo, M.S., Getz, G., Perner, S., Ding, L., Beroukhim, R., Lin, W.M., Province, M.A., Kraja, A., Johnson, L.A., et al. (2007). Characterizing the cancer genome in lung adenocarcinoma. Nature *450*, 893–898.

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., et al. (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat. Genet. *38*, 787–793.

Weiss, J., Sos, M.L., Seidel, D., Peifer, M., Zander, T., Heuckmann, J.M., Ullrich, R.T., Menon, R., Maier, S., Soltermann, A., et al. (2010). Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. Sci. Transl. Med. *2*, 62ra93.

Weng, A.P., Ferrando, A.A., Lee, W., Morris, J.P., 4th, Silverman, L.B., Sanchez-Irizarry, C., Blacklow, S.C., Look, A.T., and Aster, J.C. (2004). Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science *306*, 269–271.

Whittaker, S.R., Theurillat, J.P., Van Allen, E., Wagle, N., Hsiao, J., Cowley, G.S., Schadendorf, D., Root, D.E., and Garraway, L.A. (2013). A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. Cancer Discov. Published online January 3, 2013. http://dx.doi.org/10.1158/2159-8290.

Wiegand, K.C., Shah, S.P., Al-Agha, O.M., Zhao, Y., Tse, K., Zeng, T., Senz, J., McConechy, M.K., Anglesio, M.S., Kalloger, S.E., et al. (2010). ARID1A mutations in endometriosis-associated ovarian carcinomas. N. Engl. J. Med. *363*, 1532–1543.

Williams, S.E., Beronja, S., Pasolli, H.A., and Fuchs, E. (2011). Asymmetric cell divisions promote Notch-dependent epidermal differentiation. Nature *470*, 353–358.

Wilson, B.G., and Roberts, C.W. (2011). SWI/SNF nucleosome remodellers and cancer. Nat. Rev. Cancer *11*, 481–492.

Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science *318*, 1108–1113.

Xu, W., Yang, H., Liu, Y., Wang, P., Kim, S.H., Ito, S., Yang, C., Wang, P., Xiao, M.T., Liu, L.X., et al. (2011). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell *19*, 17–30.

Xu, X., Hou, Y., Yin, X., Bao, L., Tang, A., Song, L., Li, F., Tsang, S., Wu, K., Wu, H., et al. (2012). Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell *148*, 886–895.

Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G.J., et al. (2009). IDH1 and IDH2 mutations in gliomas. N. Engl. J. Med. *360*, 765–773.

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. Nature *478*, 64–69.

Zang, Z.J., Cutcutache, I., Poon, S.L., Zhang, S.L., McPherson, J.R., Tao, J., Rajasegaran, V., Heng, H.L., Deng, N., Gan, A., et al. (2012). Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. Nat. Genet. *44*, 570–574.

Zong, C., Lu, S., Chapman, A.R., and Xie, X.S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science *338*, 1622–1626.