The Molecular Pathology of Cancer

T.J.R. Harris and F. McCormick

Summary

Rapid technical advances in DNA sequencing, genome wide association studies and synthetic lethal screening using sh RNA is driving the discovery of both the germline and somatic mutations that are present in different cancers.
There are mutations in sets of genes involved in cellular signaling in many major cancers and some of the mutations are shared in tumors arising in very distinct anatomical locations. We review the most important molecular changes in different cancers from the perspective of what should be analysed on a routine basis in the clinic. The paradigms are EGF receptor mutations in adenocarcinoma of the lung for gefitinib treatment, KRAS mutations in colon cancer with respect to treatment with EGF receptor antibodies and the use of gene expression analysis for ER+ node negative breast cancer patients with respect to chemotherapy options. Many other examples in both solid and hematological cancers are provided including the finding of translocations of the ALK gene in neuroblastoma and non small cell lung cancer (NSCLC) and the subdivision of diffuse large B cell lymphoma (DLBCL) into ABC and GCB types by using gene expression arrays. There is an accent on how these disease subdivisions can influence therapy following the excellent examples of how imatinib is used to treat chronic myelogenous leukemia (CML) and gastrointestinal stromal tumors (GIST)
The implications of this impending molecular diagnostic revolution in cancer pathology is discussed from various perspectives including the payors, patients, physicians and the diagnostic and pharmaceutical companies. The conclusions dialogue the larger "personal medicine picture" including the economics and the fact that personalized medicine will be difficult to implement without stakeholder alignment. This paradigm shift is occurring first in cancer patient management and it is likely to continue in this arena.

TEXT BOX 1
{The year is 2019. I wake up and feel the lump under my arm that was there when I went to bed. I promised myself last night that if it was still there in the morning, the first thing to do would be to take myself off to the doctor to find out what might be wrong. Especially since my own most recent personal genetic analysis had highlighted some alleles that were associated with increased risk of lymphoma.

The appointment is set for 1130h

The doctor's assistant checks my appointment on line and brings up my electronic health record, which she sends to the hologram on the doctor's desk. I am buzzed in. I explain the problem. I am examined and the offending lymph node is subjected to a biopsy on the

spot, the tissue flash-frozen using the office NQFS (Nitrogen Quick-Freeze System) and delivered to the downstairs laboratory for a rapid molecular work-up (RMW).

The DNA will be sequenced to look for mutations in the 500 most common genes known to be involved in cancer induction and progression. Tissue sections will be analysed using high-resolution fluorescent optical images taken with third-generation Stimulated Emission Depletion microscopy (STED) profiling and Raman Spectroscopy. A blood sample is also taken for my background genomic DNA sequence analysis, concentrating on the sequence of those alleles known to predispose to cancer, both the common genes with little effect and those genes that are the rare alleles known to account for the remainder of the genetic risk predisposing to cancers of various types. Mutations in the genes coding for all the drug-metabolizing enzymes and drug distribution proteins are also obtained.

A full proteomic work-up is also undertaken to look at protein profiles and post-translational modifications. At the same time I am being prepared for New Generation Imaging so that the gross pathology of my organs can be seen in high-quality 3D. Two hours later, I review the results with the doctor on my HCD (handheld computer device), to which all the results have been downloaded and added to my electronic health record. For both my benefit and that of the doctor, the results are pre-digested and presented as a simple heat map so that a diagnosis and prognosis and drug treatment can be derived.

Fortunately, the lymph node overall molecular and cellular pathology is considered normal, and apart from the preexisting aortic valve heart disease, all other organs appear to be healthy. I am prescribed an anti-inflammatory drug consistent with potential lack of drug–drug interactions with my existing medications for valve disease.

I leave the doctor's office with a sigh of relief that all appears to be well.}

The scenario outlined in Box 1 is not as fanciful as it might at first seem, particularly as it pertains to DNA sequence-driven molecular pathological analysis of potential tumors for cancer prognosis and treatment. A large paradigm shift is occurring in cancer patient management which may actually be realized in routine clinical practice before 2019, when molecular-based patient profiles will be part of routine diagnosis (Table 1)

Table 1: Patient profiling technologies

Host Genomic DNA	Tumor Nucleic Acids	Tumor Imaging	Serum and body fluid tests
Chromosomal translocations	Gene expression profiling	Size/location	Specific proteins (PSA/CEA)
Germline mutations	Mutation analysis	Surface marker expression	Selected proteomic patterns
Genotypes at multiple loci for susceptibility alleles	Tumor specific miRNAs	Histopathology	Phospho-protein profiles
Insertions and deletions	Copy number variation	Tissue infrastructure	Peptide profiles
Gene amplification	DNA insertions/deletions	Immunohistochemistry	Other post-translational protein profiling
Epigenetic modifications (eg methylation of CpG islands).	DNA Translocations		Metabolomic analysis (e.g. hormone metabolites)
Loss of heterozygosity of alleles			
Chromatin modification by acetylation			

The development of new drugs for the treatment of cancer is taking on a much more targeted approach, in contrast to conventional cancer therapies based on general toxicity to actively dividing cells. The newer drugs are being directed at specific targets discovered by the molecular analysis of human tumors. The best examples of this are Imatinib (Gleevec) and Trastuzumab (Herceptin) respectively. The new molecular technologies that have been developed in the last five years, especially next-generation DNA sequencing, are providing a dramatic increase in the knowledge of the genomic aberrations which underlie the malignant transformation of normal cells. This includes both coding mutations, changes in the sequence of promoters and enhancers, insertions and deletions in DNA, copy number variations (CNVs) and chromosomal translocations. This knowledge is clearly defining tumor cells as having many changes that can drive their uncontrolled proliferation.

Pharmaceutical and biopharmaceutical companies are aiming their efforts at targeting these changes which drive cellular proliferation and at the same time trying to identify which tumors will be responsive to particular treatments, giving patients better outcome and a more accurate view of their prognosis. This approach is being referred to as "personalized medicine." It is more accurately defined as "more personalized medicine," where the patient and their disease can be compared molecularly to similar patients with similar disease so as to subdivide the patients into types that respond better to some treatments than to others.

Cancers occurring within a given tissue type have some distinct initiating genomic changes that define each individual tumor in terms of molecular mechanisms. Growth and survival of some cancer cells is controlled by inactivation of a single oncogene a phenomenon that has been referred to as "oncogene addiction" (Weinstein & Joe 2008). Knowledge of the molecular defects is defining a new way of classifying tumors based on molecular biology (Wood *et al*., 2007). These newer objective molecular approaches are complementing the traditionally more subjective observational techniques that are the province of the classical pathologists. These traditional tests include H&E staining, immunohistochemistry (IHC), and/or fluorescence *in situ* hybridization (FISH). An additional advantage is that it is much easier to standardize molecular tests so that identical results can be obtained between individual laboratories (Papadopoulos *et al.,* 2006).

Technology Drivers

The major technology now driving molecular diagnostics in cancer is DNA sequencing, in particular, "next-generation sequencing," defined as sequencing done using second-generation methods such as the Illumina Genome Analyzer, the Roche 454 GS FLX, or the ABI SOLiD platforms (Mardis 2008). These technologies are being used in several ways to find or examine changes in genes and their expression in numerous tumors. Cancers with mutations in mismatch repair genes will complicate this kind of analysis, but these studies and others including The Cancer Genome Atlas (TCGA) project sponsored by the NCI, are illustrative what will be done at very large scale in the next few years (Stratton *et al.*, 2009).
As the price of DNA sequencing comes down, it will inevitably replace older technologies, such as transcriptional profiling using tiled oligonucleotide arrays to examine gene expression, and the use of large-scale single nucleotide polymorphism (snp) arrays to detect deletions and amplifications in tumors. Some of this sequencing momentum is coming from efforts to sequence the human genome for $1000. The $1000 genome is a technical challenge aimed at publically funded laboratories and companies to develop technology to sequence the human genome for $1000. This project is different but related to the 1000 genomes project which aims to sequence 1000 genomes (as cost effectively as possible) to provide a comprehensive catalogue of all the common variants in the genome (Siva 2008). Both of these projects provide technology focus, but unfortunately, all these high-throughput sequencing efforts come with at least an order of magnitude (possibly two orders of magnitude) more costly data-handling bill to turn the raw sequence data into useful information. This is particularly true if that information is

to be used to drive clinical decision-making and there is also an increasing need to put the genetic information into the right biological context.

TEXT BOX 2

{It is not always appreciated that the human genome sequence data that is annotated in the existing data bases is far from base-perfect across the whole genome. A good example is provided by looking at the sequence of the region of chromosome 8q24, which is associated with susceptibility to several forms of cancer. Contrary to popular belief, the sequence in this region is not really "complete." The region corresponding to human 8q24 has been sequenced and assembled in the original human reference genome from the public effort, the original Celera assembly, and then various DNA sequences from defined individuals. There is also a "finished" sequence from chromosome 8 available (Nusbaum et al., 2006). Subsequently, a Chinese sequence (Wang et al., 2008) and data from the three genomes available so far from the 1000 genomes project have been added (www.1000genomes.org). Information is also available in this region through the analysis of BAC and fosmid ends spanning the region for 8 additional individuals from the original sequencing projects and from the most recent Korean individual DNA sequence (Kim et al., 2009). There are also individual traces at varying depths from individuals from the HapMap project and also from various sequencing centers, but it is still difficult to define a consensus-complete sequence across this region. This situation is likely to prevail for much of the rest of the genome sequence. Indeed, it has been necessary to do further sequencing to identify new snps for the more precise genetic mapping of the 8q24 region (Yeager et al., 2008). }

It would not be possible (and neither is it the point) to recreate the sequence of the human genome *de novo* from any of the new sequencing studies, although the sequences obtained will add greatly to knowledge of genome variation within individuals. Much more useful from a practical point of view however will be the precise base-perfect sequence of defined genes and their products that are known to have a role in cancer and found in tumors from many individuals.

Third generation systems (such as the Helicos, PacBio, Complete Genomics, Nanopore and other technologies) clearly are being developed with clinical utility in mind (Pushkarev *et al*., 2009). Determining the suitability of such sequencing for diagnostic applications is important. One of the prevailing issues at present (reflected above) is whether to proceed by deep-sequencing genomic DNA from patient and tumor and sorting out the information of interest from the large pool of data, or instead, sequencing regions of particular interest (Stratton *et al.*, 2009). Several selection systems for isolating defined regions of the genome are now available and others are being developed (Stratton 2008; Blow 2009). These can be used to select different types of DNA for analysis, including exons, regions of genetic association, and different types of RNA, such as transcribed regions via cDNA, miRNA, and other small RNAs of interest. Very recently, targeted capture by hybridization of 12 human "exomes" has been used to generate over 300 Mb of coding regions to identify common and rare variants in the genome, including a gene involved in a dominant inherited disorder (Ng *et al*. 2009b). This same approach will inevitably be used to find rare mutations predisposing to cancer.

Second-generation digital read-out methods for counting transcripts or for measuring gene amplification are also being developed and may form the basis of second-generation tests for diagnosis based on gene expression profiles as an alternative to RT-PCR-based methods for amplifying up- or down-regulated genes. Other transcriptome sequencing studies are being done specifically to look for fusion transcripts representing the expression of translocated genes. The Bcr-Abl transcript was rediscovered in CML cells as was the TMPRSS2-ERG translocation in a prostate cancer cell line and prostate tissues in a study of this kind (Maher *et al*., 2009).

The other major technology drive for cancer screening which has yet to demonstrate utility diagnostically or for prognosis is GWAS (Genome-Wide Association Studies). The development of Hap Map and access to relatively cheap chip-based methods to analyze single nucleotide polymorphisms (SNPs) across the human genome has led to the publication of many association studies in cancer and other diseases (Manolio *et al.,* 2008). The utility of such studies has been called into question by some (Goldstein, 2009) but as far as cancer is concerned, GWAS can lead to novel insights into the biology of the disease (Hirschorn, 2009). GWAS, however, has yet to show any utility from a cancer screening point of view, largely because the associations (by definition) are in common alleles and each on its own confers very small additional risk of disease (Kraft and Hunter, 2009). Combinations of low-risk alleles might have much greater impact, as suggested by analysis of cancer susceptibility genes in mice, but these combinations will be hard to identify and validate. In contrast, no one would question the utility of screening at-risk people for mutations in genes known to confer a much higher increase in relative risk of cancer predisposition (Foulkes, 2008). One of the best examples of this approach is measuring mutations in BRCA1 and BRCA2, which confer considerable increases in the risk of breast and ovarian cancer in women with mutations in one or the other of these genes and different alleles of BRCA1 and BRCA2 have different predisposing effects.

