Cancer Genetics 2013: Seeking Meaning in Our Patients' Genomes

Theodora S. Ross, MD, PhD

Internal Medicine Grand Rounds

University of Texas Southwestern Medical Center

March 1, 2013

This is to acknowledge that Theodora Ross, MD, PhD has disclosed that she does not have any financial interests or other relationships with commercial concerns related directly or indirectly to this program. Dr. Ross will not be discussing off-label uses in her presentation.

Theodora S. Ross, MD, PhD
Professor
Co-Director, High Risk Cancer Genetics Program
Division of Hematology/Oncology
Department of Internal Medicine

Purpose: The presentation will be to introduce the audience to cancer genetics in the age of genome sequencing by taking the audience through the history of genetics as it relates to the activities of cancer biologists, patients and doctors. I will compare the situation of the cancer patients that were diagnosed in the 1980s and 90s to those of today. The hypothesis that we will explore is that next generation sequencing of our patients' tumor and germline genomes is a new test that will revolutionize the future of medicine, especially in the area of cancer.

Overview: I will start with a brief introduction of the specific cast of cancer patients that will be discussed and then contrast cancer genetics of the 1990s with cancer genetics of the 2010s. I will provide a history of how we arrived at the "cheap" genome, highlighting the most interesting historic milestones. We will review the power of the new genetics with a basic science vignette that illustrates how the analysis of the Tasmanian devil facial tumor disease using whole genome sequencing identified much about the genome of the long dead founder devil. This will be followed by a discussion of results from sequencing human tumors and germlines of cancer patients from cancer-prone families. The hope is that in those genomes, we will understand why some patients with cancer predisposing mutations don't get cancer while others do. I will then wrap up with a summary of what the cancer genetics challenges of the future are for patients, doctors and scientists.

Objectives:

1. Learn about the history of how we have almost arrived at the $1000 genome.
2. Determine if we are really in the golden age of cancer genetics. Are we providing true or false hope to our patients with tumor sequencing?
3. Determine what the biggest cancer genetics challenges are in the age of the $1000 dollar genome.
4. Learn about important tumor and germline sequencing successes and failures that have resulted from the new genetic analysis.

Whole exome sequencing and whole genome sequencing (WGS) over the last five years have advanced our understanding of health and disease especially when genetic predisposition for a disease is well defined[1]. Examples include the identification of *PALB2* as a familial pancreatic cancer susceptibility gene[2] or mutations in the *SH3TC2* gene as a cause of Charcot-Marie-Tooth Neuropathy[3]. With the introduction of next-generation sequencing technologies, it is now possible to sequence an entire genome more rapidly, accurately, and cost-effectively than ever (Figure 1)[4].

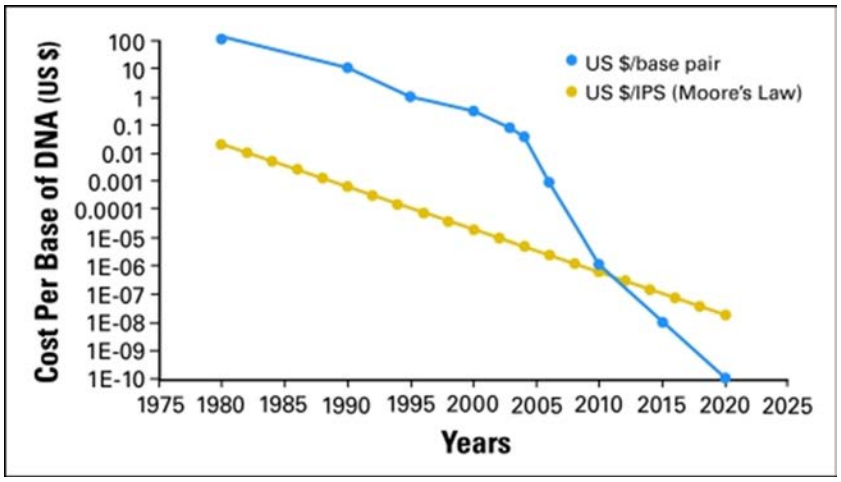

Figure 1. Moving towards that $1,000 genome. The blue line = rate of decrease in cost of sequencing and the yellow line indicates Moore's law which is the doubling of computer instruction per second per dollar.

Cancer is a good candidate problem for the use of genome sequence analysis because it is, in part, a genetic disease. Molecular alterations in cancer include chromosomal rearrangements, point mutations, insertions and deletions that have been identified over the last 50 years using Southern blots, RT-PCR, chromosome walking and more recently DNA microarrays and massively parallel sequencing (Figure 2).

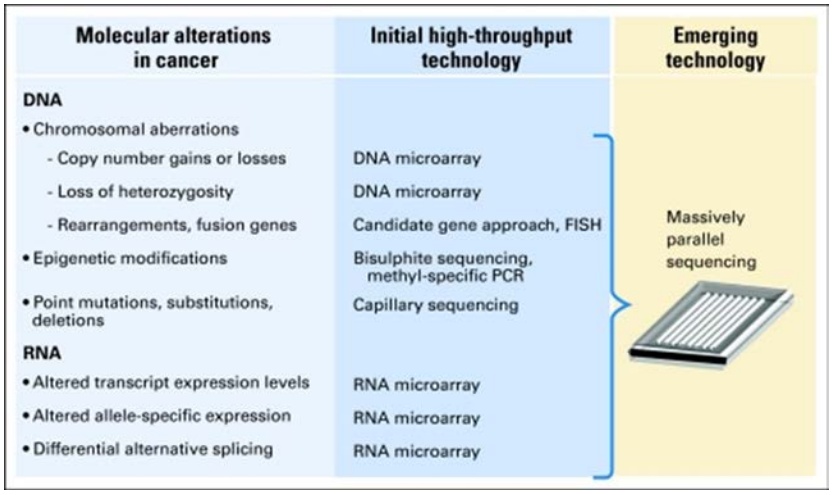

Figure 2. Genetic Aberrations in Cancer. Testing for genetic changes in cancer and how they are detected has evolved from use of allele specific tests to massively parallel sequencing of whole genomes.

The advances in massively parallel technologies have reduced the cost in a staggeringly rapid fashion due to market competition and disruptive technologies in a myriad of areas including microscopy, biochemistry and computers (Figure 2).

The successful search for and discovery of cancer drug targets using genetic analyses in the last 40 years has made the first 10 years of this century unprecedented in the development of designer therapies. Well-known targeted therapies include Imatinib[5], Trastuzumab[6], Erlotinib[7,8] and Vemurafenib[9] that inhibit Bcr/Abl fusions[10], amplified Her2/neu[11], mutant EGFR[12] and BRAF[13], respectively. Unfortunately, these are not perfect drugs. They are useful for their ability to suppress disease but not eradicate it as exemplified by BCR/ABL inhibitors[14] and *BRAF* V600E inhibitors[15]. However, the basic science to better understand why we are unable to cure diseases such as chronic myelogenous leukemia (CML) and melanoma with targeted therapies is progressing at a rapid pace and would not exist if it were not for the development of Imatinib and Vemurafenib in the first place.

Table 1. Time Cancer Targets were Discovered to Time the Designer Therapies were approved.

