

9. The Pan-Cancer Genome Atlas

David A. Wheeler, Jane H. Zhou

In 2012 with almost half the data in The Cancer Genome Atlas project collected, researchers from around the world launched an analysis that combined the data across 12 of the 25 TCGA cancers. This effort was dubbed the Pan-Cancer project and had two main goals: exploit the statistical power afforded by a collection of thousands of cancer patients to more sensitively interrogate the cancer genome, and to gain insight into the meaning of cancer subtypes by a comparative analysis of twelve adult cancers. The result has been captured in 22 papers published so far in topics relating to discovery of new cancer genes, broad classification of mutation spectra and their relationship to mechanisms of mutation in cancer, themes in common and shared intracellular signaling networks, and development of a range of internet tools. This extraordinary undertaking, which engendered participation of a few new US national and international groups, should be regarded as a pilot of an even larger effort that will involve all 10,000 TCGA patients as the project nears completion. We report on some of the most important successes of the project so far as a harbinger of more to come.

Introduction

The accumulation of cancer genome projects through The Cancer Genome Atlas (TCGA, ref: <https://tcga-data.nci.nih.gov/tcga/>) and the International Cancer Gene Consortium (ICGC, ref: <https://icgc.org/>) has resulted in a long list of adult solid and hematopoietic tumors, each involving hundreds of patients, each with a matched normal genome for comparison. The output of the major sequencing centers is collected in

the TCGA data repository, CGHub (<https://cghub.ucsc.edu/>), or the ICGC repository at EBI (<https://www.ebi.ac.uk/ega/>) and is accessible to researchers worldwide with interest and expertise in the study of cancer.

Each cancer was studied individually and the findings published in leading journals with impressive novel observations, recently reviewed in ¹⁻³). In 2012, the leadership of TCGA launched a new initiative, the Pan-Cancer Project, to analyze the large assemblage of cancer exomes at CGHub, 5,074 patients in all, spanning 12 cancers, as a single unit, with the goal of asking what molecular features are common to all cancers and which ones differ ⁴). In particular, could novel new cancer genes and pathways be brought to light through the statistical power of thousands of tumors

[Keywords]

somatic mutation, cancer genomics,
cancer subtypes

David A. Wheeler¹/Jane H. Zhou² : Human Genome Sequencing Center, and Department of Molecular and Human Genetics, Baylor College of Medicine¹ /Department of Pathology, MD Anderson Cancer Center²

combined in a single data set?

Upon the data freeze in December 2012, the consortium of analysts had amassed sufficient data from 12 cancer types (Figure 1), covered by, all or in part, six molecular platforms: whole exome (3,247), copy number (4,932), mRNA (4,080) and miRNA (4,628) expression, DNA methylation (4,855) and reverse phase protein arrays (2,674).

The Pan-Cancer project, led by J. Stuart, C. Sander and I. Shmulevich was a daunting undertaking since there existed no infrastructure to support the quality control of raw data, setting data freezes for samples across a dozen cancer types, version-control of summary data files representing individual platforms, coordination of the activities of sub-groups for this project. The team selected the web repository Synapse, from Sage Bionetworks⁵⁾, as an organizing principle for project coordination and data tracking functions. The Synapse software platform was designed to support key elements of large multi-center cooperative undertakings such as this.

As of this review, 22 papers have been published with many more in progress. The current papers, published mainly in the fall of 2013, bridge 7 topic areas (Table 1). The first three topics aggregate projects dealing with various aspects of somatic mutation, the impact of mutations in single genes or gene families, and mutational mechanisms. They consider both point mutations and copy number alterations (CNA). In these three topics, different research groups looked at the organ specificities of mutated genes and explored new techniques for recognizing significant mutation patterns that can only be discerned using very large data sets. Three groups drew insights relating to non-coding RNAs for the fourth topic and one study explored the presence and functional relationship of viral pathogen sequences in tumors for the fifth topic. The sixth topic dealt with approaches that defined novel networks of genes and pathways, and their use in stratifying tumors either biologically or for therapeutic intervention. Finally, in the seventh topic, four groups present internet resources and tools to enable other researchers to mine the data for answers to their own questions.

Here we review the highlights of the TCGA Pan-Cancer project, and critique and discuss the future prospects for the “globalization” of analysis of the TCGA and ICGC cancer genomics data sets.