The common alleles conferring low risk that have been identified by GWAS have not yet been shown to be linked to outcome (only susceptibility), and they do not account for all of the genetic risk of the cancers in question (Altshuler *et al*., 2008); hence, the idea that there are as yet unfound (from traditional linkage disequilibrium mapping in cancer families or by association studies) mutations in rare alleles of certain genes that confer the additional risk. The potential for deep sequencing of DNA to look for rare mutant alleles is alluring and will probably provide a new set of genes that will be of utility diagnostically (Ng *et al*. 2009b). These mutations will be discovered by sequencing both whole genomes and defined regions. Alternatively, there are likely to be more common genes that confer an even smaller additional risk that may be found by additional GWAS in larger cohorts of patients and in populations where the allele frequencies are different (Manolio *et al.,* 2009). Clearly, much follow-up work remains to be done on the regions found to be associated with a modest increase in risk of disease; however, one has to question the cost-effectiveness of continued GWAS in cancer unless directed towards outcome/treatment or diagnosis (Foulkes, 2008; Savage, 2008).

Two other related technologies are also driving the molecular understanding of cancer and have therapeutic implications. Although developing miRNAs as therapeutic entities may be important in the future, their utility in uncovering genes involved in cancer is

already proven (Couzin, 2008). Several loss-of-function RNAi screens have been done, where genes involved in proliferation and survival of cancer cells have been identified. Using retroviral vectors, short hairpin RNAs (shRNAs) are delivered to tumor cells. shRNAs that knock down a gene required for continued cell growth will be eliminated from the overall shRNA population after a series of cell replications; these RNAs (and hence their targets) can then be identified by scanning (by arrays) for a bar code tag attached to the individual shRNAs. In Diffuse Large B Cell Lymphoma (DLBCL), this kind of screen identified CARD 11, a gene involved in NF-κB signaling (Ngo *et al*., 2006). This approach has since been extended to investigate mechanisms of resistance of DLBCL cells to IKκB inhibition by small-molecule inhibitors (Lam *et al*., 2008).

A scaled-up version of this same kind of approach has also been applied to leukemic cells from patients. Not only were mutations found in known genes (e.g. JAK2 and KRAS), but also, mutations were found in several genes not seen before in this disease (Tyner *et al.*, 2009). It is not hard to imagine how these relatively simple technologies can be applied in the clinic, especially for leukemias and lymphomas. Pathways not previously linked to KRAS have also been found in RNAi-driven "synthetic lethal" screens recently published by several groups (Luo *et al*., 2009; Scholl *et al*., 2009; Singh *et al*., 2009). The results suggest several novel points of intervention targeting genes that do not necessarily pick up mutations themselves to drive proliferation but are nevertheless important components of the abnormal molecular drive. These may indeed provide new drug targets. Understanding the linkages between the genes, mutations, and normal and abnormal signaling pathways (although complex) is fundamental to the understanding of the molecular pathology of cancer and holds the key to not only patient stratification but also to new treatments (Chang *et al*., 2009).

THE CANCERS

Lung Cancer

It should be remembered that advanced non-small-cell lung cancer is the leading cause of cancer-related death in the world, and it has a particularly poor prognosis that has changed little over the past 30 years. This is despite great increases in knowledge about the molecular and cellular biology of the disease (Herbst *et al.*, 2008).
As in other cancers, large genome scans have been done in lung cancer to look for associated susceptibility loci. A reproducible region on chromosome 15q24-25.1 has been found. The region includes two genes coding for sub units of the nicotinic acetyl choline receptor but whether these are of biological importance remains to be seen (Amos *et al*., 2008; Hung *et al*., 2008; Thorgeirsson *et al*., 2008).

Several studies have attempted to correlate gene expression signatures with patient survival/tumour proliferation in different forms of lung cancer. In the early studies, relative small numbers of samples were tested and signatures involving small numbers of genes were established as correlating with relapse or survival (Potti *et al*., 2006; Chen *et al.*, 2007). The most comprehensive study so far has involved 440 lung cancer samples collected from six clinical centers in the same consistent way. A training set was first

established and the results verified in additional independent data sets. The study was particularly focused on prognosis and survival time. A gene constellation including cell cycle-related genes was associated with poor outcome and genes involved in a more differentiated cell phenotype correlated with better outcome (Shedden *et al*., 2008). This study has set the scene for additional large studies looking at treatment modality (e.g. types of chemotherapy), gene expression profile, and outcome (Xie and Minna, 2008). Attempts have also been made to look at the correlation of outcome/relapse with miRNA expression levels in lung cancer (Yu *et al*., 2008; Lebanony *et al*., 2009).

A major advance in the treatment of lung cancer was the finding that mutations in the EGF receptor gene that activated the kinase domain or EGFR gene amplifications were common in Asian females with adenocarcinoma of the lung who had never smoked. These cancers were particularly responsive to inhibitors of the EGFRTK, such as gefitinib and erlotinib. More than 80% of these mutations are either point mutations (L858R) in exon 21 or are exon 19 deletions. It will be important now to look for EGFR mutations in all patients (whether smokers or not) with lung cancer, as inhibition of the activated enzyme has been shown to have considerable clinical benefit. This was particularly clear in the IPASS trial with gefitinib in clinically selected Asian patients with EGFR mutations or amplifications (Mok *et al.,* 2009). It is noteworthy that treatment of patients who do not have mutations in the EGFR with enzyme inhibitors can worsen their outcome, and the inhibitors do not work at all in patients with KRAS mutations. In fact, the presence of KRAS mutations is a better predictor of response to gefitinib than the presence of EGF receptor mutations.

Sooner or later however, the cancer cells become resistant to the drugs, either by accumulating additional mutations in the EGFR gene (e.g. the T790M mutation) or by "compensatory" mutations elsewhere, which provide activation of other connected pathways to overcome the proliferative drive mediated by the EGFR (Figure 1).

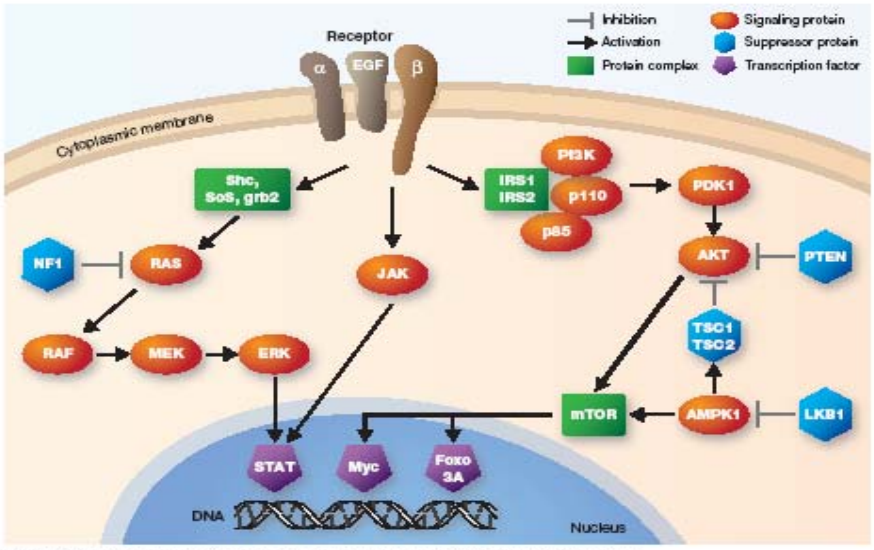

Figure 1. Schematic of downstream signaling from epidermal growth factor receptor complex

AKT:	Protein kinase B	LKB1:	Liver kinase B1	PTEN:	Phosphatase and tensin protein
AMPK1:	AMP-activated kinase	MEK:	Mitogen-activated kinase	RAF:	Ras-like factor
EGF:	Epidermal growth factor	mTOR:	Mammalian target of rapamycin	Ras:	Ras oncogene
ERK:	Extracellular regulated kinases 1-2	Myc:	Myc oncogene	Shc:	Src homology domain
Foxo 3A:	Forkhead family TF	NF1:	Neurofibromin	SoS:	Son of sevenless
grb2:	GTP exchange factor receptor bound-2	PDK1:	PI-dependent kinase	STAT:	Signal transducer/activator of transcription
IRS:	Insulin receptor substrate	PI3K:	Phosphatidylinositol 3 kinase	TSC:	Tuberous sclerosis complex
JAK:	Janus kinase				

Figure 1: The EGF receptor signaling and related pathways

These include Met gene amplifications that correlate with poor outcome (Bean *et al*., 2007; Engelman *et al*., 2007). Detection of EGFR mutations by examining lung cancer cells in the circulation or by direct measurement of mutations of DNA in the circulation may become a convenient way to screen for these mutations during initial work-up and during treatment (Maheswaran *et al*., 2008). The dramatic responses to EGFR inhibitors in subsets of patients is a strong argument for attacking mutant oncoproteins from a therapeutic perspective with cocktails of drugs, much as the current clinical practice for treating HIV and mutations that arise in the virus polymerase and protease. As in that analogy, there is a need to be able to measure in real time the genes and the mutations that are being selected (Engelman & Janne 2008).
Studies of cDNAs from a retrovirus-mediated expression library have identified a fusion gene in NSCLC cells derived from an inversion on chromosome 2p {inv(2)(p21p23)}. The gene product (EML4-ALK) is a transforming oncogene and was found in about 5–7% of the NSCLC patients (Soda *et al*., 2007). Novel iso-forms of the fusion gene have also been described (Choi *et al*., 2008). This genetic change in the tumor cells appears to be mutually exclusive to EGFR, KRAS or BRAF mutations and suggests that routine measurement of this sequence abnormality be done in clinical trials testing inhibitors of ALK. Quite remarkably, clinical data on benefits of blocking ALK with PF-02341066, a Met inhibitor that also inhibits ALK were presented at the 2009 ASCO meeting, just two years after the EML4-ALK translocation was identified and the compound has just entered Phase III clinical trials in patients with alterations in ALK. Germline and somatic changes in ALK including gene translocations, amplifications and mutations are also found in anaplastic large cell lymphoma and neuroblastoma (Mosse *et al*., 2009).

Large-scale sequencing has confirmed and extended the number of genes found with somatic mutations in lung cancer. The most frequently mutated genes include tyrosine kinases such as EGFR, ERBB4, and KDR (a VEGF receptor) and several tumor suppressor genes, including NF1, APC, RB1, and ATM. The tumor suppressor LKB1 (STK11) is also frequently deleted or mutated (Koivunen *et al*., 2008). This tumor suppressor gene, germ line alterations in which are responsible for Puetz-Jeghers syndrome, is becoming increasingly of interest as it is involved in controlling cell growth and metabolism via AMP-activated protein kinase (Shackleford and Shaw, 2009). The AMP kinase axis may be important both diagnostically and therapeutically, given the reliance of cancer cells on aerobic glycolysis (the Warburg effect) (Vander Heiden *et al.,* 2009).

The mutual exclusivity in mutations in KRAS and EGFR in the same tumor samples was also confirmed by genome sequencing, and KRAS mutations were (as shown previously) most often associated with tumors from smokers (Ding *et al.,* 2008). Certain genes, including PTEN and the EGFR, were mutated in tumors at lower than average mutation rates, suggesting that their activation or inactivation was a sufficiently positive selection factor to reduce the accumulation of additional mutations. A comprehensive picture of the mutations in the genes in signaling pathways in lung adenocarcinoma is shown below (Figure 2) (Ding *et al*., 2008).

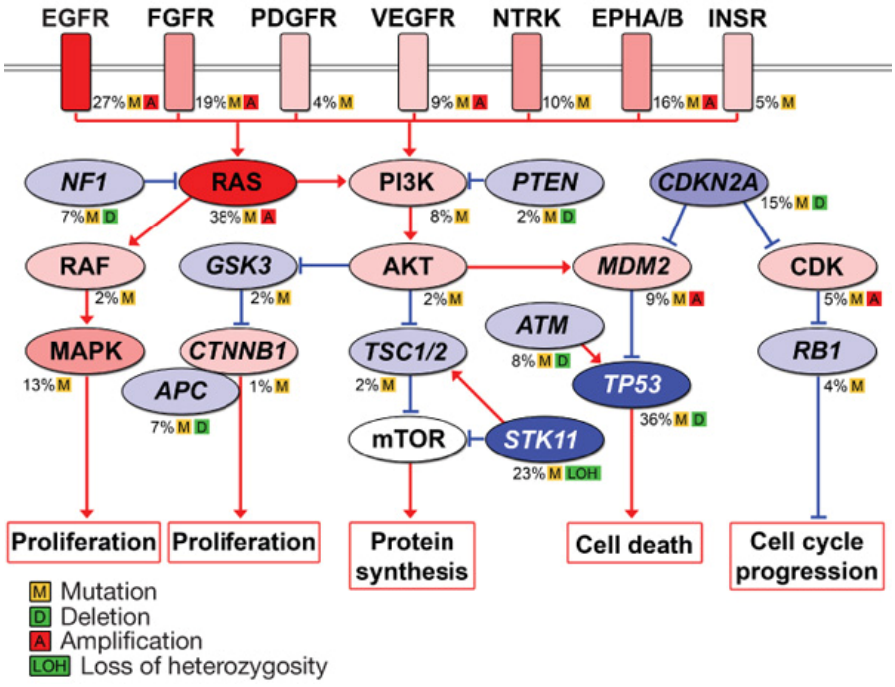

Figure 2: Significantly mutated pathways in lung adenocarcinomas

Genetic alterations in lung adenocarcinoma frequently occur in genes of the MAPK signalling, p53 signaling, Wnt signaling, cell cycle and mTOR pathways. Oncoproteins are indicated in pink to red and tumor suppressor proteins are shown in light to dark blue. The darkness of the colors is positively correlated to the percentage of tumors with genetic alterations. Frequency of genetic alterations for each of these pathway members in 188 tumors is indicated. Genes (*EGFR*, *FGFR1*, *FGFR4*, *KDR*, *EPHA3*, *KRAS*, *NRAS*, *MDM2* and *CDK6*) lying in regions of focal amplification were analyzed for the percentage of samples with copy number amplification. Samples with greater than 2.5 and fewer than 1.5 DNA copies were considered as amplified or deleted, respectively. Selected components of each pathway are shown in the figure.
From Ding et al., 2008: *Nature* 455, 1069-1075

This kind of analysis, which will inevitably be derived for other tumor types, is instructive in directing attention to potential therapeutic targets. For example, it was found that 17 genes in the mTOR pathway are mutated in more than 30% of the tumors sequenced (excluding those with KRAS mutations), indicating that rapamycin and other mTOR inhibitors in development could be examined in lung cancer patients with mutations in this pathway. Moreover, MEK inhibitors could be tested in those tumors with NF1 mutations; and in patients with KDR mutations, it would be appropriate to try semi-selective kinase inhibitors such as sorafenib or sunitinib (Sharma *et al*., 2007).