Molecular defect	Year of discovery	Cancer	Designer Therapy	FDA approval
Bcr/Abl	1973	CML/ALL	Imatinib	2001
Bcr/Abl mutants	2002	CML	Desatinib, nilotinib	2006
Tel/PDGFbR et al.	1995	CMML	Imatinib	2001
PML/RAR	1990	APL	ATRA	1987
KIT	2001	GIST	Imatinib	2001
Her2/neu	1985	BrCa et al	Trastuzumab, lapatinib, pertuzumab	2001
CD20	1988	Lymphoma	Rituximab	1997
EGFR mutations	2004	NSCLCa	Erlotinib, gefitinib	2003
ALK fusions	2007	NSCLCa	Crizotinib	2010
BRAF	2001	Melanoma	Vemurafenib	2011
KRAS	1987	Colon cancer	Cetuximab or panitumumab hold	2006

With these designer drug successes of the recent past and the added availability of next-generation sequencing, the search for additional drug targets in tumor sequences has become the subject of large multi-investigator, multi-institutional research efforts such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (TCGC). The idea has been that because cancer is a disease of the genes cataloguing and analyzing tumor associated mutations will be important (Figure 3).

But the problem of discovering which variant is the culprit and which is an innocent bystander is not solved. There are 1,000s of variants in cancer that could contribute to the cancer problem. It gets even worse in general cancer-omics as even

those culprits that are necessary for some tumors (*BRAF* V600E in melanoma) are not necessary for the survival of other tumors (*BRAF* V600E in colon cancer)[16].

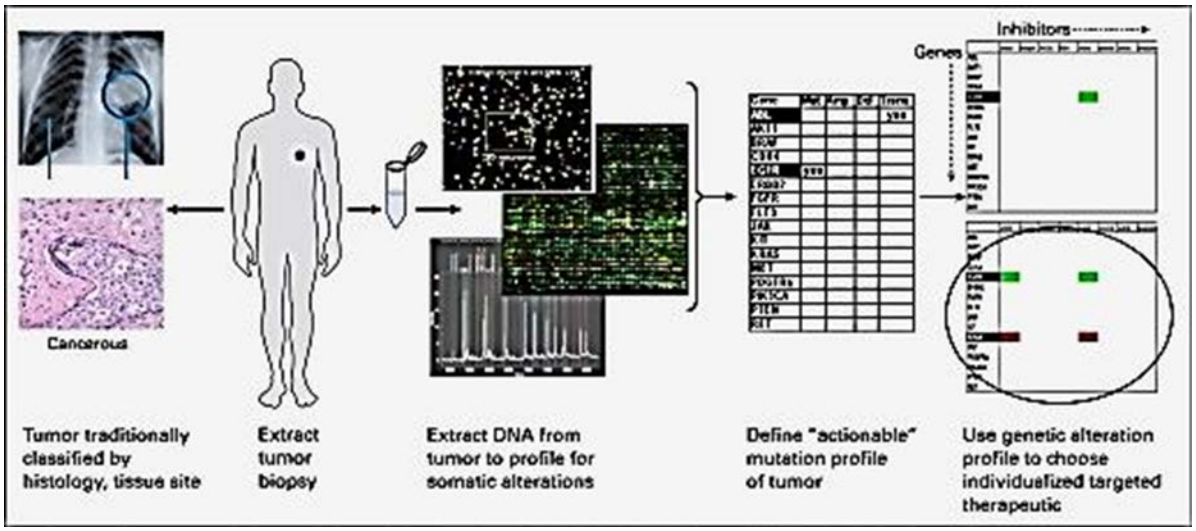

Figure 3. Tumor variant analysis to practice "precision medicine". A better way to treat suffering? Will tumors in the future be classified by "omics" rather than site of origin?
Adapted from MacConaill and Garraway 2010 JCO 28, 5219

The first publications describing "full" cancer genomes were those of AML [17,18], metastatic breast cancer [19], melanoma [20] and small cell lung cancer [21]. The AML studies identified the *IDH1* mutation in 16% of AML samples confirming an earlier identification of *IDH1* mutations in GBMs [22]. IDH1 is an isocitrate dehydrogenase that participates in the citric acid cycle. The purpose of this mutant in cancer is not understood, but its presence does support a role for altered metabolism in cancer rather than it simply being an effect of tumorigenesis. Manuscripts on tumor sequencing come out monthly and have been organized by Edwin Wang at: http://www.bri.nrc.ca/wang/cancerGenome.html. These papers emphasize that there are major challenges to WGS including bioinformatics and, most frustrating, the rate of variant discovery vastly outpaces the functional analysis rate. How to get past that wall to determine if a candidate mutation makes a contribution to the cancer problem is not known.

In additional to determining how to efficiently understand the biologic meaning of a single variant, tumor sequencing has led to other unanticipated challenges. Cancer genome analyses involve the interpretation of large datasets from genetically

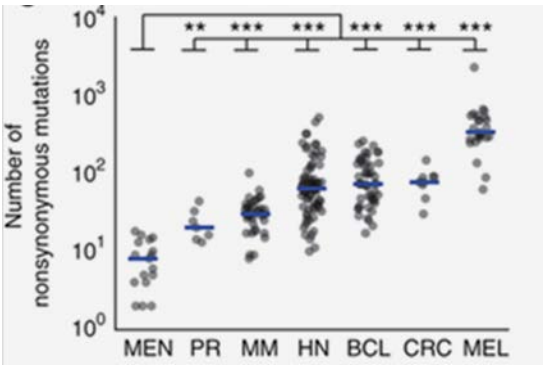

Figure 4. Genetic variation in tumors is enormous. Somatic alterations in grade I meningioma genomes (n=17) are less than prostate ca, colon ca, melanoma, B-cell lymphomas, Head and Neck cancers and multiple myelomas. Blue bar = median.
From Brastianos et al., Nat Gen. 2013 January online

heterogeneous tumors where bioinformatics methods are still developing. The number of somatic variants in one tumor or between tumors is greater than imagined. Although tumors of low histopathologic heterogeneity such as meningiomas have fewer variants than more heterogeneous tumors such as melanoma and colon cancer (Figure 4), they still have multiple candidate "pathogenic" variants. The reports of limited reproducibility and generalizability of published functional data adds yet another layer of complexity[23,24]. Still, the fact that cancer drivers discovered at the end of the last century have been verified as useful therapeutic targets rightfully encourages many to continue the search for additional targets in tumor genome sequences.

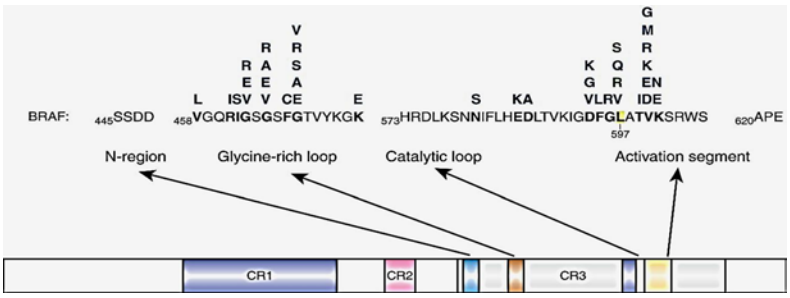

Figure 5. Schematic of oncogenic BRAF mutations. The most common mutations occur at codon 600 in the activation segment. Amino acid changes in are are in blue and are above the BRAF sequence.

A well-known cancer driver variant is the V600E/K of the BRAF serine-threonine kinase (Figure 5). Changes at this amino acid position indicate that the tumor will be sensitive to Vemurafenib a serine/threonine kinase inhibitor. However, a recent tumor sequencing study discovered a patient's melanoma had another BRAF tumorigenic mutation that substituted a serine at position L597 (Figure 5). This new mutation predicted that the tumor would not be sensitive to Vemurafenib but it would be sensitive to MEK/ERK inhibition and as suggested by the authors, it was[25].