1 Pan-Cancer observations

1) Somatic mutations landscape

Most adult solid tumors average between 1 and 10 mutations per million bases of coding DNA although the range among individual patients within a tumor type varies by two orders of magnitude⁶⁾. How many of those mutations are contributing functionally to the cancer phenotype? Kandoth et al. suggested that most tumors harbor only 2~6 driver mutations⁷⁾. Lawrence et al. go on to refine the background mutation rate calculation, critical in assessing driver status of a given gene, by taking into account the effect of replication timing and transcriptional activity. Genes that are replicated early in the S-phase of the cell division cycle, or are expressed at high levels, have lower background mutation rates⁶⁾. This approach led them to conclude that several reported genes such as CSMD1, CSMD3 and LRP1B appear to have been incorrectly identified as cancer drivers. Tamborero et al.⁸⁾ discern non-random distributions of amino acid substitution within each gene to augment the methods based solely on mutation rate computation. They identified three patterns of mutation indicative of positive selection: i. Functional bias, implied by excess nonsense frame-shift and splice site mutations; ii. Clustered missense mutations, the classical “oncogene” pattern seen in the recurrent mutation sites of KRAS or IDH1; iii. Enrichment of mutations affecting phosphorylation sites. From this approach, they find over 100 new candidate cancer driver genes.

Tumors driven by more than two cancer genes afford an opportunity to assess patterns of co-occurrence or mutual exclusivity, revealing information about signaling pathways in operation in the given cancer. In the majority of cancers in the study, mutation of TP53 is mutually exclusive with some other cancer driver⁷⁾. On the other hand, every cancer has one or more genes that are mutually exclusive uniquely in a given cancer (Figure 2).

Analysis of CNA of almost 5,000 primary cancers across 11 cancer types revealed 140 regions of recurrent amplification or deletion. Most surprising is that the majority of the regions, 102, had no known oncogene or tumor suppressor⁹⁾, indicating that there is much more to be learned about the targets of selection in affected tumors.