Colorectal Cancer (CRC)

Despite the classical molecular model of the development of colorectal cancer being put forward over 20 years ago the histopathological examination of tumor biopsies and the degree of lymph node involvement are the mainstay for the classification of colorectal cancer. Increasingly, however, molecular markers are being used to further subdivide the disease and its progression (Jones *et al*., 2008b). The high-risk hereditary genes include APC (adenomatous polyposis coli), which acts via β-catenin to control Wnt gene-directed cellular signaling, and mutations in certain DNA repair genes (MLH1, MSH2, MSH6, and PMS2), which were discovered by familial studies. Testing for mutation status of these genes is recommended in high-risk families where early onset disease is apparent. Deletion of the long arm of chromosome 18q is a common cytogenetic abnormality in CRC and correlates with poor prognosis, although studies of some of the "cancer" genes in that location (e.g. SMAD 4) have failed to show any reliable connection to outcome (Walther *et al*., 2009). For the identification of alleles conferring lower risk of disease, genome-wide association studies have shown association to 8q23/24 and to chromosome 10p14 (Zanke *et al*., 2007; Tomlinson *et al*., 2008).
Stratification of colorectal cancer by using gene expression signatures has been less useful than in many other cancers. Garman *et al*. (2008) recently published a 50-gene model that accurately predicts recurrence in early-stage CRC, but the study did not examine a very large number of patients. None of these signatures is close to being used in clinical practice at present.
The chemotherapy regimens for metastatic CRC (FOLFIRI or FOLFOX) consist of fluorouracil derivatives plus irinotecan, a topoisomerase inhibitor, or oxiplatin. It has been known for some time that the metabolism of 5FU (an inhibitor of thymidylate synthetase) is subject to variation mediated by polymorphisms in catabolic genes such as dehydropyrimidine dehydrogenase (DPD). Genetic polymorphisms in UDP-glucuronyltransferases (UGT1A1) that metabolise irinotecan are also clinically relevant in determining the risk of neutropenia, which is known to occur with this drug in some patients (Innocenti *et al*., 2009).

By far the most important recent discovery in the clinical management of colorectal cancer is the association of mutations in KRAS and the efficacy of monoclonal antibodies targeting the EGF receptor, such as vectibix and cetuximab. K-Ras is a small signaling molecule that is connected to the EGFR via Sos/Grb (See Figure 1). Some tumors harbor KRAS somatic mutations in exon 2 of the gene (codons 12 and 13), which activate

KRAS signaling activity by compromising the hydrolysis of Ras-bound GTP to GDP, resulting in constitutive activation of the MAPK pathways. KRAS mutations occur in about 40% of colorectal cancers. Recent data has shown that tumors with KRAS somatic mutations do not respond to EGF receptor inhibition (Karapetis *et al*., 2008). A European clinical trial failed to gain approval of vectibix for treatment of late-stage colon cancer, and further analysis of the data showed that only those patients with tumors with wild-type KRAS responded to treatment. Despite lack of FDA 510K approval, it is now accepted clinical practice to determine the KRAS status of all advanced colorectal cancer patients (Messersmith and Ahnen, 2008). It has also been shown that patients treated with FOLFIRI, anti-VEGFR (bevacizumab) antibody, and an anti-EGFR monoclonal antibody fare worse if their tumors have KRAS mutations (Tol *et al*., 2009).

Breast Cancer

The two most important high penetrance genes predisposing to breast cancer are the dominant genes BRCA1 and BRCA2. Carriers of these gene mutations have a 15-20 fold increase in risk of disease, compared to the general population, and are associated with early onset disease. BRCA1-related breast tumors are usually high grade, do not express hormone receptors, and are referred to as triple negative cancers. These genes affect the capacity of the cells to undergo homologous recombination and DNA repair. The prognosis of patients with breast (and ovarian) cancers that harbor BRCA mutations is generally poorer than those with tumors expressing hormone receptors. Other genes conferring lower risk have been identified in some populations, the most notable of which is CHEK 2 (1100 ΔC) (See Foulkes, 2008). Also, germ-line mutations in PTEN are associated with florid bilateral fibrocystic breast disease, and a substantially increased risk of breast cancer and mutations in CDH1 are associated with an approximate 40% risk of lobular breast cancer and diffuse gastric cancer.

Genome-wide association studies have pinpointed other regions where common alleles conferring some additional relative risk of disease have been identified, including 2q35, 16q12, and 5p12 (Easton *et al*., 2007; Stacey *et al*., 2007, 2008). Alleles of FGFR2 have also been associated with increased risk of sporadic breast cancer (Hunter *et al*., 2007). More recent GWAS have identified additional regions, including 1p11.2, 3p24 and 17q23.2, and the RAD51L1 gene involved in DNA recombination (Ahmed *et al*., 2009; Thomas *et al*., 2009).

Analysis of snps in the p53 pathway revealed allelic variants at position 309 in the Mdm2 promoter. Individuals with the G-allele express higher levels of Mdm2, and hence lower levels of p53. As a result, they may be at increased risk for multiple types of cancer, including breast cancer and prostate cancer and may respond more poorly to treatment (reviewed in Vazquez et al., 2008). The relationship between breast cancer susceptibility genes and relative risk is shown in Figure 3 and demonstrates not only how they were found (family studies, re-sequencing, or GWAS) but also the relative importance of measuring them in the course of patient management.

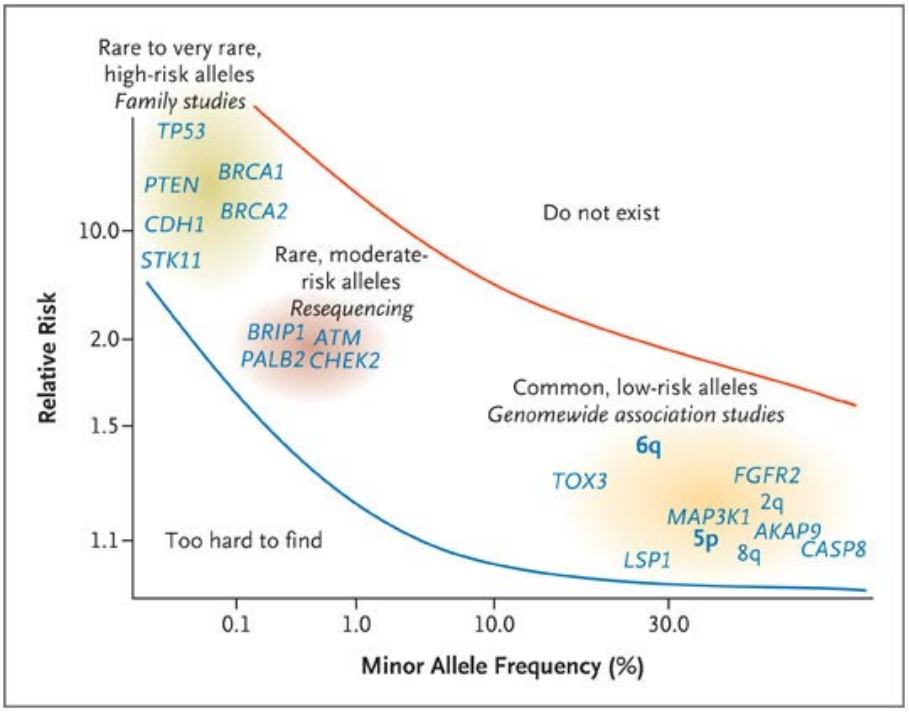

Figure 3: Germline mutations conferring susceptibility to breast cancer

All known breast-cancer susceptibility genes are shown between the red and blue lines. No genes are believed to exist above the red line, and no genes have been identified below the blue line. High-risk syndromic genes are highlighted in green. The moderate-penetrance genes (highlighted in red) have an approximate relative risk of 2.0. Some of the common, low-risk genes are shown in orange. SNPs in *FGFR2* and *TOX3*, and those on chromosomes 5p and 2q, specifically, increase the risk of estrogen-receptor–positive breast cancer (Foulkes 2008).

Recent developments in breast cancer treatment have shown that targeted therapies can have dramatic effects but these therapies must be guided with the use of companion diagnostic assays that identify the appropriate subset of patients. Trastuzumab (Herceptin) is a humanized monoclonal antibody directed at the cell surface tyrosine kinase receptor, ErbB2. Early in the development of this drug it was observed that only those tumors which overexpressed ErbB2 responded. Overexpression of ErbB2 occurs in only about 20–25% of all breast cancers. This information led to the co-development of a diagnostic test, the Herceptest, which measures the level of Her2 protein by immuno-histochemistry. FISH of chromosomal spreads is also used to look for Her2 gene amplifications. These tests identify tumors with overexpression of ErbB2 and the patients who are therefore eligible for trastuzumab treatment. Unfortunately some 20% of patients who get treated with trastuzumab are never tested for Her2 status.

Breast cancer has also been defined by gene expression profiling as comprising at least 4 major phenotypes: Luminal A, Luminal B, Her2-like, and Basal-like (Perou *et al*., 2000;

Ramaswamy and Perou, 2003). Both Luminal A and Luminal B sub-types are generally estrogen receptor-positive and therefore candidates for treatment with hormonal blockade with drugs such as tamoxifen. The Luminal B sub-type has an increase in the expression of genes associated with cell proliferation. Luminal B tumors have a poorer overall outcome than the Luminal A subtype.

The objective of using gene expression profiling tests in the clinical management of breast cancer is to provide information that enhances the predictions of clinical outcome over what can be obtained by traditional pathology. Several tests have been developed (Sotiriou and Pusztai, 2009). The MammaPrint gene test (Agendia) measures the expression of 70 genes derived from a set of 78 patients with node-negative breast cancer who had received no adjuvant therapy. It provides a prognostic score which stratifies patients into good or poor prognosis and works on fresh or frozen tissue. The Oncotype DX test (Genomic Health) measures the expression of 21 genes by RT-PCR and can be used with formalin fixed paraffin embedded tissue (FFPET). This test predicts the risk of recurrent disease in ER-positive, node-negative breast cancer and identifies patients with low risk of recurrence (Table 2). Information from these tests is designed to inform clinical decision-making about if and when to start chemotherapy. The Oncotype DX test provides a score that puts patients into three categories: low, intermediate or high risk of recurrence. Low risk patients would not be managed with chemotherapy whereas high risk patients would be. Unfortunately, many patients have an intermediate profile which is not clarifying one way or the other. Also, as pointed out by Sotiriou and Pusztai (2009), another limitation of these tests is their inability to subdivide ER-negative or triple-negative patients into higher- or lower-risk categories.
Now that prospective studies that are underway (MINDACT and TAILORx) testing MammaPrint and Oncotype DX, respectively, more precise information about the utility of these (expensive) tests in the clinical management of breast cancer patients will be forthcoming. In the meantime, other signatures are being described, including those reflecting additional genetic changes in tumors such as loss of PTEN or response to treatment (Paik, 2007; Saal *et al*., 2007).