They tested the oral MEK inhibitor TAK-733 in the patient with the L597S mutation. After 4 cycles of treatment they assessed the response (Figure 6). There was a limited response that was expected since WGS of the tumor normal pair also uncovered a number of other somatic mutations many of which were relevant to melanoma biology. These included KRAS, APC, BRCA2, NOTCH1, PTEN and NF1!

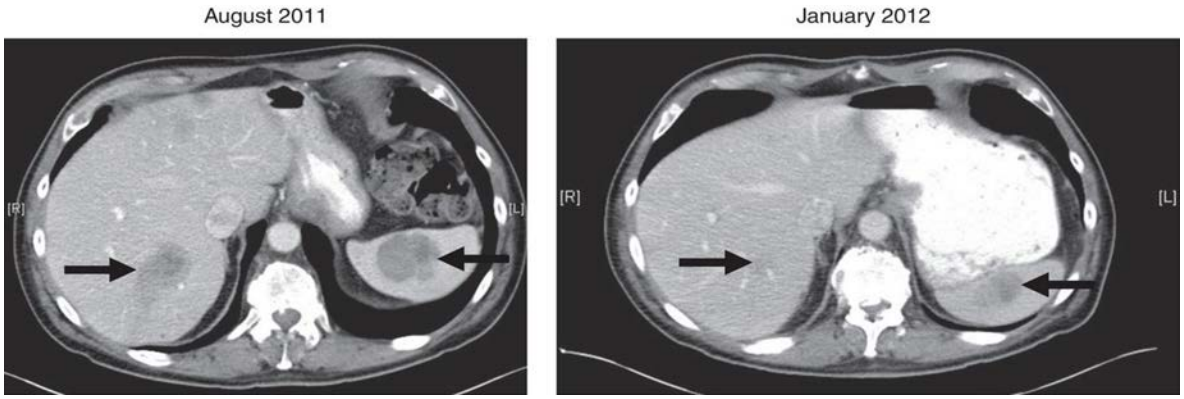

Figure 6. Response of metastatic melanoma to TAK-733. Arrows identify the melanoma in the liver and spleen. After cycle 4 (right) there is a limited but measurable response to the MEK inhibitor TAK-733.

Correctly selecting which of the many rare nonsynonymous variants in such tumors contribute to the maintenance or progression of cancer is a formidable task. One way to assess biological significance of a mutation is to observe whether or not such aberrations occur repeatedly in tumors from different patients. For example, Mullighan and colleagues have discovered that somatic mutations in the c-AMP response element binding protein (*CREB*) binding protein (*CBP* or *CREBBP*) recur in B-cell acute lymphoblastic leukemias (Figure 7)[26]. In support of *CREBBP*'s potential functional significance, several groups have since found recurrent *CREBBP* mutations in B-cell lymphomas[27,28], transitional cell cancers of the bladder[29] and small cell lung cancers[30]. CREBBP is a molecular scaffold that binds a large number of transcription factors and

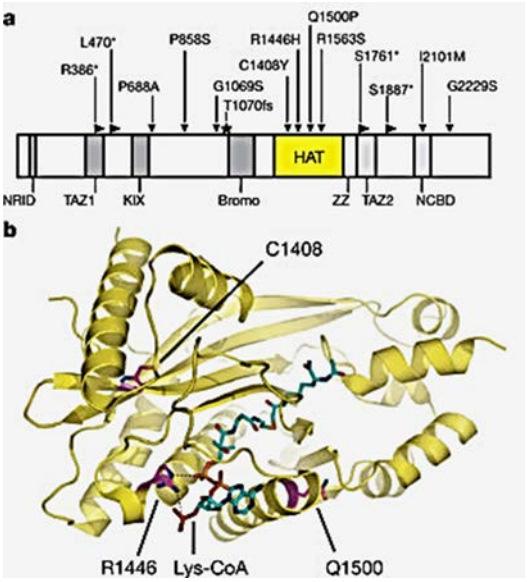

Figure 7. CREBBP sequence mutations in relapsed ALL identified with WGS.

regulates transcription by acetylation of histone and non-histone targets[31]. Because of CREBBP's role as an acetyltransferase, mutations in this gene possess potential therapeutic significance as the cellular loss of acetylation activity leads to an increase in relative de-acetylase activity. This raises the possible therapeutic application of histone deacetylase (HDAC) inhibitors to restore the acetylation/de-acetylation balance in patients with tumors containing *CREBBP*-deactivating mutations.

From tumor genome sequencing it is too early to know if such ideas for therapy will pan out in the clinic. The phenomenal basic science discoveries have arrived, but so far no new tumor therapies have been developed as a result of genome sequencing. Past hypothesis rather than non-biased sequencing studies have identified transforming alterations in RAS, RAF, EGFR, PI3Kinase, Her2/neu, Myc, Abl, Ets and ALk. These genes have been to some extent successfully targeted and this success encourages an aberration-centric rather than histo-centric approach to therapeutic development. In fact this omics view may be very useful and as impactful as frequent "druggable" aberrations in one tumor type as low frequencies in many tumor types can add up across tumors and be similar to those at high frequency in one tumor type. The success of crizotinib in ALK rearranged Non-Small Cell Lung Cancer and lymphoma supports this[32]. But this is not a given as exemplified with Vemurafenib which does not show effects in *BRAF* V6600E mutant colon cancers. As mentioned earlier, this brings us back to the major challenge which is our limited understanding of the true role of rare variants in cancer pathophysiology.

In addition to finding a variant repeatedly in many tumors, another way to address the biological significance of a variant is to turn to the simpler germline of patients at a high risk for a cancer diagnosis. But this could be a limited strategy as it is

well known that much of neoplastic disease is not familial as shown by Lichtenstein et al[33] and more recently Roberts et al[34] where they compared the cancer rates of monozygotic and dizygotic twins. These data indicate that sequencing every cancer patient is likely to be an uninformative test; most of cancer is "nurture rather than nature." The strategy of both of these groups was that if twins show a higher concordance for cancer when identical (share all genes) than the non-identical twins (share 50%) or siblings, genetic effects are important. On the other hand, if the risk is the same in both monozygotic and dizygotic twins then shared environmental effects are probably more important components of the cancer causing equation. Analysis of cohorts of twins from the United States, Sweden, Denmark and Finland indicated that genetics for most cancers makes a minor contribution to cancer susceptibility and instead the environment (cigarettes and alcohol) plays a principal role in causing sporadic cancer. Of interest was that colon, breast and prostate cancers do have a solid genetic component (Table 2). And of course in some situations such as adenomatous polyposis coli (APC), Li Fraumeni (p53) and hereditary breast and ovarian cancer syndromes caused by *APC*, *p53* and *BRCA1/2* mutations respectively, the genetic component is well defined.

Table 2. Absolute risks of cancer (concordance) in twins of an affected person up to the age of 75.

Cancer Site	Monozygotic	Dizygotic
colon	0.11	0.05
breast	0.13	0.09
prostate	0.18	0.03

Adapted from Lichtenstein et al NEJM 2000 343, 78

BRCA1 and *BRCA2* were discovered in 1994[35,36] and 1995[37], respectively, by old fashioned genetics and when they were discovered there was a debate about the value of the discovery. However, one use was not up for debate. Both BRCA1 and BRCA2 were and continue to be fascinating to cancer biologists. It has been found that BRCA1 is a multi-domain protein (figure 8) that has several functions and works together with BRCA2 to repair damaged DNA (figure 9). The clinical debate

Figure 8. Schematic of BRCA1 domain structure.

when these genes were cloned 18 years ago was whether we were going to offer such testing to patients and whether anyone would want to know information when there was nothing they could do about it. The heated discussions were widespread among ethicists, patients, genetic counselors, doctors and the media.