2) Somatic mutation mechanism

Three projects lead to new insights into mutational

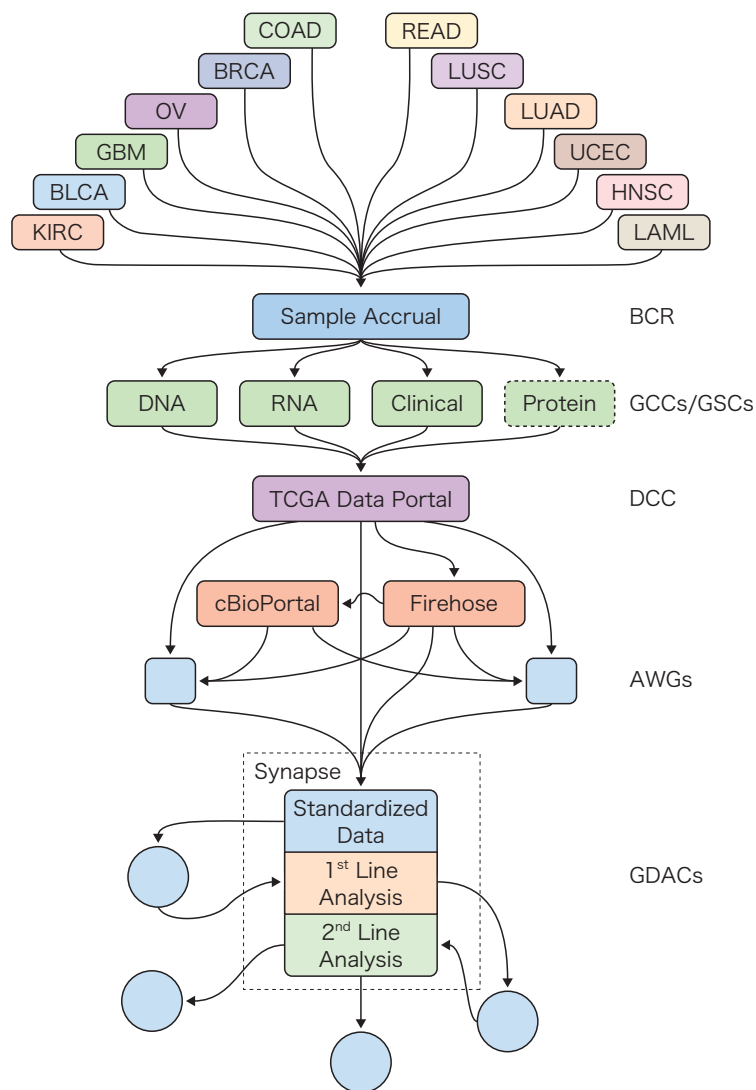


Figure 1 Organization and data flow in the TCGA Pan-Cancer project

Synapse, dashed line box, which lay outside of the TCGA project, was a key enabling technology for version control and coordination of the analyses. TCGA Data Portal is at <https://tcga-data.nci.nih.gov/tcga/>, the cBioPortal at <http://www.cbioportal.org/public-portal/> and Firehose at <https://confluence.broadinstitute.org/display/GDAC/Home>. BCR, the Biospecimen Core resource, prepares all analytes, DNA, RNA and protein, for distribution to TCGA data-generating centers; AWG, the TCGA disease-specific Analysis Working Groups [Reproduced from (4)]

mechanisms. Two of those delved into the prevalence of the mutation signature produced by the cytidine deaminase, APOBEC3B. This gene can convert cytosine to uracil, and is thought to play a role in RNA editing in normal cells but in some cancers it also mutates DNA, producing specific patterns of nucleotide substitutions at C residues. Cancers of bladder, cervix, lung, head and neck, and breast ^{10,11}, all show

significant contribution to the overall mutation profile from APOBEC-associated DNA damage.

The third study examined the relationship between point mutation (M-class) and CNA (C-class) and revealed an intriguing inverse relationship between these two types of mutation when data from all cancers were combined ¹². Ciriello et al. stratified these classes into disease-specific signatures, which could be

Table 1

Theme	Ref	Title	Lead Institution	Journal	Pubmed ID
PanCancer Overview	4	The Cancer Genome Atlas Pan-Cancer analysis project	TCGA	Nat Genet	24071849
Mutation Landscape	9	Pan-cancer patterns of somatic copy number alteration	Broad	Nat Genet	24071852
	7	Mutational landscape and significance across 12 major cancer types	WashU	Nature	24132290
	8	Comprehensive identification of mutational cancer driver genes across 12 tumor types	UPF, Toronto, WashU	Sci Rep	24084849
	6	Mutational heterogeneity in cancer and the search for new cancer-associated genes	Broad	Nature	23770567
Mutation Impact	13	Effects of TP53 mutational status on gene expression patterns across 10 human cancer types	BCM	J Pathol	24374933
	15	The mutational landscape of phosphorylation signaling in cancer	U Toronto	Sci Rep	24089029
	14	A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events	Broad	PLoS One	24498085
Mutation mechanism	10	Evidence for APOBEC3B mutagenesis in multiple human cancers	U Minnesota	Nat Genet	23852168
	11	An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers	NIEHS, Broad	Nat Genet	23852170
	12	Emerging landscape of oncogenic signatures across human cancers	MSKCC	Nat Genet	24071851
ncRNA	20	Identification of a pan-cancer oncogenic microRNA superfamily anchored by a central core seed motif	BCM	Nat Commun	24220575
	21	Analysis of microRNA-target interactions across diverse cancer types	MSKCC	Nat Struct Mol Biol	24096364
	22	The expression level of small non-coding RNAs derived from the first exon of protein-coding genes is predictive of cancer status	BCCR	EMBO Rep	24534129
Pathogens	19	The landscape of viral expression and host gene fusion and adaptation in human cancer	U Gothenburg	Nat Commun	24085110
Networks & Subtypes	16	Network-based stratification of tumor mutations	UCSD	Nat Struct Mol Biol	24037242
	17	Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types	MDACC	Nat Commun	24488081
		A pan-cancer proteomic perspective on The Cancer Genome Atlas	MDACC	Nat Commun	24871328 ^{a)}
	18	Inferring tumour purity and stromal and immune cell admixture from expression data	MDACC	Nat Commun	24113773
Tools and Portals	25	TCPA : a resource for cancer functional proteomics data	MDACC	Nat Methods	24037243
	23	Exploring TCGA Pan-Cancer Data at the UCSC Cancer Genomics Browser	UCSC	Nat Methods	24084870
	24	IntOGen-mutations identifies cancer drivers across tumor types	UPF	Nat Methods	24037244
	5	Enabling transparent and collaborative computational analysis of 12 tumor types within The Cancer Genome Atlas	Sage Bionet-works	Nat Genet	24071850

Institutions: Broad, Broad Institute of MIT and Harvard; BCM, Baylor College of Medicine; MSKCC, Memorial Sloan Kettering Cancer Center; MDACC, MD Anderson Cancer Center; UCSD, University of California, San Diego; UCSC, University of California, Santa Cruz; NIEHS, National Institute of Environmental Health Sciences; UPF, Pompeu Fabra University; Wash U, Washington University, St. Louis. a) PMID 24871328 has become available since submission of the manuscript and could not be discussed. It is included for completeness.