Table 2: Gene Expression Profiles of Clinical Utility

Test	Profile	Utility
Mammaprint (Agendia)	70-gene expression profile for frozen tissue	Prognostic for relapse and time to progression in node-negative breast cancer
Oncotype DX (Genomic Health)	21-gene signature (FFPET)	Chemotherapy decision for ER^+ node-negative breast cancer
DLBCL/Burkitt's lymphoma	Gene expression array distinguishes Burkitt lymphoma from DLBCL	EPOCH-R for Burkitt's lymphoma. CHOP for DLBCL with or without Rituxan
DLBCL	Gene expression signature distinguishes ABC from GCB tumours	ABC tumors have worse prognosis. Potential for Bortezomid or other NFkB-regulating agents in treatment of ABC DLBCL
MYCN/CD44	Expression levels plus 17 other genes	Predictive value for outcome in neuroblastoma
Multiple Myeloma	Expression profile of multiple genes	Defines follow-up of bortezomib treatment
AML	2-gene identifier	Sensitivity to farnesyl transferase inhibitor tipifarnib

A key observation by the Ashworth group in the UK was that breast cancer cells harboring mutations in BRCA1 or BRCA2 are profoundly sensitized to the new poly(ADP-ribose) polymerase (PARP) inhibitors (Farmer *et al*., 2005). Several PARP inhibitors are now in clinical trials in breast cancer targeting triple-negative patients with BRCA1 mutations. An inhibitor of PARP has also been used in a Phase 0 trial where the "profiling" of the patient is part of the trial process (Kummar *et al*., 2009). Unfortunately, the complexity of the mutations in the BRCA1 locus and the price of the testing may affect the ability to do this kind of screening routinely to identify the right patients for clinical trials (Dalla Palma *et al*., 2008). Synthetic lethal screens using siRNA are not only being applied directly to cancer drug discovery (Fong *et al*., 2009; Inglehart and Silver, 2009) but are also being used to find other cellular targets that are synthetically lethal in combination with the PARP inhibitors. CDK5, a cell cycle kinase which is abnormal in a small percentage of breast cancers, has been identified by this means. Patients with gene loss by 7q deletion or reduced CDK5 expression and patients with BRCA mutations may therefore also benefit from PARP inhibitors (Turner *et al*., 2008). Additionally, tests that can predict responses to chemotherapy are also needed. Increasing data are showing that the efficacy of tamoxifen, which is metabolised to the active drug endoxifen by CYP2D6, is compromised by inherited insufficiency of the enzyme (poor metabolizer phenotype caused by gene loss or mutation) or by co-administration of drugs that are also metabolized by CYP2D6, including some SSRIs. Genotype- and phenotype-based cytochrome P450 testing is available and should be used to determine the status of Cyp2D6 and the potential consequences of drug–drug interactions in women eligible for tamoxifen treatment (Hoskins *et al.,* 2009).

Ovarian cancer

Ovarian cancer cure rates are low, owing to the generally late presentation of patients with the disease. In those cases caught early, treatment with taxol and platinum-based compounds leads to a good outcome. Some screening, using trans-vaginal sonography and monitoring of CA125, a protein antigen found in serum, is available for high-risk patients. Ovarian cancer can be differentiated into low-grade and high-grade tumors based on genetic information. High-grade tumors are often associated with BRCA1 and BRCA2 mutations and LOH on chromosome 7q and 9p. Mutation and loss of the TP53 gene occur in a high proportion of both familial and other tumors. Low-grade tumors have mutations in KRAS, BRAF and PI3KC (Bast *et al.,* 2009). As these authors suggest, analysis of p53 and KRAS to differentiate low-grade serous and non-serous type I cancers from high-grade serous and endometroid type II cancer would be useful from a treatment perspective. Type I cancers should respond better than type II cancers to drugs targeting the PI3K and Ras-MAPK pathways, and gene expression studies directed to this end have recently been published (Crinjs *et al*., 2009). Nevertheless, the clinical management of ovarian cancer (and other cancers, for that matter) targeting mutations in the PI3K pathway will be complex. There will be a need to look for not only mutations in PI3K and the regulatory sub-unit, but also PTEN (deletions), AKT mutations (e.g. E17K), and amplifications of other genes such as AKT2. Dual PI3K-mTOR inhibitors, isoform specific PI3K inhibitors, and specific AKT- or mTOR-targeted molecules may have potential use. Several compounds of this kind are in clinical development (Engelman, 2009; Liu et al., 2009).

Low levels of two different microRNAs, specifically Dicer and Drosha, involved in splicing pre-mRNA, have been shown to be associated with advanced ovarian tumors and poor surgical resection, both of which are already known as markers of poor prognosis (Merritt *et al*., 2008).

In a relatively uncommon form of ovarian cancer (granulosa cell tumors) it has recently been shown by transcriptome sequencing that mutation of the FOXL2 gene, a transcription factor, is a potential driver for the uncontrolled proliferation of granulosa cells. Analysis of this gene for mutations might be useful to aid the diagnosis of this particular form of ovarian cancer (Shah *et al*., 2009). Although the TCGA consortium results have not been published, the accumulating data suggests that there are several molecular subtypes of ovarian cancer, based on loss of suppressor genes or patterns of mutations in genes in different signaling pathways (http://cancergenome.nih.gov).

Prostate Cancer

Rising levels of prostate-specific antigen (PSA), followed by examination and biopsy, is the current approach to the diagnosis of prostate cancer. However, the widespread adoption of PSA screening has led to a major shift toward the diagnosis of small, early-stage tumors, and has raised difficult questions regarding their clinical significance (Dall'Era *et al.,* 2008). In contrast to other solid tumors it has been difficult to determine the underlying genetics of prostate cancer, although there is a significant familial risk. With the exception of BRCA2, which confers an increased risk for lifetime prostate

cancer of over 20-fold, there appear to be no other predisposing genes of large effect. GWAS of prostate cancer have highlighted several regions of the genome that are associated with susceptibility to disease. Of particular note is the association of multiple alleles in 8q24, especially in black African-Americans. Some of these alleles are also implicated in susceptibility to many other solid tumors including breast and colon cancer (Eeles *et al.,* 2009). These low risk susceptibility alleles account for about 20% of the familial risk of prostate cancer. Although the associated alleles are some distance from the MYC gene, increasing evidence indicates that these susceptibility regions operate by long-range control of MYC gene expression through enhancer/promoter interactions (Freie and Eisenman, 2008; Pomerantz *et al*., 2009; Sotelo *et al*., 2009; Tuupanen *et al*., 2009; Jia *et al.,* 2009). Some recent evidence also suggests that the effects of the 8q24 association interact with regions of cancer susceptibility on 17q (Zheng *et al*., 2008).

A variant snp in close proximity to the MSMB (β microseminal protein) gene on chromosome 10q11.2 has also been associated with prostate cancer (Eeles *et al*., 2008; Thomas *et al*., 2008). The variant appears to be involved in the control of expression of MSMB, which is a potential candidate gene largely owing to the location of the expressed protein (Lou *et al*., 2009).

Gene fusions involving transcription factors have been found in prostate tumors. Fusions of the 5′ untranslated region of the TMPRSS2 gene (a serine protease) on chromosome 21 (21q22.3) with ETS transcription factor family members (ERG, ETV1, or EYV4) have been identified by FISH in many late stage tumors. TPMRSS2-ERG fusions, apparently caused by an internal deletion on chromosome 21 (ERG is at 21q22.2), are associated with an androgen-independent, highly aggressive and lethal subtype of the disease (Mehra *et al*., 2008) and absence of these fusions or PTEN loss correlates with a favorable outcome and less likely disease recurrence (Yashimoto *et al.,* 2008). It remains to be determined if this subtype would benefit from especially aggressive treatment early on, but clearly examining TPMRSS fusions and PTEN status in prostate cancer will become important in the future (Kumar-Sinha *et al*., 2008; Squire 2009).

It is somewhat surprising, given that prostate cancer is a common disease and often lethal, that there is rather less of an understanding of the genetic basis of this cancer compared to others (Witte, 2009).

Pancreatic Cancer

Pancreatic cancer is one of the most rapidly fatal of all cancers, owing primarily to lack of early detection. Several genetic lesions have been found in these cancers, including mutations in the KRAS, TP53, CDKN2A, and SMAD4 genes. Recent genome-wide studies of gene expression and genome organization in 24 pancreatic tumors have revealed many point mutations, which defined alterations in a core set of 12 signaling pathways (Jones *et al*., 2008a). The pathways included: apoptosis, DNA repair, cell cycle control, Wnt/notch pathway, Hedgehog signaling, and KRAS. The implications for therapy are important, given that the studies highlight the hedgehog and the Wnt/notch pathways, both of which are targets for new therapies (Scales & de Sauvage 2009). Quite recently, by candidate gene studies, the PALB2 gene (with a not well-understood function), has been implicated in familial pancreatic cancer (Jones *et al*., 2009). This

finding follows other studies showing involvement of PALB2 in Fanconi's anemia and breast cancer (Turnbull *et al*., 2008). The other major familial pancreatic cancer susceptibility gene appears to be BRCA2 (the product of which binds to the PALB2 protein), implicating, particularly, DNA repair abnormalities in this disease. It also suggests that pancreatic cancer patients with abnormal signaling through these pathways might benefit from exposure to therapies targeting DNA repair mechanisms such as PARP inhibitors (Xia *et al*, 2006).

Renal cell carcinoma

Studies of familial renal cancer have unlocked the basic understanding of the molecular basis of the sporadic disease and led to more rational treatment regimens (Guibellino *et al*., 2009). Four major familial forms of the disease exist, each one of which has contributed useful and connected information. Von Hippel-Lindau (VHL) disease patients suffer from clear cell renal cancer that is primarily caused by dominant germ-line mutations in the VHL tumor suppressor gene on chromosome 3. Inactivation of the suppressor has multiple effects, including interference with the degradation of HIF1α. Papillary renal cell carcinoma (PRC) is the second most common form of the disease and can be subdivided into two types, based on gene expression profile and histopathology. Mutations in MET, the receptor for a growth factor (HGF), which are common in type I hereditary PRC, were discovered through classic linkage genetics (Schmidt *et al*., 1997).

Since that time, activation of Met signaling through mutation and overexpression has been described not only in hereditary renal cancer but also in sporadic forms of the disease and in many other cancers. Many different mutations in the TK domain of the Met protein have been described. This indicates that the MET gene should be sequenced in all renal cell carcinoma (RCC) patients to determine mutation and/or amplification status.

The HGF/Met interaction has suggested several ways forward clinically for PRC, including the development of anti-HGF monoclonal antibodies. Given that several approved TK inhibitors exist, the development of combination therapies targeting the particular pathways activated in renal cell carcinoma, including not only Met but also enzymes downstream of Met, is a real clinical option. Sunitinib, an effective inhibitor of Met, is now being used as first line therapy to treat RCC, pazopanib from GSK,a more selective TK inhibitor had a recent favourable ODAC review for the treatment of RCC and XL880 a multikinase and potent Met inhibitor from Exelixis is in Phase II clinical trials in RCC. As in other malignancies, the potential to combine these drugs with rapamycin analogues that target mTOR is particularly encouraging.

One of the other forms of hereditary renal cancer is caused by inactivation of fumarate dehydratase (hereditary leiomyomatosis and RCC), which may suggest that further study of genes coding for metabolic enzymes in cancer is warranted, given the recent findings of IDH 1 mutations in both glioma and AML (Downing, 2009).

Hepato-cellular carcinoma

Cancer of the liver is associated with environmental insults such as toxin exposure and chronic infection with hepatitis B and C virus and is more common in Africa and Asia than in Europe and the United States. One of the problems of this disease, given that patients with pre-exposure to insult are often carefully monitored, is predicting recurrence after treatment that results in cure in some patients but not in others. FFPET is the predominant source of archived tumor samples and hence is readily available for most patients at the time clinical outcome becomes available. Gene expression in FFPE hepato-cellular carcinoma biopsy tissue demonstrated that samples from 90% of the patients yielded data of high quality. Some of the samples were over 20 years old. The study supports the conclusion that FFPET can be used to derive transcriptional profiles that correlate expression signature with outcome (Hoshida *et al*., 2008).
Hepatocellular carcinoma cells frequently contain genetic abnormalities, including chromosomal deletions, amplifications and loss of heterozygosity, some of which involve known oncogenes. In a recent comprehensive study of the molecular aberrations in HCC, overexpression of the VEGF gene via 6p21 gain was identified (Chiang *et al*., 2008). Good evidence has also been found for disruption of WNT-β catenin signaling via mutations in AXIN1 or CTNNB1 (see Minquez *et al*., 2009). There is some potential for partitioning patients using gene expression levels, since 50% of tumors have been shown to have WNT or AKT pathway activation (Boyault *et al*., 2007). It is encouraging that sorafenib, an inhibitor of several tyrosine kinases affecting different signaling pathways, including Raf1, B-Raf, VEGFR2, PDGFR, and CKit, has shown efficacy in the treatment of this disease (Llovet *et al*., 2008). Aberrant activation of the Ras/Raf/ERK pathway in HCC at least in part occurs through decreased levels of the Sprouty-related protein Spred, and expression of this protein is inversely correlated with the incidence of tumor invasion and metastasis (Yoshida *et al*., 2006). This may account, in part, for the sensitivity of HCC to the Raf kinase/ KDR inhibitor sorafenib, since patients with elevated MAPK activity show better clinical responses to this agent. Despite the fact that combination therapies will still be needed, the prospect of being able to subdivide HCC patients based on the activation status of different signaling pathways in the tumor cells is an exciting therapeutic prospect (Llovet and Bruix, 2008).