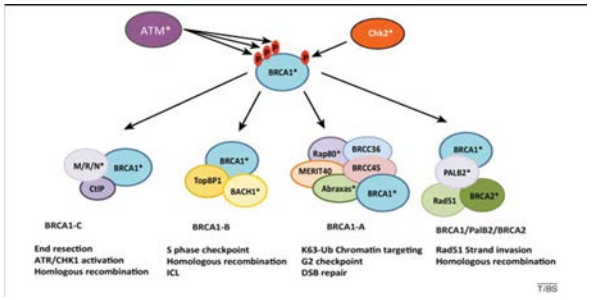

Figure 9. BRCA1 interactions are myriad.

But almost 20 years later we have in fact adopted widespread testing because we realize that there were many women who died of breast and ovarian cancer unnecessarily. These deaths could have prevented if we knew they had a genetic mutation. A series of landmark studies by Rebbeck et al in in the early 2000s showed that prophylactic surgeries are effective a reducing breast and ovarian cancers by 90-95%[38,39] (Figure 10). Early detection and then treatment with the possible risks of morbidly and mortality of a cancer diagnosis is not equivalent to surgical prevention. So, while bilateral prophylactic mastectomy (BPM) is considered ideal from a doctor's view, it is of course optional for some from a quality of life view. In contrast bilateral salpingo-oopherectomies is a standard of care that is strongly recommended after childbearing. This is because there is no effective screening method for ovarian cancer.

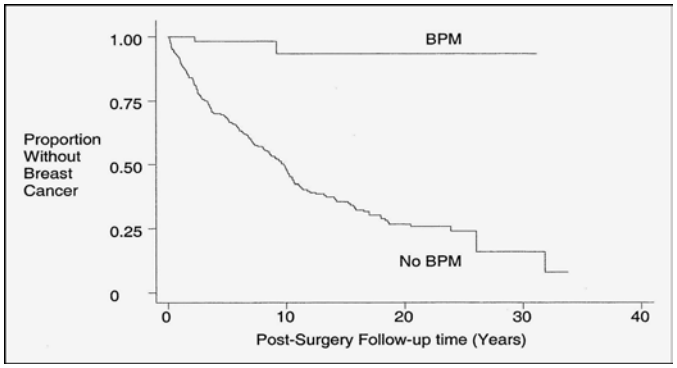

Figure 10. Breast cancer prevention in *BRCA1/2* mutation carriers with bilateral prophylactic mastectomy (BPM).

The risk of occurrence of ovarian and breast cancers by age 70 is increased approximately 10 fold (44% and 87%, respectively) in patients with germline mutations in *BRCA1*[40]. There is also a wide range of breast and ovarian cancer risk associated with *BRCA1* (and *BRCA2*) mutations between different families, suggesting that the presence of additional cooperating germline mutations and/or environmental stresses

contribute to the development of “genetic” cancers. It is of interest to note that there is a significant fraction (on average 25%) of mutation carriers that live into their 90s without having cancer. Furthermore, the youngest age at breast cancer diagnosis in a family can help predict when other family members will develop breast cancer[41]. This familial risk suggests that there may be more than the *BRCA1/2* mutations in the genome to potentiate or abrogate the effects of the *BRCA1/2* mutations. Could whole genome sequence analysis help us improve on the precision of our predictions for *BRCA1/2* patients? The answer to this could also help us understand if sporadic mutations in tumors, such as the *CREBBP* mutations (Figure 7), are functionally significant. If rare hypomorphic variants in the germ lines of patients with early-onset cancer syndromes such as those involving the *BRCA1/2* genes, were also found sporadically in tumors the evidence for their driver role would be compelling. The Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) asks a similar question but focuses on determining

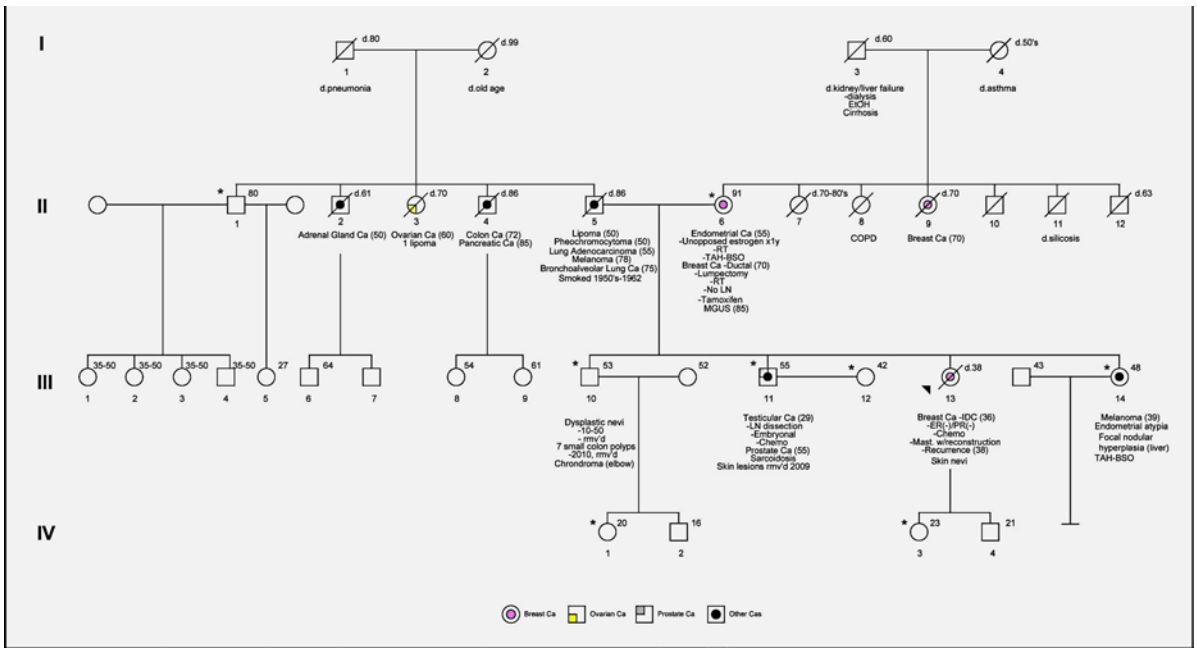

Figure 11. Pedigree of index *BRCA1* 5382insC family. Four generations of the family of the index patient (arrow). Pertinent medical information is noted.

potential BRCA1 and BRCA2 modifiers using genome-wide association studies[42].

We have in fact obtained data that expands upon CIMBA’s approach through the use of WGS analysis of a cancer-prone *BRCA1*-mutant family (Figure 11) to investigate

potential cooperating mutations that increase their cancer risk. We chose this family because the *BRCA1* mutation was, to our surprise, not inherited from the paternal side of the family despite the fact that this side of the family had several cancers such as ovarian, pancreatic, melanoma that are all known to be associated with BRCA mutations[43]. We reasoned that if we were to sequence key family members, we might discover a pathogenic mutation from the father that potentiated the effects of the maternal *BRCA1* 5382insC mutation.