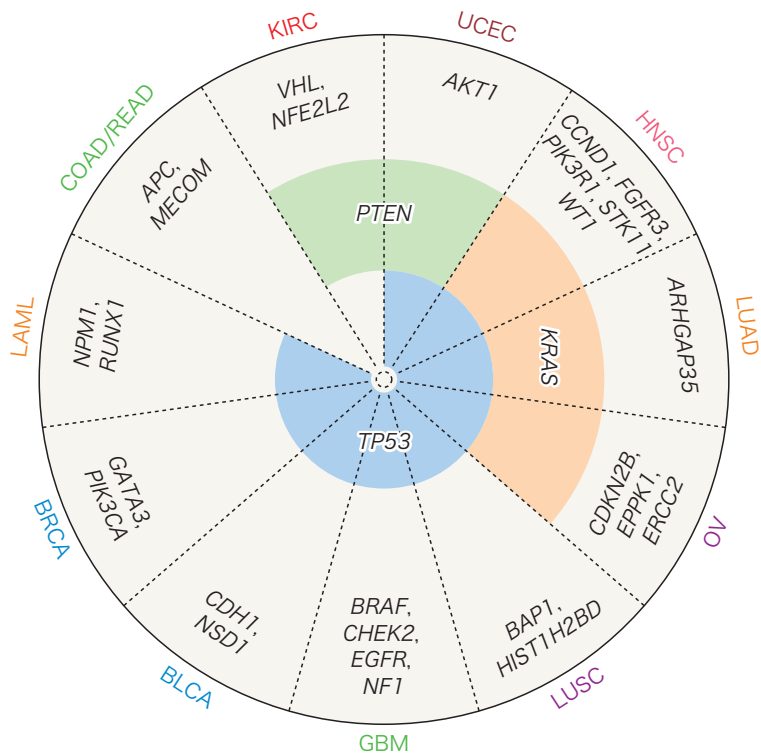


Figure 2 Mutual exclusivity of mutation within and between cancers

The program Dextrix⁷ was applied to all pairwise combinations of genes recurrently mutated in the 11 of the cancers under consideration in the Pan-Cancer project (labeled in outer ring) to identify patterns of mutual exclusivity in driver genes that co-occur in individual tumors. TP53 mutation is mutual exclusive with some other cancer gene in the majority, 9, of the cancers. KRAS and PTEN were also mutually exclusive with another gene in multiple cancers. All other genes were found to be mutually exclusive in only one cancer [Reproduced from (7)]

characterized by their content of therapeutically actionable targets¹².

3) Effects of mutated genes

Two studies focused on the consequences of mutation of single genes for the transcriptional program of the cell. TP53 is the most commonly mutated gene in human cancers, and is perhaps the most thoroughly studied cancer gene of all. In absence of wild type p53, upregulation of G2/M transcription factors, including FOXM1 and MYBL2, leads to activation of transcription programs that drive mitosis. Parikh et al. show that a high fraction of the genes composing the kinetochore are highly upregulated in the absence of wild type p53 in multiple human cancers, providing important mechanistic insight into the regulation of cell division by TP53¹³.

Brooks et al. ask what happens to the overall splicing capability of the cancer cell when the splicing

factor, U2AF1, is mutated. They observed 30 genes with significant alteration in splice pattern, including the well-known oncogenic transcription factor, beta catenin¹⁴. How the specificity arises to target only 30 these genes is still unclear, but this analysis revealed an intriguing hidden link between U2AF1 and WNT signaling.

Post-translational phosphorylation of proteins plays a central role in intracellular signaling. Reimand et al. use the enrichment of mutations at phosphorylation sites (i.e., pattern iii from (8), above) to analyze the relationships among mutated signaling proteins, and use this information to predict how signaling pathways may be rewired in cancer cells¹⁵.