Glioma

One of the largest DNA sequencing efforts in cancer is the Cancer Genome Atlas study (TCGA) sponsored by the NCI. The pilot project has looked at copy number, gene expression, and other molecular changes in 206 glioblastomas. An integrated view of the glioblastoma genome from this study identified changes in genes in three major pathways: RTK signaling, the p53 and the RB suppressor gene pathways. There were for example, frequent deletions and mutations in PTEN, and 86% of the samples had a change in the RTK/PI3K pathway. Somatic mutations in the p85 regulatory gene (PIK3RI) were found in about 10% of the gliomas and it is becoming increasingly clear from other sequencing studies that mutations in the gene coding for the catalytic sub unit of PI3K (p110) are quite common in other cancers as well. Mutations in the genes

coding for the two PI3K sub units appear to be mutually exclusive. For a comprehensive review of the role of PI3K in cancer cell signaling see Engelman (2009).
It had already been shown previously in glioma that distinct mutations in the extracellular domain of the EGF receptor predict sensitivity to EGFR TK inhibitors but only when PTEN is active (Mellinghoff *et al*., 2005). In the p53 pathway, there were amplifications of MDM 2 and 4 and frequent ARF deletions. In the RB pathway, the most common deletion was in the CDKN2A/CDKN2B locus on chromosome 9q21 and amplification of CDK4.
One of the most important biomarkers in glioma appears to be the methylation status of the promoter region of the MGMT gene, which predicts sensitivity to temozolomide, the standard of care treatment for the disease, and has implications for treatment regimens in the future management of the disease (Brandes *et al*., 2006).
Mutations in the NF1 gene, conventionally associated with neurofibromatosis, have also been described. Loss of NF1 is expected to lead to upregulation of the Ras/MAPK pathway, and could sensitize these tumors to therapies based on Raf or MEK inhibitors.
Another genome-wide study of glioma showed unexpected recurrent mutations in the active site of isocitrate dehydrogenase1 (IDH1), an enzyme involved in central glucose metabolism (Parsons *et al*., 2008). These studies have been confirmed and extended to include the IDH2 gene, suggesting the possibility that modulation of IDH activity may provide some therapeutic opportunity for those patients in whom these enzymes are inactivated (Yan *et al*., 2009).

Melanoma

Germline mutations in p16 and CDK4 were discovered several years ago as major familial melanoma predisposing genes, and a third familial gene (p14 ARF) is derived from an alternative reading frame of the CDKN2A gene locus. Some progress has been made in finding lower penetrance melanoma genes in families, but very large GWAS need to be done to find loci conferring lower susceptibility. Some gene expression signature studies have been done in melanoma, one of the most important of which has led to the derivation of a test to measure the expression (as proteins by immunohistochemistry) of five genes that can distinguish malignant melanoma from benign naevi (Kashani-Sabet *et al*., 2009). Other studies are being directed to correlation of gene expression signature with progression and outcome.
More informative has been the analysis of somatic mutations. One example is an activating mutation of BRAF (V600E) found in the majority of all melanomas. A novel therapeutic agent (PLX 4720) has been developed that specifically targets this mutant form of BRAF, and the compound is in clinical trials (Tsai *et al*., 2008). Other cancers such as thyroid and lung cancer also harbor BRAF mutations. The BRAF gene is part of the mitogenic pathway downstream of KRAS (Fig 1) and activating mutations in BRAF could also interfere with drugs that target mechanisms upstream in the pathway, suggesting that PLX4720 may be useful for cancers other than melanoma. Since the drug specifically targets the mutant form of BRAF, it will be important to have a companion diagnostic test that can distinguish the BRAF (V600E) from the wild-type gene. Analysis of somatic mutations in tyrosine kinase genes in cutaneous melanoma recently identified

recurrent mutations in ERBB4 which could lead to therapeutics directed specifically towards patients with tumours with mutations in this protein (Prickett *et al.,* 2009). Some 5–10% of melanomas appear to have mutations in cKit and respond dramatically to imatinib treatment. Distinguishing these patients from those with BRAF mutations (who do not respond to imatinib) is clearly of considerable importance clinically (Curtin *et al*., 2006).

Solid pediatric tumors

The small, round, blue cell tumors (SRBCTs), including neuroblastoma (NB), rhabdomyosarcoma (RMS), non-Hodgkin's lymphoma (NHL), and Ewing sarcoma (ES), represent four of the most aggressive solid cancers in the pediatric population and accurate diagnosis is critical to the management of these patients. For example, patients with high-stage RMS require a combination of high-dose chemotherapy, surgery, and radiation treatment, whereas patients with non-Hodgkin's lymphoma (NHL) require intrathecal chemotherapy.

These cancers are difficult to distinguish by light microscopy, and currently no single test can precisely distinguish these cancers. To confirm the diagnosis, pathologists rely on several techniques, including immuno-histochemistry, cytogenetics, FISH and RT-PCR. Molecular techniques to measure tumor-specific chromosome translocations are used to diagnose EWS containing the EWS-FLI1 translocation and alveolar rhabdomyosarcoma (ARMS) containing the PAX3-FKHR translocation.

Using microarray gene expression profiling and artificial neural networks (ANNs), 93 genes have been identified as a gene expression signature which was capable of stratifying these SRBCTs to specific diagnostic categories (Khan *et al*., 2001). Further work has demonstrated that a measurement of the expression of a subset of the genes were sufficient for stratification, making an RT-PCR-based test a possibility. The Khan lab has also shown recently that up-regulation of and mutations in the FGFR4 gene appear to promote metastasis in RMS patients (Taylor *et al*., in press).

Neuroblastoma

Neuroblastoma (NB) is a tumor derived from the peripheral sympathetic nervous system and is the most common solid extracranial tumor of childhood. Patients are currently stratified into high, intermediate, and low risk, based on age, tumor staging, histology, MYCN amplification, and DNA ploidy. Patients <18 months of age or with lower-stage diseases (1 or 2) have a better outcome (>80% survival) than older patients or those with advanced stage disease. The MYCN gene, located on chromosome 2p24, is amplified in ~22% of all NB tumors and is an independent predictor for poor prognosis. Although this risk stratification accurately predicts that >80% patients with low risk and <30% of patients with high-risk disease will survive with modern therapy, it cannot predict the outcome of individual patients with NB. Microarray studies have demonstrated for the first time that NB tumors have a diagnostic-specific gene expression profile, compared with normal tissues or other malignancies. It has also been shown that the gene expression profile of low-stage tumors is markedly different from those derived from high-stage tumors (Krasnoselsky *et al*., 2005; Wei *et al*., 2004; Ohira *et al*., 2005; Schramm *et al*., 2005). Furthermore, 19 genes, including MYCN and CD44, had an

expression profile that had a positive predictive value of >80% for poor outcome for patients with NB of all stages. The majority of the 19 genes are involved in early neural development, suggesting that the more aggressive phenotype is characterized by less differentiated cells. Other gene expression studies have also correlated expression of gene sets with outcome for NB (Asgharzadeh *et al.*, 2006, Oberthuer *et al.,* 2006) (Table 2).

In recent surveys of germline and somatic mutations in neuroblastoma high-level amplification (>10 copies) and mutation of the ALK (anaplastic lymphoma kinase) gene have been observed especially in advanced cases (Chen *et al*., 2008; George *et al*., 2008; Janoueix-Lerosy *et al*., 2008; Mosse *et al*., 2008) This is interesting in the context that ALK has been shown to be deregulated in other cancers, particularly by translocation events for example, EML4-ALK in non-small-cell lung cancer and NPM ALK in non-Hodgkin's lymphoma (Chiarle *et al*., 2008) suggesting that TK inhibitors that target ALK may be useful in the clinical management of neuroblastomas with activated ALK.

Non-Hodgkin's Lymphoma

Non-Hodgkin's lymphomas account for about 5% of all cancers in the U.S.A., and the incidence is rising in an ageing population. NHLs are classified as either T cell or B cell in origin and can be indolent or aggressive in character.
Seminal studies by the Staudt group at the NCI, working with the Leukemia Lymphoma Molecular Profiling Program (LLMPP) over the last 10 years, has demonstrated the feasibility of using gene expression microarray technology for identification and classification of a variety of lymphoma sub-classes. Burkitt's lymphoma is an aggressive form of non-Hodgkin's lymphoma associated with a translocation of the MYC gene [t(8;14)] that can be difficult to diagnose using histopathology and sometimes can be confused with Diffuse Large B-Cell Lymphoma (DLBCL). Successful treatment of Burkitt's lymphoma requires a much more aggressive chemotherapy regimen than DLBCL. Accurate diagnosis and clear differentiation of Burkitt's lymphoma from DLBCL is critical to patient outcome. Several studies have demonstrated that gene expression arrays can be used to identify Burkitt's lymphoma in patients with disease diagnosed as DLBCL by histopathology (Hummel *et al*., 2006). In a separate study, Dave *et al*. (2006) identified a Burkitt's lymphoma signature that showed decreased expression of genes in the NF-kB pathway and increased expression of genes that were the targets of cMyc. A subset of high-grade Burkitt's lymphoma patients who had been treated with CHOP (cyclophosphamide, adriamycin, vincristine, and prednisolone - the standard of care for DLBCL) were identified by this means and shown to have inferior survival to similar patients treated with more aggressive therapy [etoposide, doxorubicin, vincristine, prednisolone, cyclophosphamide (EPOCH) plus rituximab, an anti CD20 antibody]. A diagnostic test based on these gene expression differences has the opportunity to significantly prolong the lives of those Burkitt's lymphoma patients who are difficult to distinguish from DLBCL.
DLBCL has been considered to be a single, albeit heterogeneous disease and is generally treated with the same therapy (CHOP) sometimes given in conjunction with rituximab. Novel molecular gene expression signatures have now also identified distinct biological subtypes of DLBCL specifically, Germinal Center B Cell-like (GCB), Activated B Cell-

like (ABC), and primary mediastinal B cell lymphoma (PMBL). These forms of lymphoma differ in both prognosis and potential therapy. The t(14;18) translocation involving BCL2 and amplification of the c-REL locus on chromosome 2p is seen exclusively in GCB tumors. In contrast, the NF-κB pathway is activated in ABC cells but not GCB cells. The GCB subtype responds well to traditional therapy (CHOP) and CHOP plus rituxumab (R CHOP), whereas the ABC subtype has a generally poor prognosis and needs additional therapeutic approaches. New drugs targeting NF-κB are being tested in this poor prognosis group. The proteasome inhibitor bortezomib, which inhibits the NF-κB pathway, has been tested in conjunction with chemotherapy of relapsed and refractory DLBCL in a small Phase II clinical trial conducted at the NIH clinical center (Dunleavy *et al*., 2009). In addition to these subtypes, the microenvironment of the tumor cells has been shown to be important and gene expression signatures have recently defined two distinct stromal subtypes (Lenz *et al*., 2008a). The stromal signatures correlate with survival after R-CHOP treatment.

Array technology and DNA sequencing have also been used to identify gene copy number changes and mutations in DLBCL. Recurrent oncogenic mutations affecting the NF-κB pathway have been identified by the LLMPP, in particular in CARD11, a gene coding for a cytoplasmic scaffold protein involved in NFκB activation pointing to the coiled-coil domain of CARD11 as a potential therapeutic target (Lenz *et al*. 2008b). Other studies have shown that there is frequent inactivation of A20 (TNFAIP3), a negative regulator of NF-κB in B cell lymphomas, presumably of the ABC type (Compagno *et al*., 2009; Kato *et al.*, 2009). All these studies demonstrate that activation of NF-κB is caused by genetic lesions in several genes, the loss or activation of which promote cellular proliferation by activating NF-κB. This strongly suggests using drugs directed at this transcription factor for DLBCL (Baud and Karin 2009), and it also may be appropriate to test the new NEDD8-activating enzyme inhibitors in molecularly profiled (ABC) patients with DLBCL (Soucy *et al*., 2009).

Multiple Myeloma

Multiple myeloma is a malignancy of plasma cells with a very variable outcome following standard or high-dose treatment. For some time, myelomas have been categorized molecularly by chromosome analysis. Myelomas can be segregated into non-hyperploid or hyperploid in character. In the latter, there are often trisomies in chromosomes 3, 5, 11,15, 19, and 21, which can be subdivided by array CGH into tumors that have the trisomies above (except 11) and have additional chromosomal alterations, including gains of chromosome 1q and chromosome 7 and deletion of chromosome 13. Non-hyperploid disease can be subdivided into two groups: one with deletions of chromosomes 8 and 13 and one with amplification of chromosome 1q and deletion of 1p and 13. Immunoglobulin heavy-chain region translocations (14q32) involving several genes is also apparent in up to 40% of myeloma patients.

The myeloma group in Arkansas has recently further extended the classification of myeloma using gene expression arrays (Shaughnessy *et al*., 2007). In their study, two high-risk profiles were obtained, one based on the chromosome 14q32 translocations and the other on the expression of genes expressed during proliferation on the background of

the other chromosome abnormalities classification. This builds on data pointing to the importance of chromosome 1 abnormalities, including the fact that gains of 1q21 lead to inferior survival. The objective of the original study was to identify patients with high-risk disease. This approach may provide a more powerful and cheaper alternative to the extensive chromosomal analysis routinely used to subdivide patients and to identify those at higher risk for early fatal disease.
Gene expression profiling in myeloma has since been extended to be an important component of the clinical monitoring of patients treated with bortezomib an inhibitor of NF-κB activation. The gene expression stratification for newly diagnosed multiple myeloma treated with high-dose chemotherapy is predictive of outcome in relapsed disease treated with bortezomib or high-dose steroids (Zhan *et al*., 2008, Mulligan *et al*., 2009). Gene expression is also being used to look at patients with respect to proliferation free survival time after relapse (Nair *et al*., 2009). More information on the somatic mutations found in myeloma cells will be forthcoming from the whole genome sequencing studies sponsored by the Multiple Myeloma Research Consortium. This will point to other drugs that can be used in subsets of patients, based on the genetic lesions in their tumor cells.