After consenting the family members we obtained skin biopsies and blood samples and provided DNA to Complete Genomics, Inc. for an unbiased WGS test for other coopering mutations. After samples were sequenced reads were aligned to the reference genome and all differences were identified through the process of variant calling and we received a long list of genomic variations (Figure 12). First there were common variants or synonymous variants, then there are those that were less common and then a lot of bioinformatics occurs. This ultimately led to the identification of mutations in *TSC2* and *CREBBP* that may cooperate with the *BRCA1* mutations. We found these mutations in three members of the second paternal generation (II:2, 4 and 5) and 6 members of the next generation (III:6,7,8,10,11,14). Only the two living but affected members of the third generation (III:11 and III:14) to our knowledge had these mutations as well as the *BRCA1* 5382insC. Of interest, is the fact that these two individuals had earlier cancer diagnoses (20s-30s) than all other affected family members. It is also possible that the index *BRCA1* 5382insC patient (III:13) who was diagnosed with breast cancer at age 35 was *CREBBP/BRCA1* double mutant, as well.

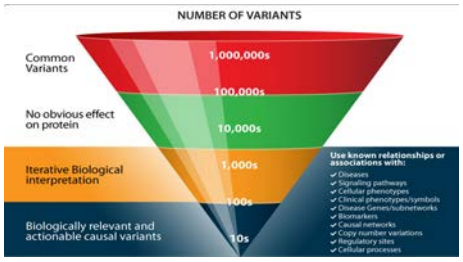

Figure 12. Schematic of the JJ Rios and Ingenuity bioinformatics

To determine if the identified *CREBBP* and *TSC2* mutations were unique to our index *BRCA1*-mutant family, we sequenced both variants in an independent cohort of 98 *BRCA1*-mutant breast cancer patients (representing 86 families). Although no additional individuals with the *TSC2* variant were identified, two sisters were identified as carriers of the *CREBBP* mutation. Both of these women carry the *BRCA1* 5382insC

mutation of the index family (Figure 11) and were diagnosed with infiltrating ductal carcinoma of the breast at the ages of 57 (III:5) and 60 (III:9) (Figure 13). The paternal and maternal sides of this second family were also afflicted with early onset cancer diagnoses, but inheritance of the *BRCA1* and *CREBBP* mutations could not be determined with the available material. The maternal great aunt (I:5) was diagnosed with breast cancer at age 30 in addition to two maternal aunts (II:7 and II:8) with breast cancer diagnoses in their 50s. A maternal cousin was diagnosed with a brain stem tumor and died at age 21 (III:10) and a paternal cousin died of ovarian cancer at age 50 (III:1). Finally, the child of another paternal cousin (III:2) died of leukemia at age 30 (IV:1) (Figure 13). The discovery of the *CREBBP* variant in a second family with a *BRCA1* mutation in this group lends support to the hypothesis that rare *CREBBP* variants may increase the penetrance of *BRCA* mutations' cancer phenotypes.

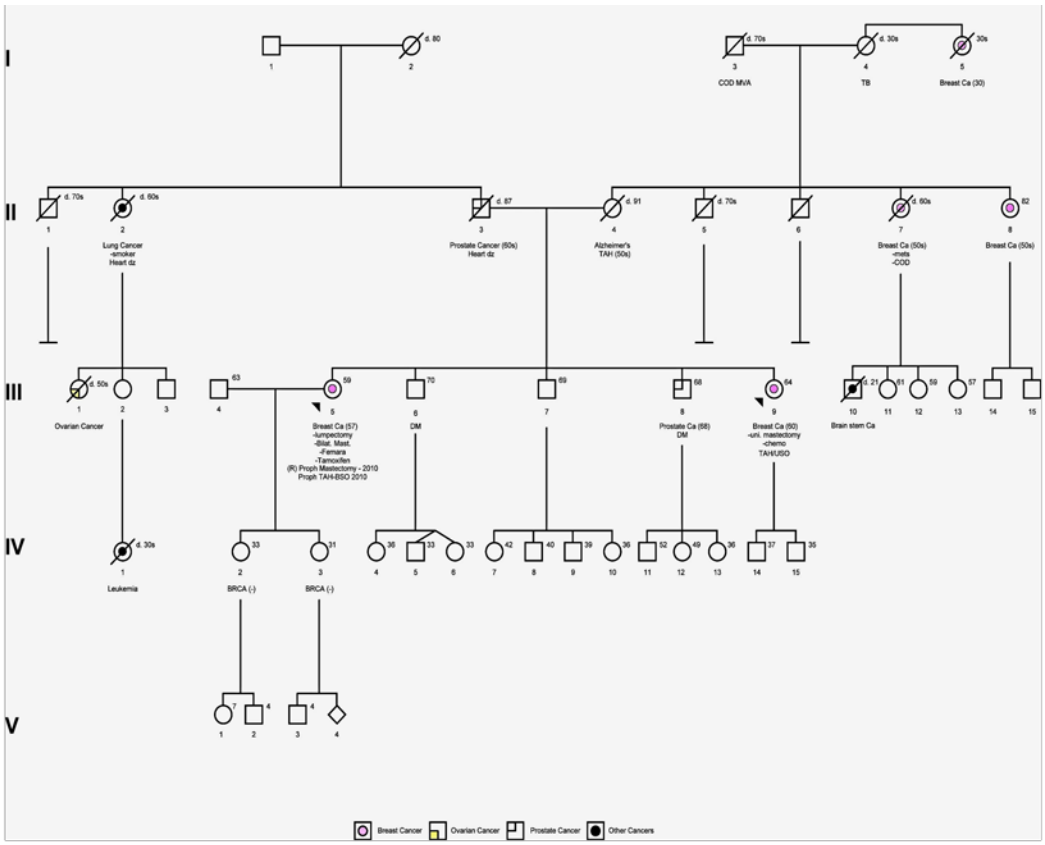

Figure 13. Pedigree of *BRCA1* 5382insC Family no. 2. Arrow head indicates *BRCA1* 5382ins C and CREBBP mutant positive patients.

To further test this modifier hypothesis, we sequenced 100 additional *BRCA2*-mutant individuals from 82 families for both the *CREBBP* and the *TSC2* P91L variants. Again, while no individuals with the *TSC2* variant were identified, two unrelated individuals (2.43%; 2/82 BRCA2 families) were positive for the *CREBBP* mutation. In family 3, the index patient was diagnosed with breast cancer at age 43 and her *BRCA2* 6672delT frame-shift mutation was paternally inherited. Of interest, this individual's offspring were all diagnosed with attention deficit disorder, which has been associated with *CREBBP* mutations[44]. The index patient from the fourth *CREBBP* mutant family was diagnosed with breast cancer at age 53 and carries the *BRCA2* 7297delCT frame-shift mutation. In total, the frequency of the *CREBBP* variant in *BRCA1/2* mutant families (2.3%; 4/168 BRCA1/2 families) is significantly different from the 1000 Genomes Project's frequency (0.37%; 4/1092 genomes[45]; Fisher's exact test, two-tailed P=0.0138) and encouraged us to further evaluate potential molecular differences between cells that were either mutant or wild type for the *BRCA1* and *CREBBP* genes.

Table III. Enrichment of the *CREBBP* mutation in BRCA1/2 families.