4) Stratification of cancer by network analysis

The goal of systems biology is to identify prognostic and predictive features of the cancer cell using global attributes of multi-dimensional genomics data. The

Network-Based Stratification of cancer mutations is one of a class of methods that maps genomic data, in this case mutation data, onto known gene and pathway networks. Hofree et al. clustered patients with mutations in similar network regions that were predictive of tumor histology, response to therapy, and patient survival¹⁶. Co-expression networks were used by Yang et al to identify prognostic mRNA genes, which are groups of highly interconnected genes that show cross-cancer conservation¹⁷. Taken together, these studies suggest that cancers of different organ systems, but with the similar prognostic gene networks, may one day be treated using similar clinical approaches.

Yoshihara et al. stratified tumors by the fractions of infiltrating reactive stromal and immune cells within the tumor as inferred computationally from gene expression data¹⁸. Although the method developed for achieving this was designed for the purpose of estimating tumor purity, i.e. the fraction of cells in a tumor that are cancer cells, tissue composition plays an important role in cancer biology and clinical prognosis. Thus, there is great interest in computationally derived characterization of the cell mixtures in a tumor.

5) Pathogens in cancer

Viruses cause >10% of human cancer and in the first large-scale assessment of virus expression, Tang et al. analyze over 700 billion RNA Seq reads to quantify the presence of viral sequences and understand their relationship to the host. Besides the common association of HPV with cervical and oropharyngeal squamous cell carcinoma, and HBV and HCV with liver cancer, trace levels of viral RNAs from different species were seen in kidney clear cell, endometrial and lung cancers. Most interesting was the observation of highly expressed transcripts representing fusions between virus and host sequences, including recurrent fusions with cancer genes such as ERBB4, MLL4, PTV1 and RAD51B¹⁹.

6) Role of non-coding RNA

In spite of the known importance of miRNA to gene regulation at the transcriptional level, the relationships between a given miRNA and its target genes are poorly defined. Two of the miRNA studies use clever approaches to infer high-confidence sets of miRNA with their cognate mRNA. Hamilton and colleagues integrate the Pan-Cancer miRNA data with databases of miRNA binding sites defined by cross-linking

studies and find mutational evidence for disruption of the binding of specific miRNAs to key tumor suppressors in the PI3K, TGF β , and p53 pathways²⁰. Jacobsen et al., compared expression levels of miRNA and mRNA to infer miRNA-target relationships²¹. They found regulation of DNA demethylation pathway members TET1 and TDG. Both of these studies break new ground in our understanding of how miRNA directly impacts key intracellular signaling pathways.

Zovoilis et al. investigated the differential expression of small non-coding RNAs, a class of RNA associated with the transcription initiation start site of genes, and found that their abundance could discriminate cancer and normal tissues, and predict cancer status²². This is the first reported association between these enigmatic short non-coding RNAs and cancer and may eventually lead to the discovery of new avenues to understanding of tumorigenesis.

7) Tools and internet portals

The sum total of cancer genomic data amassed thus far occupies petabytes of data storage. Therefore tools and portals are essential for exploring and analyzing TCGA data. Reported for the Pan-Cancer Project include UCSC Cancer Genomics browser (<http://genome-cancer.ucsc.edu>), the IntOGen-mutations platform (<http://www.intogen.org/mutations/>), The Cancer Proteome Atlas (TCPA, <http://bioinformatics.mdanderson.org/main/TCPA:Overview>) and the aforementioned Synapse software platform.

The UCSC Cancer Genomics browser is an interactive visualization and exploration portal for browsing and analyzing gene expression, copy number, DNA methylation, and somatic mutation in the context of the genome browser²³. It includes the ability to layer onto genomic data, pathway information and clinical features to enable data mining in real time. It also provides Pan-Cancer subtype classifications and genomic biomarkers. The IntOGen-mutations web-site²⁴ provides a browser for the driver mutation data reported by the aforementioned work of Tamborero et al. providing ready access to frequencies, functional impact and spatial clustering tendencies of mutations by patient and by cancer. And finally, The Cancer Proteome Atlas (TCGP) tracks the abundance of 200 proteins and phosphoproteins, measured by antibody staining in reverse phase protein arrays, and allows users to correlate pair-wise protein levels across the entire data set; evaluate differential expression levels between different groups of samples, and link these