Leukemias

Routine histopathology and cytogenetic analysis are mainly used for leukemia diagnosis. However, molecular techniques are increasingly being used to refine the diagnosis and subdivide the patients (Staudt 2006).

Chronic Myelogenous Leukemia (CML)
CML is held up as a poster child for personalized genetic medicine. A known translocation {The Philadelphia chromosome t(9;22)(q34;q11)} gives rise to one of three fusion proteins (Bcr-Abl) that activate the Abl kinase gene, resulting in uncontrolled cell proliferation. The Abl kinase is a target for a selective tyrosine kinase inhibitor (Imatinib mesylate, Glivec). This drug was being developed by CIBA-Geigy and found to have the ability to inhibit Abl kinase. This led to the trial of the compound in CML patients with good efficacy and to rapid approval for the treatment of CML (Lydon 2009). Owing to the ability of the compound to inhibit both cKit and PDGF receptor TK in addition to Abl, imatinib is also being used to treat GIST and other cancers where these enzymes are activated or overexpressed. In the chronic phase of CML, imatinib works by preventing Bcr-Abl-positive stem cells producing progenitor cells that differentiate down the myeloid pathway, but the compound does not inhibit fully differentiated cells. In blast crisis, where the progenitor cells are replicating in an uncontrolled fashion, imatinib kills the majority of the cells, but owing to the number and speed of proliferation, cells with drug resistance inevitably emerge. Patients who develop resistance to imatinib are being treated with other more recently developed inhibitors dasatinib and nolotinib (Shah *et al*., 2007). One notable resistant mutant (T315I) is not inhibited effectively by any of these compounds. Sequencing of the Bcr-Abl gene in cells from patients with imatinib-resistant disease is an important part of CML patient management today.
Despite being an excellent example of personalized medicine at work, CML and the development of imatinib are unusual in several respects. Firstly, cell proliferation is

driven by changes in a single gene in this disease; and secondly, imatinib was being developed as a PDGFR TK inhibitor, rather than targeted to the Abl kinase initially. Nevertheless, the paradigm is very important. Imatinib has not only fundamentally changed the life expectancy of patients with GIST and CML, but it has also fully justified the decision of Novartis to develop an agent for a disease with an incidence of only some 5,000 new patients a year. It should also provide a salutary message to the Pharmaceutical and Biotechnology industry that targeted therapies can be successful.

Acute Myeloid Leukemia

AML patients are generally stratified into good, intermediate, or poor outcome categories, depending on the cytogenetic abnormalities in their cells. AML is a heterogeneous disease with several subtypes, including those based on chromosome translocations (e.g. (t(8;21)[RUNX1(AML1)-RUNX1T1(ETO)] associated with good outcome, inv(16)/t(16;16)[CBF beta-MYH11]), and cases with MLL rearrangements on chromosome 11q23 associated with a poorer prognosis. Some point mutations have been found in genes such as KRAS, FLT 3, cKIT, and TP53. Mutations in the juxta-membrane domain through internal tandem duplications (ITDs) and in the kinase domain of FLT 3 are common in AML. FLT 3 ITDs are associated with increased relapse rates and shorter survival. These studies have implications for clinical trials involving kinase inhibitors targeting FLT3, where it would be appropriate for all activating mutations of the gene to be measured as part of the entry criteria into the clinical trial (Frohling *et al*., 2007).

High-throughput digital gene analysis using the Nanostring technology with acute promyelocytic leukemia with the t(15;17) RAR alpha gene fusion has been done with limited amounts of clinical material and showed a profile specific to samples with the PML-RARA fusion. This indicates that a gene expression array can be used instead of or as well as chromosome (FISH) analysis to differentiate AML subtypes (Payton *et al*., 2009). A two-gene classifier has also been developed in AML for predicting the response to the farnesyl transferase inhibitor tipifarnib, an inhibitor with rather poor and variable efficacy (Raponi *et al*., 2008) (Table 2).

Further genomic analysis of 111 children with AML has revealed that in contrast to other cancers there are in fact limited somatic changes in the genome of these cancer cells (Radtke *et al*., 2009). Although the analysis was based on snps and copy number variation rather than on sequencing, this study has confirmed the published data derived by sequencing the genome of an individual with AML (Ley *et al*., 2008) and both studies showed low level changes in genes in pathways already implicated in this cancer (e.g. FLT3 and TP53). In a follow-up study with an another patient with AML, additional mutations were found, some of which were in the same genes as those found in the first patient, and are therefore unlikely to occur by chance. Of potential importance was the finding of mutations in the IDH1 gene, although these were not the same mutations as found in this gene in the glioma studies (Mardis *et al*., 2009).

Acute Lymphoblastic Leukemia

The cure rate for ALL using conventional chemotherapy is more than 80%, but for those unfortunate children who do relapse, the prognosis is grim. Gene copy number analysis has been used to compare diagnosis and relapse samples in an attempt to dissect those

genes involved in relapse and treatment response. The evidence so far points to changes in the copy number of genes involved in cell cycle regulation and B cell development (Mullighan *et al*., 2007; Yang *et al*., 2009). More than 50 copy number abnormalities were found in a cohort of 221 children with high risk B-cell progenitor ALL. Alteration of the IKZF1 gene, which encodes a lymphoid tissue transcription factor (IKAROS), was found to be a predictor of poor outcome. The gene expression signature of the poor outcome patients showed increased expression of haematopoietic stem cell genes and reduced expression of genes of the B cell lineage. These signatures were similar to those obtained for Bcr-Abl positive ALL, a high-risk group with a high frequency of IKZF1 deletions (Mullighan *et al*., 2009). The extension of these studies should make it possible to identify those children who are more likely to relapse and identify the most effective treatment modalities. A recent GWAS in ALL has also shown association of the IKZ gene (chr.7p 12.2), with susceptibility to disease, and also reported two other regions of interest, both of which involve transcriptional regulation of the B cell progenitor lineage (Papaemmanuil *et al*., 2009).
About one-third of patients with myelodysplastic syndromes and myeloproliferative disorders progress to AML. Myeloproliferative disorders such as polycythemia vera frequently present with a somatic activating mutation in the Janus kinase, JAK 2 (V617F) (Morgan and Gilliland, 2008). Mutations in a candidate tumor suppressor gene (TET2) have recently been described in about 15% of patients with various myeloid cancers. These are thought to arise before the JAK2 mutations (Delhommeau *et al*., 2009).

Chronic Lymphoblastic leukemia (CLL)
CLL is the most common human leukemia. It is generally indolent in character but not curable. The expression of non mutated versions of the immunoglobulin heavy chain variable region locus (IgHV) and leukemic cell expression of CD38 is associated with more aggressive forms of the disease. The expression of ZAP 70, an intracellular tyrosine kinase associated with T-cell signaling, is a marker of unmutated IgHV and thus a biomarker for less indolent disease (Wiestner *et al.*, 2003). Patients with ZAP 70-positive CLL cells should be monitored closely, as they require chemotherapy before the more indolent forms of this disease (Rassenti *et al*., 2008).

Implications and Conclusions

The implications from this analysis are quite profound. The understanding of the underlying germline and somatic mutations involved in cancer (*Supplementary materials Tables 3 and 4*) is already allowing better diagnosis and treatment. That is not to say that this understanding will allow cure of the disease; however, the ability to diagnose and treat more effectively, using combinations of drugs directed to the molecular defects that can be seen in the particular tumors, will change the clinical management of cancer into that of managing a more chronic disease state. Tables 3 and 4 show that there are multiple germline and somatic gene alterations in cancer, knowledge of which can direct therapeutic options. Many of the larger cancer centers (MSKCC, MD Anderson, MGH and Dana Farber) are already triaging cancer patients of different types for the routine analysis of biopsy slides for mutations in many of the genes. Some of these are done by

reimbursable tests and some in the course of research programs. There is also a compelling case to be made for gene-directed clinical trials in NSCLC by segregating the KRAS mutant, EGFR mutant, and EML4-ALK tumors and also in breast cancer, using PI3K inhibitors in cancers that have failed trastuzumab and have PI3K mutations or ERBB2 amplifications. Nevertheless, the combination of the mutations and the selectivity of the inhibitors are going to be complex even for compounds targeting one pathway (see Engelman 2009; Liu *et al*., 2009).

Several issues, if not addressed, will either slow down or even prevent the ability of the general primary care cancer physician to deliver this genetic information to the patient. One of the main barriers is making sure that the biomarkers that are found are relevant for disease progression monitoring and treatment. As shown above, considerable progress is being made but the complexity of the data is very high. The general physician will need guidance not only as to which tests to order to examine tumors for their mutations, but also to describe what the consequences of the results are with respect to treatment and follow-up. This is a great deal more complicated but analogous to blood results that are requested and acted on in cardiovascular and other diseases today. The bioinformatics needed to provide digested and useful information to the clinician to guide his decision-making is not trivial. Educating physicians to appreciate molecular profiling will be important, because it will be a clear economic driver for health in the future. As much effort needs to go into planning how to provide the "molecular information content" to electronic health records as into planning the development and delivery of the electronic health records themselves (Jha *et al*., 2009; Steinbrook 2009).

The speed with which the information can be derived from patient presentation to biopsy to result will be important and will increase the pathologist's role in "active diagnosis." There will be a need for a larger number of "interventional" molecular pathologists. For example, the finding of EGFR mutations in NSCLC is important for the lung cancer patients in real time, since EGFR inhibitors do work in these patients at least until resistance emerges. Further interesting examples of biomarker-directed treatment of solid tumors is described by Sawyers (2008). All this serves to reinforce the view that there needs to be much more of a partnership between the molecular pathologist and the physician in the future.

There is an increasing interest commercially in providing "relative genetic risk" information to individuals from GWAS. Several companies, including Navigenics, DeCodeMe, and 23andme, have been set up to provide this information in the context of "life style management and wellness" programs. This "genetic risk information" is designed to enable people to do something prospectively to ameliorate their genetic predisposition. Unfortunately, given the influence of environment on the outcome of their predisposition, this is not going to be (by definition) very accurate information (Ng *et al.,* 2009a). As elegantly stated by Bruce Ponder: "the clinical use of single common low penetrance genes is limited but a few susceptibility alleles may distinguish women who are at high risk for breast cancer from those at low risk, particularly in the context of population screening" (Pharoah *et al*., 2008). The Scripps Clinic genomics group has recently announced a 3500-person screening initiative to look at the common snps predisposing to breast cancer in women who have had at least five years of

mammography. This view of the utility of low-risk genetic information is shared by others: it is a fact that it is gene function in the context of the environment that confers outcome (Vineis *et al*., 2001; Kraft and Hunter, 2009). Even highly penetrant gene mutations (such as the δ508 mutation in the CFTR gene) give rise to widely different phenotypes, reflected in the severity of the resultant cystic fibrosis disease, little of which is accounted for by other genetic effects. Humans are in general notoriously incapable of dealing with the concept of relative risk: whether relative genetic risk will be taken any more seriously by the majority of an increasingly socially networked population remains to be seen. The costs of the delivery and the utility of this kind of genetic information, however (about $2,500-$4,000), need to be put into the proper perspective, given that the cost of a BRCA1 test is about the same and yields much more valuable information.

Somatic mutation analysis of tumors can make a dramatic difference to the economic fate of the patient. The KRAS mutation test is a $400 test for an $80,000 decision (the cost of an EGFR Mab therapy course). It has been estimated that KRAS tests will save $20 million for every 1,000 metastatic colon cancer patients progressing to second-line therapy directed at EGF receptor inhibition, saving approximately $600 million per year. Using the Oncotype DX test to avoid unnecessary chemotherapy could also save about $2000 per patient based on a 34% reduction in the use of chemotherapy.

Alignment of the stakeholders

Several stakeholders (payers, physician providers, pharmaceutical and biotechnology companies, and molecular diagnostics companies) need to be aligned for the effective delivery of molecular diagnostics in the context of therapeutics; i.e., the Dx/Rx model. Although there are several challenges, including the complexity of the data and the operational delivery of the tests, by far, the most important challenge is to align the stakeholders economically (Aspinall and Hamermesh, 2008). Issues include making sure that there are incentives to the payers to pay for diagnostic tests that may have longer-term consequences that are no longer that particular payer's concern. Additionally, reimbursement practice must be in line with the utility of the drug for a sub set of patients (e.g. no reimbursement for EGF receptor-targeted drugs unless accompanied by a KRAS wild-type designation for the patient). The intellectual property issues could also be considerable when, for accurate patient profiling for optimal drug treatment and follow-up, the changes in several genes (panel profiling) need to be measured. A good case in point would be lung cancer, where KRAS, EGFR (activating and resistant mutations), EML-ALK, cMet amplifications and PTEN deletions will need to be analyzed in the original tumor and in any metastatic lesions.