BRCA1/2 mutation	*CREBBP* mutation	Total (n)	frequency
BRCA1	4	100	4%
BRCA2	2	100	2%
BRCA1/2	6	200	3%
1000 Genome	4	1092	0.35%

Since individuals III:11 and III:14 from the index family (Figure 11) possess rare variants with predicted loss of function changes in both the *BRCA1* and *CREBBP* genes, it is possible that cancer cells from these individuals could have abnormal *BRCA1/CREBBP* cooperation, contributing to the loss of tumor-suppressing function of *BRCA1* or vice versa. To begin to understand the effects of *BRCA1* and *CREBBP* mutations on gene expression in the index family, RNA from 24 induced pluripotent stem (iPS) cell lines generated from this family were analyzed for differential expression. IPS cell lines were generated from fibroblasts using non-integrating mRNA transfection methods[46], and these iPS lines were found to be pluripotent and capable of differentiating into endodermal, mesodermal and ectodermal derived tissues (Soyombo et al, in preparation). We chose to analyze expression differences in iPS cells because de-differentiation from fibroblast to stem cell (iPS cell generation) is considered a surrogate step in the path to transformation. As such, the gene expression patterns of

iPS cells have been deemed similar to tumor cells[47]. Changes in gene expression greater than 2.0-fold were compared in iPS cells from individuals with both the *BRCA1* and *CREBBP* mutations (III:11 and III:14; n= 6) and iPS cells from other family members (II:1, II:6, III:10, III:12, IV:1, IV:3; n=18). Three genes (*PIWIL2, ALG11* and *ODF2L*) displayed significant changes in gene expression consistently in all test iPS lines (Figure 14A). *PIWIL2*, the gene with the greatest change in expression, was induced approximately 4-fold in the *BRCA1/CREBBP* mutants compared to *BRCA1* mutants or wild-type cells (Figure 14A and B). Increased *PIWIL2* expression in double-mutant iPS cells was confirmed by qPCR (Figure 14C). In contrast, we were not able to independently confirm differential expression of either *ALG11* or *ODF2L* in *BRCA1/CREBBP* mutant

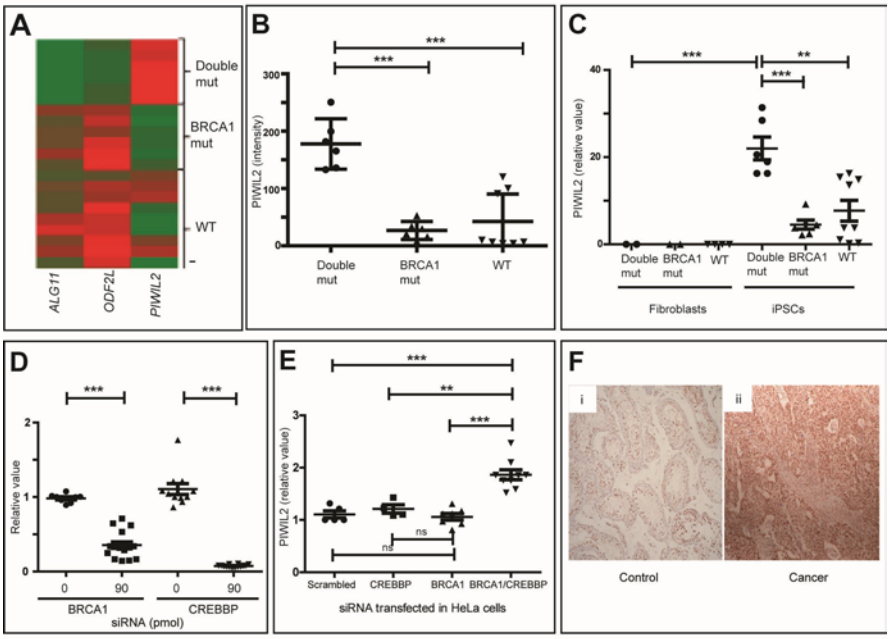

Figure 14. PIWIL2 expression patterns in BRCA1/CREBBP mutants

A. Heat map representing relative expression levels of *ALG11, ODF2L,* and *PIWIL2* in iPS cells from double mutant (*BRCA1/CREBBP* mutant), *BRCA1* mutant, and wild type individuals. Green: low expression, red: high expression.

B. Relative *PIWIL2* expression levels (intensity) from the RNA expression array were compared amongst double mutant (*BRCA1/CREBBP* mutant), *BRCA1* mutant, and wild type.

C. qPCR indicates that fibroblasts express minimal PIWIL2 relative to iPS cells and that PIWIL2 is elevated in the double mutant iPS cells.

D. Knockdown of *CREBBP* and *BRCA1* in HeLa cells.

E. a significant increase in the BRCA1/CREBBP double knockdown cells compare to control or single gene knockdown cells. Unpaired t-test for statistical significance was performed: *p<0.05; **p<0.01; ***p<0.001; ns = not significant.

F. PIWIL2 immunohistochemistry of (i) normal testis and (ii) testicular cancer from III:11.

iPS cells. Increased expression of *PIWIL2* occurred in iPS cells derived from III:11 and III:14 fibroblasts; however, as expected due to its classification as a cancer testes antigen[48,49], *PIWIL2* expression was not detected in any of the terminally differentiated fibroblasts from the index family (Figure 14C). *PIWIL2* is a cancer-testis antigen required for stem cell self-renewal and has been shown to be overexpressed in a wide variety of tumors[48,49]. PIWIL2, like other PIWI proteins, binds to PIWI-interacting RNAs and is thought to promote tumorigenesis by decreasing genetic stability and epigenetic changes[50].

To determine whether a change in *PIWIL2* expression was coincidental or an artifact of the iPS cell system and limited sample size, we sought to confirm the relationship between *BRCA1*/*CREBBP* and *PIWIL2* expression in an independent cell-based assay. Because BRCA1 and CREBBP proteins have been shown to interact physically and functionally[51], we investigated the possibility that both BRCA1 and CREBBP cooperatively alter *PIWIL2* expression using siRNA knockdown in HeLa cells, a cell type where we achieved successful knockdown (Figure 14D). Similar to iPS cells, *PIWL2* expression increased two-fold when both BRCA1 and CREBBP were knocked down in HeLa cells; however, *PIWIL2* expression remained unchanged when either BRCA1 or CREBBP alone was knocked down (Figure 14D and E). We also stained the testis tumor from the index patient's brother (III:11) for PIWIL2 expression and observed significant expression compared to normal testis tissue (Figure 14F). Future analysis of PIWIL2 levels in other *CREBBP* and/or *BRCA1* mutant cancers will determine the extent to which PIWIL2 is expressed in these cancers[50].

Although researchers have identified a number of familial cancer-causing mutations, the genetics of why some patients with these mutations develop cancer and others do not are poorly understood. The CIMBA investigators focus on determining potential *BRCA1* and *BRCA2* modifiers or factors that cooperate with *BRCA1* and *BRCA2* using common single nucleotide polymorphisms (SNPs)[*42*]. Results of the sequencing of the index family presented here (Figure 11) highlight the added value of WGS analyses to identify rare variants associated with the cancer phenotype. We sought to identify a potential modifying gene that cooperatively increases cancer risk in

a *BRCA1*-mutant cancer-prone family. WGS analysis confirmed a maternally-inherited *BRCA1* 5382insC Ashkenazi founder mutation and identified paternally-inherited mutations in the tumor-suppressor genes *CREBBP* and *TSC2*. Mutations in these genes have been reported previously in patients with tuberous sclerosis (*TSC2*) and Rubinstein-Taybi syndrome (*CREBBP*), syndromes that are associated not only with increased predisposition to cancer but also other phenotypes[52-56]. Of interest, the other phenotypes reported for Rubinstein-Taybi syndrome (short stature, cognitive deficits) are absent in the index family. In fact, this might be expected as the severe loss of function mutations in *CREBBP* that cause florid dysmorphologies and cognition problems are rarely transmissible in the germline but are rather *de novo* mutations[57]. It is possible that less severe mutations, such as those that alter single conserved amino acids in the transcriptional transactivation domain of CREBBP, will lead to increased cancer risks in the absence of other phenotypes.