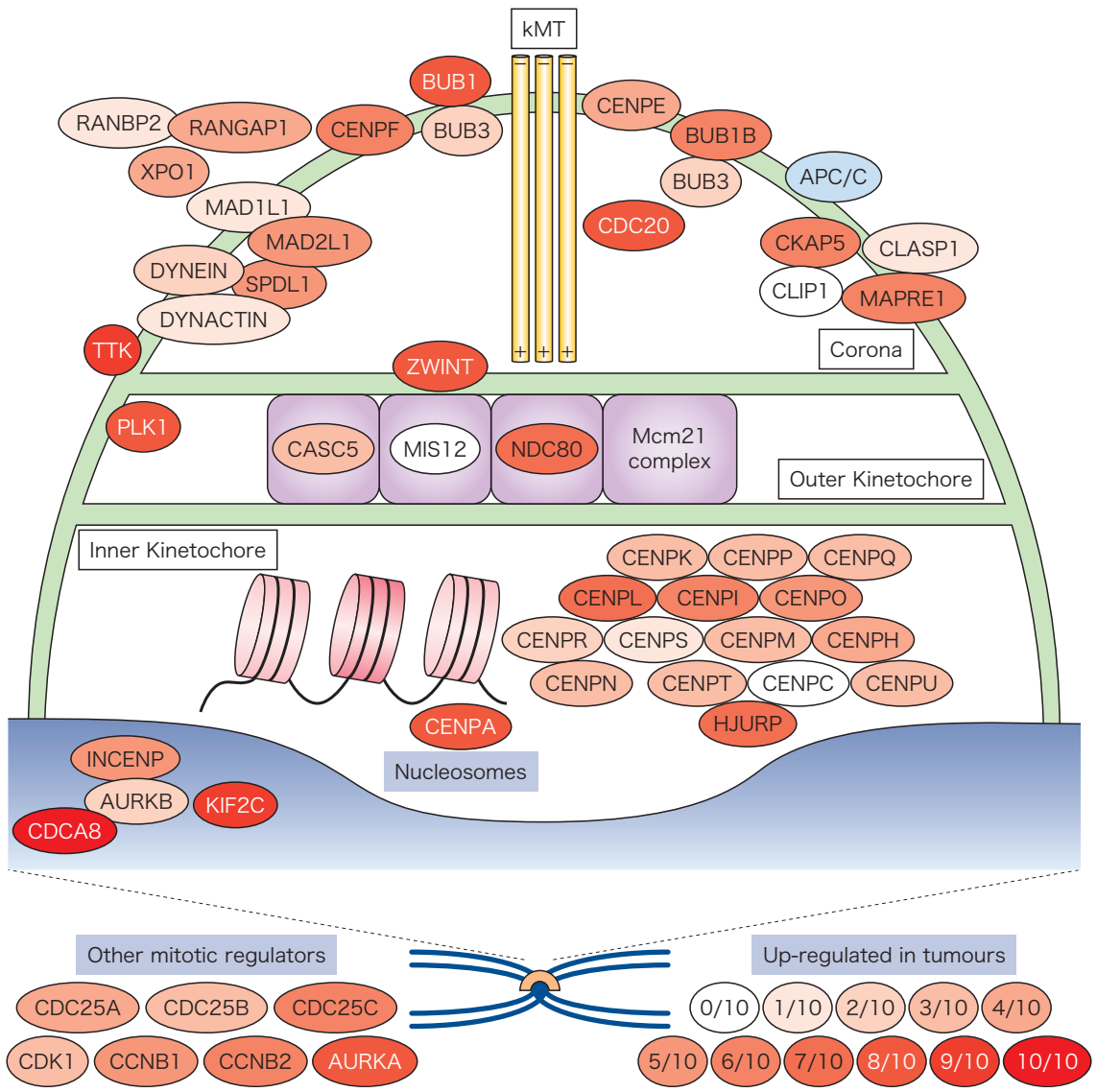


Figure 3 Components of the kinetochore up-regulated by loss of wild-type TP53 in cancer
 Many of the ten human cancers with mutant TP53 alleles exhibited dramatic upregulation of cell cycle regulatory genes, particularly mitosis promoting genes and genes that make up the kinetochore, the junction structure where centrosome-directed spindle fibers attach to the mitotic chromatids. These results are consistent with literature reports that p53 directly suppresses genes that drive the G2/M cell cycle transition, such as cyclin B1 and FOXM1. This association of TP53 mutation and G2/M gene upregulation had been noted previously in individual cancers such as breast and ovarian cancers, but the widespread nature of this phenomenon became more evident in the study by Parikh et al. [Reproduced from (13)]

results to patient survival²⁵. Together, these internet portals form an extraordinary resource for the investigation of cancer genes that should enhance the study of cancer for basic and clinical researchers alike.