A report by McKinsey recently elaborated four major catalysts to help the adoption of a more personalized approach to medicine (Davis *et al*., 2009). These are maximizing the transparency/efficiency of the regulatory process (510K, etc.); increasing the pace and predictability of payer coverage for appropriate tests; aligning reimbursement practices to encourage appropriate diagnostic use by physicians; and encouraging pharmaceutical and biotechnology companies to take a long-term investment view. A report to the President by the President's council of advisors on science and technology has echoed these sentiments and provided further clear recommendations about what can be done by the government to promote the rapid adoption of personalized medicine (*Priorities for Personalized Medicine, 2008).*

Unfortunately, given the short-term pressures on maintenance of shareholder value quarter by quarter for both large biotech and pharmaceutical companies, encouraging them to take a long-term view without incentives/disincentives is not going to work. The pharmaceutical/biotech industry has to come to terms with the fact that the patient population responding well to their drug might be <20% of the overall patients with that particular disease, thus reducing their potential market by 80%. It is appropriate in today's environment that the very expensive cancer drug treatments (which sometimes only increase survival on average by a few weeks) are used selectively in patients whose molecular profile would indicate efficacy. Industry or reimbursement practices must make this happen. The pricing of small-molecule TK inhibitors based on the price of a drug efficacious in the majority of patients (Imatinib in CML) is completely unjustified, unless the drugs have the same efficacy profile in the overall patient populations in which they are used. Most of them do not.

Perhaps a better model to consider would be an "own the patient" model not unlike what happens now for a patient with diabetes. As soon as a potential cancer patient walks into the doctor's office (as above), the pharmaceutical company pays for all the profiling, in exchange for a fixed patient fee for managing all the tests, drugs, and procedures necessary to manage that patient's disease. The data on average costs from diagnosis to death for most cancers is known. The anticipation is that these costs would then be reimbursable by the insurance companies or paid for by Medicare. The fact that patients treated with a targeted therapy would be expected to live longer and therefore need the drugs for longer may make the economics rather better than expected for the Pharmaceutical companies. The use of targeted EGFR inhibitors in profiled stage IV lung cancer patients with EGFR mutations for example, can give striking increases in survival from the 18-week average to 2 to 3 years. From an economic perspective, the experience and success of Gleevec as a targeted therapeutic definitively supports this view.
The "own the patient" scenario will require some co-operation between the pharmaceutical companies with oncology franchises, since no one of them will have all the targeted drugs needed in their own stable. This system would also place a premium on accurate diagnosis and prognosis, and the current molecular pathological profiling that can be done is fully consistent with this approach.

In conclusion, it is most likely that cancer will be the disease where the full adoption of personalized medicine will happen first, owing to the molecular information that is becoming available. Organizations like the Veterans Administration, Kaiser Permanente, and the Ascension Hospitals, who tend to take a longer-term view of their patient welfare and keep patients in their system for longer, should help to drive the process of adopting more personalized medicine by advocating the comprehensive use of molecular diagnostic tests in cancer.

References

Ahmed S et al. (2009). Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet* **41**: 585-590.

Amos CI et al. (2008). Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* **40**: 616-622.

Alizadeh AA et al. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**(6769): 503-511.

Asgharzadeh S et al. (2006). Prognostic significance of gene expression profiles of metastatic neuroblastomas lacking MYCN gene amplification. *J Natl Cancer Inst*, **98**(17): 1193-12203.

Aspinall M, Hamermesh RG. (2007). Realizing the promise of personalised medicine. *Harvard Bus Rev* R0710F.

Altshuler D, Daly M, Lander ES. (2008). Genetic mapping in human disease. *Science* **322**: 881-888.

Bast R, Hennessey B, Mills GB. (2009). The biology of ovarian cancer: new opportunities for translation. *NatRev Cancer* **9**: 415-428.

Baud V, Karin M. (2009). Is NF-κB a good target for cancer therapy? Hopes and pitfalls. *Nat Rev Drug Discov* **8**: 33-39.

Bean J et al., (2007). Met amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. Proc Nat Acad Sci USA **104**: 20932-20937.

Blow N. (2009). Genomics: catch me if you can. *Nat Meth* **6**: 539-542.

Boyault S et al. (2007). Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* **45**: 42-52.

Brandes AA et al. (2006). Correlations between 06-methylguanine DNA methyl transferase promoter methylation status, 1p and 19q deletions, and response to temozolomide in anaplastic and recurrent oligodendroglioma: a prospective GICNO study. *J Clin Oncol* **24**: 4746-4752.

Campbell P et al. (2008). Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet* **40**: 722-729.

Chang JT et al. (2009). A genomic strategy to elucidate modules of oncogenic pathway signaling networks. *Molec Cell* **34**: 104-114.

Chen Y et al. (2008). Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* **455**(7215): 971-974.

Chen H-Y et al. (2007). A five-gene signature and clinical outcome in non small cell lung cancer. *N Engl J Med* **356**: 11-20.

Chiang DY et al. (2008). Focal gains of *VEGFA* and molecular classification of hepatocellular carcinoma. *Cancer Res.* **68**: 6779-6788.

Chiarle R et al. (2008). The anaplastic lymphoma kinase in the pathogenesis of cancer. *Nat RevCancer* **8**: 11-23.

Choi YL et al. (2008). Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer. *Cancer Res* **68**: 4971-4976.

Compagno M. (2009). Mutations of multiple genes cause deregulation of NF-kB in diffuse large B cell lymphoma. *Nature* **459**: 717-722.

Couzin F. (2008). MicroRNAs make big impression in disease after disease. *Science* **319**: 1782-1783.

Crinjs APG et al. (2009). Survival-related profile, pathways and transcription factors in ovarian cancer. *PLoS Medicine* **6**: 1-13.

Curtin JA et al. (2006). Somatic activation of Kit in distinct subtypes of melanoma. *J Clin Oncol* **24**: 4340-4346.

Dalla Palma M et al. (2008). The relative contribution of point mutations and genomic rearrangments in BRCA1 and BRCA2 in high-risk breast cancer families. *Cancer Res* **68**: 7006-7014.

Dall'Era MA et al. (2008). Active surveillance for early stage prostate cancer: review of current literature. *Cancer* **112**(8) 1650-1659.

Dave SS et al. (2006). Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med* **354**(23): 2431-2442.

Davis JC et al. (2009). The microeconomics of personalized medicine: today's challenge and tomorrow's promise. *Nat Rev Drug Discov* **8**: 279-286.

Delhommeau F et al. (2009). Mutation in TET2 in myeloid cancers. *N Engl J Med* **360**: 2289-2301.

Ding L et al. (2008). Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**: 1069-1074.

Downing J. (2009). Cancer Genomes – Continuing Progress. *N Engl J Med* **360**: 1111-1112

Dunleavy K. et al. (2009). Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse B cell lymphoma. *Blood* **113**: 6069-6076.

Easton D. (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**:1087-1093.

Eeles RA et al. (2008). Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet* **40**: 316-321.

Eeles RA et al. (2009). Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet* **41**: doi:10.1038/ng.450

Engelman JA et al. (2007). *MET* amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**: 1039-1043.

Engelman JA. Janne PA. (2008). Mechanisms of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small cell lung cancer. *Clin Cancer Res* **14**: 2895-2899.

Engelman JA. (2009). Targeting PI3K signaling in cancer: opportunities, challenges and limitations. *Nat Rev Cancer* **9**: 550-562.

Farmer H et al. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**: 917-921.

Fong PC. et al. (2009). Inhibition of Poly(ADP-Ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med* **361**:123-134.

Foulkes WD. (2008). Inherited susceptibility to common cancers. *N Engl J Med* **359**: 2143-2153.

Freie BW, Eisenman RN. (2008). Ratcheting Myc. *Cancer Cell* **14**: 425-426.

Frohling S et al. (2007). Identification of driver and passenger mutations of Flt3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. *Cancer Cell* **12**: 501-513.

Garman KS et al. (2008). A genomic approach to colon cancer risk stratification yields biological insights into therapeutic opportunities. *Proc Natl Acad Sci USA* **105**: 19431-19436.

Geiss GK et al. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotech* **26**: 317-325.

George RE et al. (2008). Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature* **455**(7215): 975-978.

Goldstein D. (2009). Common genetic variation and human traits. *N Engl J Med* **360**: 1696-1698.

Greenman C. et al. (2007). Patterns of somatic mutation in human cancer genomes. *Nature* **446**: 153-158.

Guibellino A, Linehan WM, Bottaro DP. (2009). Targeting the Met signaling pathways in renal cancer. *Expert Rev Anticancer Ther* **9**: 785-793.

Herbst RS, Heymach JV, Lippman SM. (2008). Lung cancer *N Engl J Med* **359**: 1367-1380.

Hirschorn JN. (2009). Genome-wide association studies—Illuminating biologic pathways. *N Engl J Med* **360**: 1699-1701.

Hoshida Y et al. (2008). Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* **359**: 1995-2004.

Hoskins JM, Carey L, McLeod HL. (2009). *CYP2D6* and tamoxifen: DNA matters in breast cancer. *Nat Rev Cancer* **9**: 576-586.

Hung RJ et al. (2008). A susceptibility locus for lung cancer maps to nicotinic acetyl choline receptor subunit genes on 15q25. *Nature* **452**: 633-637.

Hummel M et al. (2006). A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med* **354**: 2419-2430.

Hunter D et al. (2007). A genome-wide association study identifies alleles in FGFR2 associated with risk of spoaradic postmenopausal breast cancer. *Nat Genet* **39**: 870-874.

Inglehart JD, Silver DP. (2009). Synthetic lethality—a new direction in cancer drug development. *N Engl J Med* **361**: 189-191.

Innocenti F et al. (2009). Comprehensive pharmacogenetic analysis of irinotecan, neutropenia, and pharmacokinetics. *J Clin Oncol* **27**: 2604-2614.

Janoueix-Lerosy et al. (2008). Somatic and germline mutations of the ALK kinase receptor in neuroblastoma. *Nature* **455**: 967-971.

Jha A et al. (2009). Use of electronic health records in U.S. hospitals. *N Engl J Med* **360**: 1628-1638.

Jia L et al. (2009). Functional enhamcers at the gene-poor 8q24 cancer-linked locus. *PLOS Genetics* **5** e1000597

Jones S et al. (2008a). Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**: 1801-1806.

Jones S et al. (2008b). Comparative lesion sequencing provides insights into tumor evolution. *Proc Natl Acad Sci USA* **105**: 4283-4288.

Jones S et al. (2009). Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* **324**: 217.

Karapetis CS et al. (2008). K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med* **359**: 1757-1765.

Kato M et al. (2009). Frequent inactivation of A20 in B cell lymphomas. *Nature* **459**: 712-716.

Kashani-Sabet M et al. (2009). A multimarker assay to distinguish malignant melanoma from benign nevi. *Proc Natl Acad Sci USA* **106**: 6268-6272.

Kim JB et al. (2007). Polony multiplex analysis of gene expression (PMAGE) in mouse hypertrophic cardiomyopathy. *Science* **316**: 1481-1484.

Kim J-I et al. (2009). A highly annotated whole genome sequence of a Korean individual. *Nature* **460**: 1011-1016.

Khan J, Wei JS et al. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* **7**(6): 673-679.

Koivunen JP et al. (2008). Mutations in the LKB1 tumour suppressor are frequently found in tumours from Caucasian but not Asian lung cancer patients. *Br J Cancer* **99**: 242-252.

Krasnoselsky AL, et al. (2005) Altered expression of cell cycle genes distinguishes aggressive neuroblastoma. *Oncogene* **24**(9): 1533-1541.

Kraft P, Hunter D. (2009). Genetic risk prediction—are we there yet? *N Engl J Med* **360**: 1701-1702.

Kumar-Sinha C, Tomlins SA, Chinnaiyan AM. (2008). Recurrent gene fusions in prostate cancer. *Nat Rev Cancer* **8**: 497-511.

Kummar S et al. (2009). Phase 0 clinical trial of the Poly (ADP-Ribose) polymerase inhibitor ABT-888 in patients with advanced malignancies. *J Clin Oncol* **2**: 2705-2711.

Lam LT et al. (2008). Compensatory IκKα activation of classical NF-κB signaling during IκKB inhibition identified by an RNA interference sensitization screen. *Proc Natl Acad Sci USA* **105**: 20798-20803.

Lebanony D et al. (2009). Diagnostic assay based on hsa-miR-205 expression distinguishes non-squamous from non-small-cell lung cancer. *J Clin Oncol* **27**: 2030-2037.

Lenz G et al. (2008a). Stromal gene signatures in large B cell lymphomas. *N Engl J Med* **359**: 2313-2323.

Lenz G et al. (2008b). Oncogenic CARD 11 mutations in Human Diffuse Large B Cell Lymphoma. *Science* **319**: 1676-1679.

Ley T et al. (2008). DNA sequencing of a cytogenetically normal acute myeloid leukemia genome. *Nature* **456**: 66-72.