As mentioned earlier, Mullighan and colleagues reported that somatic *CREBBP* mutations are commonly associated with acute lymphoblastic leukemia relapse[26] and others have found somatic mutations in B-cell lymphomas, small cell lung cancers and bladder cancers[27-30]. In fact, the specific variant identified in the four families with *BRCA1/2* mutations in our work was reported as associated with acute lymphoblastic leukemias[26]. Furthermore, the CREBBP amino acid position is conserved in eight species including mouse[53] and the tumors observed in Rubinstein-Taybi patients include not only breast and ovarian cancers but also pheochromocytomas and testicular tumors, tumor types diagnosed in this study's index family [57,58]. *CREBBP* is a transcriptional co-activator and histone acetyltransferase that physically and functionally interacts with *BRCA1*[51]. Therefore, cancer patients with *BRCA1* and *CREBBP* mutations may benefit from treatment with histone deacetylase (HDAC) inhibitors. To our knowledge, the experimental use of HDAC inhibitors in these cancer patients has not yet been attempted most likely due to the lack of pre-clinical mouse data. However, mice with heterozygous deletion of *CREBBP* develop lymphomas[59], making them a useful model from which pre-clinical data will need to be collected. If HDAC inhibitors are therapeutic in such mice, this will provide evidence supporting the functional

significance of these mutations and encourage further studies in patients with *CREBBP* mutant tumors whether germline or sporadic.

The observation that *PIWIL2*, a stem-cell self-renewal gene[60], is elevated in cells from members of our index family with both *BRCA1* and *CREBBP* mutations may also be of future clinical significance. *PIWIL2* has been reported to be specifically expressed in testes and several tumor types including testes, breast and endometrial[48]. As such, *PIWIL2* has been classified as a cancer-testis antigen and is a novel potential target for therapy in these individuals. Further studies to investigate the role of *PIWIL2* in the cancer biology of *CREBBP* mutant tumors originally identified by Mullighan et al[26] will be necessary to understand if its elevated expression is downstream of *BRCA1* and/or *CREBBP* to promote tumorigenesis.

Family-based WGS and gene expression analyses together can provide a detailed genetic and phenotypic understanding of the molecular pathways involved in a family's cancer phenotype and can identify putative new molecular modifiers of BRCA1 and BRCA2, which may lead to novel therapies in such patients. These types of human studies are important to combine with mouse studies as the genetic background in humans in which tumors develop is, unlike uniform genetic backgrounds of mouse models, heterogeneous. Background heterogeneity may contribute to the penetrance and severity of human tumor phenotypes. Further studies of cancer-prone *BRCA1/2* families using WGS and stem cell analyses will contribute to a better understanding of cancer's complexity in humans.

Summary: The hypothesis that WGS of our patients' tumor and germline genomes is a new test that will revolutionize the future of cancer medicine is partially correct. To quote Tony Blair after the first draft of the human genome in 2000 came out:

"Ever so often in the history of human endeavor, there comes a breakthrough that takes human kind across a frontier into a new era. Today's announcement is such a breakthrough, a breakthrough that open the way for massive advancement in the treatment of cancer and hereditary diseases. And this is only the beginning." - Tony Blair

This was 13 years ago. Whether he was right or wrong is up for debate. And he gets a gold star for being positive. As Yogi Bera would say: **"Prediction is very hard, especially about the future." – Yogi Bera**

The identification of *IDH1* mutations in AML and GBM, *CREBBP* mutations in many different tumors and *CREBBP* as a potential *BRCA1/2* mutation modifier are a few among many discoveries that suggest that the null hypothesis that this new WGS test will *not* revolutionize cancer medicine is incorrect. However, it is apparent that all of the approved therapies that are designer drugs resulted from targets identified with low throughput hypothesis driven studies. And it should be noted these drugs are not perfect and few cures result[14,15]. The bottle neck that is preventing this WGS test from realizing its potential quickly is a need to find a way to comprehend what these rare variations are doing to the physiology and pathophysiology of human cells. Solving this problem will necessarily take us back to model organisms (eg mus musculus) and the lab bench. Basic research will point the way. "Pointing and clicking" will not be sufficient. We should definitely use this WGS test in research especially for cancer patients with strong family histories of cancer. But we should use it in combination with solid, creative *in vitro* and *in vivo* investigations to comprehend what our genomes mean.

Future: The Naysayers in the past about the clinical use of the *BRCA1/2* discoveries were wrong. The Naysayers in the past about J. Craig Venter's sequencing technique being impractical and inaccurate, were wrong. The *BRCA1/2* story may be considered baby steps in our search for useful knowledge of how genes go bad and what to do about it. Such knowledge is a power that can be put to work to prevent cancer in the future. It has taken a long time for the clinical world to accept *BRCA1/2* testing. Even 5 years after Rebbeck et al's classic studies defining the effectiveness of prophylactic surgeries[38,39], there were too few physicians who recommended women with early onset breast cancers to have *BRCA* testing. The hope is that WGS will enhance our ability to determine how to prevent cancer before we have to treat it and that we will implement in the clinic sooner, rather than later.

References

1 Gonzaga-Jauregui, C., Lupski, J. R. & Gibbs, R. A. Human genome sequencing in health and disease. *Annu Rev Med* **63**, 35-61, doi:10.1146/annurev-med-051010-162644 (2012).

2 Jones, S. *et al.* Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* **324**, 217, doi:10.1126/science.11712021171202 [pii] (2009).

3 Lupski, J. R. *et al.* Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med* **362**, 1181-1191, doi:10.1056/NEJMoa0908094NEJMoa0908094 [pii] (2010).

4 Mardis, E. R. & Wilson, R. K. Cancer genome sequencing: a review. *Hum Mol Genet* **18**, R163-168, doi:ddp396 [pii]10.1093/hmg/ddp396 (2009).

5 Druker, B. J. *et al.* Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* **344**, 1031-1037, doi:10.1056/NEJM200104053441401 (2001).

6 Slamon, D. J. *et al.* Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* **344**, 783-792, doi:10.1056/NEJM200103153441101 (2001).

7 Kris, M. G. *et al.* Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. *JAMA* **290**, 2149-2158, doi:10.1001/jama.290.16.2149290/16/2149 [pii] (2003).

8 Miller, V. A. *et al.* Pilot trial of the epidermal growth factor receptor tyrosine kinase inhibitor gefitinib plus carboplatin and paclitaxel in patients with stage IIIB or IV non-small-cell lung cancer. *J Clin Oncol* **21**, 2094-2100, doi:10.1200/JCO.2003.12.008JCO.2003.12.008 [pii] (2003).

9 Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* **364**, 2507-2516, doi:10.1056/NEJMoa1103782 (2011).

10 Rowley, J. D. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **243**, 290-293 (1973).

11 King, C. R., Kraus, M. H. & Aaronson, S. A. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. *Science* **229**, 974-976 (1985).

12 Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* **350**, 2129-2139, doi:10.1056/NEJMoa040938NEJMoa040938 [pii] (2004).

13 Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949-954, doi:10.1038/nature00766nature00766 [pii] (2002).

14 Savona, M. & Talpaz, M. Getting to the stem of chronic myeloid leukaemia. *Nat Rev Cancer* **8**, 341-350 (2008).

15 Flaherty, K. T. *et al.* Inhibition of mutated, activated BRAF in metastatic melanoma. *N Engl J Med* **363**, 809-819, doi:10.1056/NEJMoa1002011 (2010).