2 Future directions

The goal of TCGA is to collect multi-platform genomic data on 10,000 cancer patients. The project is stratified into more than 25 cancers, the majority of

which include at least 500 patients. The reason for the deep population analysis is to minimize disease heterogeneity. Indeed, breast cancer is so well characterized into distinct sub-types that TCGA took the unusual step of extending that cohort to over 1,000 patients. The Pan-Cancer seeks to pool all the patients to ask two fundamental questions: what common features among all cancers can be discovered with the extraordinary statistical power of thousands of patients; what new features of each disease might emerge from multi-platform integrated analysis? With the current analysis we find ourselves at the about the halfway point on the road to completion and the findings, as reviewed here, are impressive. Based on the above results, we may expect still more novel genes and pathways to emerge. Lawrence et al. have argued forcefully, that for genes mutated at the 1~3% level we have only scratched the surface of the mutation space²⁶). One might argue that for genes mutated so infrequently, the clinical implications are small and time would be better spent in pursuit of therapeutic options for the more frequently mutated genes. However, the network analyses reported above suggest the opposite is true. The infrequently mutated genes are often participants in pathways that relate to targeted therapies, and might one day help stratify patients into therapeutic subgroups with much improved outcomes. As the TCGA data generation effort completes its final year, and the data for all 10,000 patients becomes available, Pan-Cancer analyses can only become more revealing.

The current analysis is not without its limitations, and future rounds of analysis will seek to further improve on this work. The current data set consists of mutation data derived from the three US National Sequencing Centers: the Broad Institute of MIT and Harvard, The Genome Institute of Washington University and the Human Genome Sequencing Center at Baylor College of Medicine. Subtle differences in DNA sequence alignment and somatic mutation calling have probably led to “batch effects” in the data [see for example (27~29)]. Mutation profiling in a Pan-Cancer context will become more accurate and sensitive after uniform alignment of all the cancer data with a single aligner, followed by recalling mutations using the greatly improved mutation callers that have emerged from five years of experience in cancer analysis. As with somatic mutation calling in whole exome data, wherein multiple mutation calling approaches

leads to more sensitive mutation calls, analysis of RNA Seq data will also benefit from the application of multiple algorithms, especially for detection of alternative splicing and gene fusion transcripts.

Both of these improvements will require efficient pipelines implemented in “the cloud” wherein the computing resource can be matched to the gigantic scale of the data. The ICGC in collaboration with TCGA has begun an analysis of comparable size with analysis of 2,000 whole genome sequence patients. Though fewer in number, a matched tumor-normal whole genome sequence data set is roughly the size of 10 whole exomes. This extraordinary experiment is exploiting the resources of the cloud already and should provide valuable experience for the next round of Pan-Cancer TCGA analysis.

The use Synapse for data aggregation and versioning has afforded valuable experience in support of the coordination of large-scale multi-institutional project. And whereas the first round of analyses was predominated by researchers in, or affiliated with TCGA analysis centers, there were several groups participating from the outside (Table 1). As more national and international researchers become aware of the opportunity this vast data set affords, participation will expand, and with that, new ideas and approaches will lead to new insights into the disease. Thus, the Pan-Cancer analysis of TCGA data is a living process, inviting participation worldwide to seek better understanding for clinicians and better outcomes for patients.

Acknowledgments

The authors thank Joshua Stuart for helpful suggestions on this manuscript. JHZ is supported by NIH grant T32CA163185.

References

- 1) Garraway LA & Lander ES : Cell, 153 : 17–37, 2013
- 2) Vogelstein B, et al : Science, 339 : 1546–1558, 2013
- 3) Wheeler DA & Wang L : Genome Res, 23 : 1054–1062, 2013
- 4) Weinstein JN, et al : Nat Genet, 45 : 1113–1120, 2013
- 5) Omberg L, et al : Nat Genet, 45 : 1121–1126, 2013
- 6) Lawrence MS, et al : Nature, 499 : 214–218, 2013
- 7) Kandath C, et al : Nature, 502 : 333–339, 2013
- 8) Tamborero D, et al : Sci Rep, 3 : 2650, 2013
- 9) Zack TI, et al : Nat Genet, 45 : 1134–1140, 2013
- 10) Burns MB, et al : Nat Genet, 45 : 977–983, 2013