Liu P et al. (2009). Targeting the phosphoinositide 3-kinase pathway in cancer. *Nat Rev Drug Disc* **8**: 627-643.

Llovet J, Bruix J. (2008). Molecularly targeted therapies in hepatocellular carcinoma. *Hepatology* **48**: 1312-1327.

Llovet J et al. (2008). Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* **359**: 378-390.

Lou H et al. (2009) Fine mapping and functional analysis of a common variant in MSMB on chromosome 10q11.2 associated with prostate cancer susceptibility. *Proc Natl Acad Sci USA* **106**: 7933-7938.

Luo Ji. (2009). A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the RAS oncogene. *Cell* **137**: 835-848.

Lydon,N (2009). Attacking cancer at its foundation. *Nature Medicine* **15**: 1153-1157
Maher C et al. (2009). Transcriptome sequencing to detect gene fusions in cancer. *Nature* **458**: 97-103.

Maheswaran S et al. (2008). Detection of mutations in *EGFR* in circulating lung-cancer cells. *N Engl J Med* **359**: 366-377.

Manolio TA, Brooks LD, Collins FS. (2008). A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* **118**: 1590-1605.

Manolio TA, et al. (2009). Finding the missing heritability of complex diseases. *Nature* **461**: 747-753.

Mardis ER. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet* **24**: 133-141.

Mardis ER et al. (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* **10**:1058-1066.

Mehra R et al. (2008). Chracterisation of TMPRSS2-ETS gene aberrations in androgen-independent metastatic prostate cancer. *Cancer Res* **68**: 3584-3590.

Mellinghoff IK. (2005). Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N Engl J Med* **353**: 2012-2024.

Merritt WM et al. (2008). Dicer, Drosha and outcomes in patients with ovarian cancer. *N Engl J Med* **359**: 2641-2650.

Messersmith WA, Ahnen DJ. (2008). Targeting the EGFR in colorectal cancer. *N Engl J Med* **359**: 1834-1836.

Minquez B et al. (2009). Pathogenesis of hepatocellular carcinoma and molecular therapies. *Curr Opin Gastroenter* **25**: 186-194.

Mok TS et al. (2009) Gefitinib or Carboplatin-Paclitaxel in pulmonary adencocarcinoma. *N Engl J Med* **361**: 947-957.

Morgan KJ, Gilliland G. (2008). A role for Jak2 mutations in myeloproliferative diseases. *Ann Rev Med* **59**: 213-222.

Mosse YP et al. (2008). Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* **455**: 930-935.

Mosse YP et al. (2009). Inhibition of ALK signaling for cancer therapy. *Clin Cancer Res*. **15**: 5609-5614.

Mullighan CG et al. (2007). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**: 1377-1380.

Mullighan CG. (2009). Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med* **360**: 470-480.

Mulligan G et al. (2009). Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. *Blood* **109**: 3177-3188.

Nair B et al. (2009). Gene expression profiling of plasma cells at myeloma relapse from total therapy 22 predicts subsequent survival. *Blood* **113**: 6572-6575.

Ng PC et al. 2009a). An agenda for personailzed medicine. Nature 461: 724-726.

Ng SB et al. (2009b). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**: 272-276.

Ngo VN et al. (2006). A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* **441**: 106-110.

Nusbaum C et al. (2006). DNA sequence and analysis of human chromosome 8. *Nature* **439**: 331-335.

Oberthuer A et al. (2006). Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. *J Clin Oncol* **24**(31): 5070-5078.

Ohira M et al. (2005). Expression profiling using a tumor-specific cDNA microarray predicts the prognosis of intermediate risk neuroblastomas. *Cancer Cell* **7**(4): 337-350.

Paik S. (2007). Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with tamoxifen. *Oncologist* **12**(6): 631-635.

Papadopoulos N, Kinzler KW, Vogelstein B. (2006). The role of companion diagnostics in the development and use of mutation-targeted cancer therapies. *Nature Biotech* **24** 985-995.

Papaemmanuil E et al. (2009). Loci on 7p12.2,10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet* on line 16th August 2009.

Parsons DW et al. (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**(5897): 1807-1812.

Payton J et al. (2009). High-throughput digital quantification of mRNA abundance in primary human acute myeloid leukemia samples. *J Clin Invest* **119**: 1714-1726.

Perou CM et al. (2000). Molecular portraits of human breast tumours. *Nature* **406**: 747-752.

Pharaoh PDP et al. (2008). Polygenes, risk prediction and targeted prevention of breast cancer. *N Engl J Med* **358**: 2796-2803.

Pomerantz MM et al. (2009) The 8q24 cancer risk variant shows long range interaction with MYC in colorectal cancer. *Nat Genet* **41**: 882-884.

Potti A et al. (2006). A genomic strategy to refine prognosis in early stage non-small-cell-lung cancer. *N Engl J Med* **355**: 570-580.

Prickett TD et al. (2009). Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in *ERBB4. Nature Genetics on line Aug 30th 2009*

Pushkarev D, Neff NF, Quake S. (2009). Single -molecule sequencing of an individual human genome. *Nat Biotechnol* on line August 2009.

Radtke I et al. (2009). Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia. *Proc Natl Acad Sci USA* **106**: 12944-12949.

Ramaswamy S, Perou CM. (2003). DNA microarrays in breast cancer: the promise of personalised medicine. *Lancet* **361**(9369): 1576-1577.

Raponi M et al. (2008). A 2-gene classifier for predicting response to the farnesyl transferase inhibitor tipifarnib in acute myeloid leukemia. *Blood* **111**: 2589-2596.

Rassenti LZ et al. (2008). Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia. *Blood* **112**: 1923-1930.

Saal LH et al. (2007). Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. *Proc Natl Acad Sci USA* **104**: 7564-7569.

Savage S. (2008). Cancer genetic association studies in the genome-wide age. *Pers Med* **5**: 589-597.

Sawyers CL. (2008). The cancer biomarker problem. *Nature* **452**: 548-552.

Scales S, de Sauvage FJ. (2009) Mechanisms of Hedgehog pathway activation in cancer and implications for therapy. *TIPS* **30**: 303-312.

Schmidt L et al. (1997). Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. *Nat Genet* **16**: 68-73.

Scholl C et al. (2009). Synthetic lethal interaction between oncogenic KRAS dependency and STK33 supression in human cancer cells. *Cell* **137**: 821-834.

Schramm A et al. (2005). Prediction of clinical outcome and biological characterization of neuroblastoma by expression profiling. *Oncogene* **24**(53): 7902-7912.

Shackleford DB Shaw RJ. (2009). The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nat Rev Cancer* **9**: 563-575.

Shah NP et al. (2007). Sequential ABL kinase inhibitor therapy selects for compound drug-resistant BCR-ABL mutations with altered oncogenic potency. *J Clin Invest* **117**; 2562-2569.

Shah SP et al. (2009). Mutation of FOXL2 in granulosa-cell tumours of the ovary. *N Engl J Med* **360**: 2719-2729.

Sharma SV et al. (2007). Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* **7**: 169-181.

Shaughnessy JD et al. ,(2007). A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* **109**: 2276-2283.

Shedden K. (2008). Gene expression-based survival prediction in lung carcinoma: a multisite, blinded validation study. *Nat Med* **14**: 822-827.

Singh A et al. (2009). A gene expression signature associated with "KRAS addiction" reveals regulators of EMT and tumor cell survival. *Cancer Cell* **15**: 489-500.

Siva N (2008). 1000 genomes project. *Nat.Biotech* **26** 256.

Soda M et al. (2007). Identification of the transforming EML4-ALK fusion gene in non-small cell lung cancer. *Nature* **448**: 561-566.

Sotelo J et al. (2009). Long-range enhancers on 8q24 regulate cMyc. *Proc Natl Acad Sci USA* in press.

Sotiriou C, Pusztai L. (2009). Gene expression signatures in breast cancer. *N Engl J Med* **360**(8): 790-800.

Soucy TA et al. (2009). An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**: 732-737.

Squire JA. (2009). TMPRSS2-ERG and PTEN loss in prostate cancer. *Nature Genetics* **41**: 509-510.

Stacey SN et al. (2007). Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Gen* **39**: 865-869.

Stacey SN et al. (2008). Common variants on 5p12 confer susceptibility to estrogen-positive breast cancer. *Nat Gen* **40**: 703-706.

Staudt LM. (2006). Molecular diagnosis of hematological cancers. *N Engl J Med* **348**: 1777-1785.

Staudt LM, Dave S. (2005). The biology of human lymphoid malignancies revealed by gene expression profiling. *Adv Immunol* **87**: 163-208.

Stebbing J. (2008). FLEX data, KRAS and ERCC1 testing in oncology. *Fut Oncol*. 4: 471-473.

Steinbrook R. (2008). Personally controlled on-line health data—the next big thing in medical care? *N Engl J Med* **358**: 1653-1656.

Stratton M. (2008). Genome resequencing and genetic variation. *Nat Biotech* **26**: 65-66.

Stratton MR, Campbell P, Futreal PA. (2009). The Cancer Genome. *Nature* **458**: 719-724.

Taylor JG et al. (2009). FGFR4 is a mutationally activated oncogene that promotes metastasis in rhabdomyosarcoma. *J Clin Invest* in press.

Thomas G et al. (2008). Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Gen* **40**: 310-315.

Thomas G et al. (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Gen* **41**: 579-584.

Thorgeirsson TE et al. (2008). A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**: 638-642.

Tomlinson I et al. (2008). A genome-wide association study identifies colorectal cancer susceptibility loci on chromosome 10p14 and 8q23.3. *Nat Gen* **40**: 623-630.

Tol J et al. (2009). Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. *N Eng J Med* **360**: 563-572.

Tsai J et al. (2008). Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc Natl Acad Sci USA* **105**: 3041-3046.

Turner,NC et al. (2008). A synthetic lethal siRNA screen to identify genes mediating sensitivity to a PARP inhibitor. *EMBO J* **27**: 1368-1377.

Tuupanen S et al. (2009). The common colorectal cancer predisposition SNP rs 698267 at chromosome 8824 confers potential to enhanced Wnt signaling. *Nat Gen* **41**: 885-890.

Tyner JW et al. (2009). RNAi screen for rapid therapeutic target identification in leukemia patients. *Proc Natl Acad Sci USA **106:** 8695-8700*

Vander Heiden MG, Cantley LC, Thompson CB. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**: 1029-1033.

Vasquez A. et al. (2008). The genetics of the p53 pathway, apoptosis and cancer therapy. *NatureRevDrug Disc* **7**: 979-987

Vineis P, Schulte P, McMichael AJ. (2001). Misconceptions about the use of genetic tests in populations. *Lancet* **357**: 709-712.

Walther A, et al. (2009). Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer* **9**: 489-499.

Wang J et al. (2008). The diploid genome sequence of an Asian individual. *Nature* **456**: 60-65.

Wei JS et al. (2004). Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. *Cancer Res* **64**(19): 6883-6891.

Weinstein B & Joe A. (2008). Oncogene addiction. *Cancer Res*. **68**: 3077-3079.

Wiestner A et al. (2003). ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome and distinct gene expression profile. *Blood* **101**: 4944-4951.

Witte J. (2009). Prostate cancer genomics: towards a new understanding. *Nat Rev Gen* **10**: 77-82.

Wood L D et al. (2007). The genomic landscapes of human breast and colorectal cancers. *Science* **318**: 1108-1113.

Xia B et al. (2006). Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* **22**: 719-729.

Xie Y, Minna J. (2008). Predicting the future of people with lung cancer. *Nat Med* **14**: 812-813.

Yan H et al.(2009). IDH1 and IDH2 mutations in gliomas. *N Engl J Med* **360**: 765-773.

Yang JJ et al. (2009). Genome-wide interrogation of germline genetic variation associated with treatment response in childhood acute lymphoblastoid leukemia. *JAMA* **301**: 393-403.

Yashimoto M et al. (2008). Absence of *TMPRSS2:ERG* fusions and *PTEN* losses in prostate cancer is associated with a favorable outcome. *Modern Pathology* **21**: 1451-1460.

Yeager M et al. (2008). Comprehensive resequence analysis of a 136kb region of human chromosome 8q24 associated with prostate and colon cancers. *Hum Genetics* **124**: 161-170.

Yoshida T et al. (2006). Spreds, inhibitors of the RAS/ERK signal transduction are dysregulated in human hepatocellular carcinoma and linked to the malignant phenotype of tumors. *Oncogene* **25**: 6056-6066.

Yu SL et al. (2008). MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell* **13**:48-57.

Zanke BW et al. (2007). Genome–wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Gen* **39**: 989-994.

Zhan F et al. (2008) High-risk myeloma: a gene expression-based risk-stratification model for newly diagnosed multiple myeloma treated with high dose therapy is predictive of outcome in relapsed disease treated with single-agent bortezomib or high-dose dexamethasone. *Blood* **111**: 968-969.

Zheng SL et al. (2008). Cumulative association of five genetic variants with prostate cancer. *N Engl J Med* **358**: 910-919.

Acknowledgements:

The authors would like to thank Drs Javed Khan, David Munroe, Bob Stephens, Lou Staudt, Snorri Thorgeirsson and Mickey Williams for providing information and for helpful comments on the manuscript.

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.