16 Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* **483**, 100-103, doi:10.1038/nature10868nature10868 [pii] (2012).

17 Ley, T. J. *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **456**, 66-72, doi:10.1038/nature07485nature07485 [pii] (2008).

18 Mardis, E. R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* **361**, 1058-1066, doi:10.1056/NEJMoa0903840NEJMoa0903840 [pii] (2009).

19 Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809-813, doi:10.1038/nature08489nature08489 [pii] (2009).

20 Pleasance, E. D. *et al.* A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**, 191-196, doi:10.1038/nature08658nature08658 [pii] (2010).

21 Pleasance, E. D. *et al.* A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184-190, doi:10.1038/nature08629nature08629 [pii] (2010).

22 Parsons, D. W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807-1812, doi:10.1126/science.11643821164382 [pii] (2008).

23 Huang, Y. & Gottardo, R. Comparability and reproducibility of biomedical data. *Brief Bioinform*, doi:bbs078 [pii]10.1093/bib/bbs078 (2012).

24 Ioannidis, J. P. Why most published research findings are false. *PLoS Med* **2**, e124, doi:04-PLME-E-0321R2 [pii]10.1371/journal.pmed.0020124 (2005).

25 Dahlman, K. B. *et al.* BRAF(L597) mutations in melanoma are associated with sensitivity to MEK inhibitors. *Cancer Discov* **2**, 791-797, doi:2159-8290.CD-12-0097 [pii]10.1158/2159-8290.CD-12-0097 (2012).

26 Mullighan, C. G. *et al.* CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* **471**, 235-239, doi:nature09727 [pii]10.1038/nature09727 (2011).

27 Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189-195, doi:10.1038/nature09730nature09730 [pii] (2011).

28 Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303, doi:10.1038/nature10351nature10351 [pii] (2011).

29 Gui, Y. *et al.* Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nat Genet* **43**, 875-878, doi:10.1038/ng.907ng.907 [pii] (2011).

30 Peifer, M. *et al.* Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet* **44**, 1104-1110, doi:10.1038/ng.2396ng.2396 [pii] (2012).

31 Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev* **14**, 1553-1577 (2000).

32 Kwak, E. L. *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* **363**, 1693-1703, doi:10.1056/NEJMoa1006448 (2010).

33 Lichtenstein, P. *et al.* Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* **343**, 78-85, doi:10.1056/NEJM200007133430201 (2000).

34 Roberts, N. J. *et al.* The predictive capacity of personal genome sequencing. *Sci Transl Med* **4**, 133ra158, doi:scitranslmed.3003380 [pii]10.1126/scitranslmed.3003380 (2012).

35 Futreal, P. A. *et al.* BRCA1 mutations in primary breast and ovarian carcinomas. *Science* **266**, 120-122 (1994).

36 Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).

37 Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789-792, doi:10.1038/378789a0 (1995).

38 Rebbeck, T. R. *et al.* Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. *J Clin Oncol* **22**, 1055-1062 (2004).

39 Rebbeck, T. R. *et al.* Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. *N Engl J Med* **346**, 1616-1622 (2002).

40 Antoniou, A. C. *et al.* Breast and ovarian cancer risks to carriers of the BRCA1 5382insC and 185delAG and BRCA2 6174delT mutations: a combined analysis of 22 population based studies. *J Med Genet* **42**, 602-603 (2005).

41 Panchal, S. *et al.* Does family history predict the age at onset of new breast cancers in BRCA1 and BRCA2 mutation-positive families? *Clin Genet* **77**, 273-279, doi:10.1111/j.1399-0004.2009.01328.xCGE1328 [pii] (2010).

42 Mulligan, A. M. *et al.* Common breast cancer susceptibility alleles are associated with tumour subtypes in BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2. *Breast Cancer Res* **13**, R110, doi:bcr3052 [pii]10.1186/bcr3052 (2011).

43 Ross, T. S. Grace Rock's consent. *Endocr Relat Cancer* **19**, C29-31, doi:ERC-12-0186 [pii]10.1530/ERC-12-0186 (2012).

44 Lopez-Atalaya, J. P. *et al.* Histone acetylation deficits in lymphoblastoid cell lines from patients with Rubinstein-Taybi syndrome. *J Med Genet* **49**, 66-74, doi:10.1136/jmedgenet-2011-100354jmedgenet-2011-100354 [pii] (2012).

45 Abecasis, G. R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632nature11632 [pii] (2012).

46 Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618-630, doi:S1934-5909(10)00434-0 [pii]10.1016/j.stem.2010.08.012 (2010).

47 Riggs, J. W. *et al.* Induced pluripotency and oncogenic transformation are related processes. *Stem Cells Dev*, doi:10.1089/scd.2012.0375 (2012).

48 Lee, J. H. *et al.* Stem-cell protein Piwil2 is widely expressed in tumors and inhibits apoptosis through activation of Stat3/Bcl-XL pathway. *Hum Mol Genet* **15**, 201-211, doi:ddi430 [pii]10.1093/hmg/ddi430 (2006).

49 Lee, J. H. *et al.* Pathways of proliferation and antiapoptosis driven in breast cancer stem cells by stem cell protein piwil2. *Cancer Res* **70**, 4569-4579, doi:0008-5472.CAN-09-2670 [pii]10.1158/0008-5472.CAN-09-2670 (2010).

50 Suzuki, R., Honda, S. & Kirino, Y. PIWI Expression and Function in Cancer. *Front Genet* **3**, 204, doi:10.3389/fgene.2012.00204 (2012).

51 Pao, G. M., Janknecht, R., Ruffner, H., Hunter, T. & Verma, I. M. CBP/p300 interact with and function as transcriptional coactivators of BRCA1. *Proc Natl Acad Sci U S A* **97**, 1020-1025 (2000).

52 Gomez, M. R. *Tuberous sclerosis*. 3rd edn, (Oxford University Press, 1999).

53 Bartsch, O. *et al.* DNA sequencing of CREBBP demonstrates mutations in 56% of patients with Rubinstein-Taybi syndrome (RSTS) and in another patient with incomplete RSTS. *Hum Genet* **117**, 485-493, doi:10.1007/s00439-005-1331-y (2005).

54 Kishimoto, M. *et al.* Mutations and deletions of the CBP gene in human lung cancer. *Clin Cancer Res* **11**, 512-519, doi:11/2/512 [pii] (2005).

55 Miller, R. W. & Rubinstein, J. H. Tumors in Rubinstein-Taybi syndrome. *Am J Med Genet* **56**, 112-115, doi:10.1002/ajmg.1320560125 (1995).

56 Schorry, E. K. *et al.* Genotype-phenotype correlations in Rubinstein-Taybi syndrome. *Am J Med Genet A* **146A**, 2512-2519, doi:10.1002/ajmg.a.32424 (2008).

57 Roelfsema, J. H. & Peters, D. J. Rubinstein-Taybi syndrome: clinical and molecular overview. *Expert Rev Mol Med* **9**, 1-16, doi:S1462399407000415 [pii]10.1017/S1462399407000415 (2007).

58 Ward, R., Johnson, M., Shridhar, V., van Deursen, J. & Couch, F. J. CBP truncating mutations in ovarian cancer. *J Med Genet* **42**, 514-518, doi:42/6/514 [pii]10.1136/jmg.2004.025080 (2005).

59 Kung, A. L. *et al.* Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev* **14**, 272-277 (2000).

60 Cox, D. N. *et al.* A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev* **12**, 3715-3727 (1998).