- 11) Roberts SA, et al : Nat Genet, 45 : 970-976, 2013
- 12) Ciriello G, et al : Nat Genet, 45 : 1127-1133, 2013
- 13) Parikh N, et al : J Pathol, 232 : 522-533, 2014
- 14) Brooks AN, et al : PLoS One, 9 : e87361, 2014
- 15) Reimand J, et al : Sci Rep, 3 : 2651, 2013
- 16) Hofree M, et al : Nat Methods, 10 : 1108-1115, 2013
- 17) Yang Y, et al : Nat Commun, 5 : 3231, 2014
- 18) Yoshihara K, et al : Nat Commun, 4 : 2612, 2013
- 19) Tang KW, et al : Nat Commun, 4 : 2513, 2013
- 20) Hamilton MP, et al : Nat Commun, 4 : 2730, 2013
- 21) Jacobsen A, et al : Nat Struct Mol Biol, 20 : 1325-1332, 2013
- 22) Zovoilis A, et al : EMBO Rep, 15 : 402-410, 2014
- 23) Cline MS, et al : Sci Rep, 3 : 2652, 2013
- 24) Gonzalez-Perez A, et al : Nat Methods, 10 : 1081-1082, 2013
- 25) Li J, et al : Nat Methods, 10 : 1046-1047, 2013
- 26) Lawrence MS, et al : Nature, 505 : 495-501, 2014
- 27) Cancer Genome Atlas Research Network : Nature, 487 : 330-337, 2012
- 28) Cancer Genome Atlas Research Network : Nature, 490 : 61-70, 2012
- 29) Cancer Genome Atlas Research Network : Nature, 499 : 43-49, 2013

David A. Wheeler : Dr. Wheeler received Bachelor of Science degrees in Biochemistry and Zoology from the University of Maryland and a Master of Science in Biochemistry from the George Washington University. His postgraduate study at, George Washington University, included the study of immunochemistry of RNA tumor virus reverse transcriptases, under Dr. Robert Gallo; and RNA tumor viruses transcription under Dr.

Gordon Hager. Dr. Wheeler was the first to demonstrate transcription of the long open reading frame gene in the LTR of endogenous mouse mammary tumor virus and the effects of CpG methylation on its transcription. Subsequently Dr. Wheeler did postdoctoral research in behavioral genetics at Brandeis University under Dr. Jeffrey Hall. Dr. Wheeler participated in the cloning of the period locus, the first gene with an established role in regulating behavior (circadian rhythms) in any organism. Through this work in the late 1980s, Dr. Wheeler became interested in the new area of computational biology. He joined the faculty at Baylor College of Medicine in 1991 to develop computational tools for molecular biology. He was Director the Molecular Biology Computation Resource at Baylor College of Medicine for 10 years and in 2001 joined the Human Genome Sequencing Center at BCM where he guided the finishing of the *D. melanogaster* chromosome 3 and X genome sequence followed by the human genome sequence, chromosomes 3, and 12. Since the completion of the human genome he has developed methods for discovery of genome variation in human and animal populations using DNA sequencing technologies with the goal of relating polymorphism to human disease. In 2008 he led the analysis of the Watson genome, the first human genome sequenced using next generation sequencing technology. Over the past 4 years, his work on genome variation has been applied to somatic mutation analysis in cancer.

本総説の日本語訳は、下記書籍にてご覧いただけます。



実験医学増刊 Vol.32 No.12

個別化医療を拓く がんゲノム研究

一解き明かされるがんの本質と分子診断・治療応用への展開

柴田龍弘／編

B5判 231ページ ISBN978-4-7581-0340-4

<http://www.yodosha.co.jp/jikkenigaku/book/9784758103404/>

【発行元】

株式会社 羊 土 社

〒101-0052

東京都千代田区神田小川町2-5-1

TEL 03(5282)1211(代表)

FAX 03(5282)1212

E-mail eigy@yodosha.co.jp

URL <http://www.yodosha.co.jp/>

© YODOSHA CO., LTD. 2014

本誌に掲載する著作物の複製権・上映権・譲渡権・公衆送信権(送信可能化権を含む)は(株)羊土社が保有します。本誌を無断で複製する行為(コピー、スキャン、デジタルデータ化など)は、著作権法上での限られた例外(「私的使用のための複製」など)を除き禁じられています。研究活動、診療を含み業務上使用する目的で上記の行為を行うことは大学、病院、企業などにおける内部的な利用であっても、私的使用には該当せず、違法です。また私的使用のためであっても、代行業者等の第三者に依頼して上記の行為を行うことは違法となります。