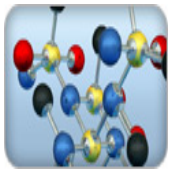
NATIONAL CANCER INSTITUTE
Office of Cancer Genomics

Published on *Office of Cancer Genomics* (https://ocg.cancer.gov)



# Issue 2 : July, 2010



## DIRECTOR'S NOTE

### Director's Note

On July 13, 2010, President Barack Obama issued a letter on a National HIV/AIDS strategyOpens in a New Tab. He wrote: “Thirty years ago, the first cases of human immunodeficiency virus (HIV) garnered the world’s attention.

## GUEST EDITORIAL

### Linking Tumor Genomic Features to Cancer Therapeutics

Stuart L. Schreiber, Ph.D., a collaborator for several NCI initiatives, has developed systematic ways to explore the field of chemical biology by using small molecules. Most recently Dr. Schreiber was awarded the Fourth Annual AACR Award for Outstanding Achievement in Chemistry in Cancer Research.

## TARGET PROGRAM HIGHLIGHTS

### TARGET Investigator Highlights the Importance of Translating Cancer Research to Patient Care

Cheryl Willman, M.D., is one of those clinical researchers who inspires cancer patients to hope for a cure. A two-time cancer survivor herself, Dr. Willman is passionate about changing the outlook for patients in her community by ensuring that everyone in New Mexico has world-class cancer care. Dr.

## NCI GENOMIC PROGRAM HIGHLIGHTS

### TCGA Researchers Discover Novel Molecular Subtype in Patients with Distinct Clinical Features

A study conducted by The Cancer Genome AtlasOpens in a New Tab Research Network and published in Cancer Cell used epigenomic

profiling, which maps specific chemical changes or

## CONNECTING THE DATA

### OCG Data Portals: Opening the Doors to Cancer Genomics Discovery

Current lower-cost, high-throughput methods developed for genomic sequencing and genotyping have opened a literal floodgate for the generation of colossal amounts of genotypic information that further require advanced analyses, predominantly through bioinformatics means.

## FEATURED RESEARCHERS

### Perspectives from the Next Generation

If analyzing next [2nd] generation sequence data is like trying to drink from a fire hose, then, as a graduate student focused on the application of next [2nd] generation sequencing to cancer research, I have swallowed a lot of water and have gotten very wet over the last few years.

## PATIENT PERSPECTIVE

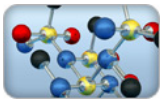
### What Do Patients Know About Cancer Genomics and Personalized Medicine?

“The goal of getting your genome done is not to tell you what you will die from . . . it’s how to take action to prevent disease.” – George M. Church, M.D., Ph.D. (in “On a Mission to Sequence the Genomes of 100,000 People” by David Duncan, 6/7/10, New York Times.)

---

# DIRECTOR'S NOTE

## Director's Note



On July 13, 2010, President Barack Obama issued a letter on a National HIV/AIDS strategyOpens in a New Tab. He wrote: “Thirty years ago, the first cases of human immunodeficiency virus (HIV) garnered the world’s attention. Since then, over 575,000 Americans have lost their lives to AIDS and more than 56,000 people in the United States become infected with HIV each year. Currently, there are more than 1.1 million Americans living with HIV. Moreover, almost half of all Americans know someone living with HIV. Our country is at a crossroads. Right now, we are experiencing a domestic epidemic that demands a renewed commitment, increased public attention, and leadership.”

With an increase in the number of Americans affected by AIDS and their associated

risk for certain types of cancers, the National Cancer Institute (NCI) Office of Cancer Genomics [2] and the NCI Office of HIV and AIDS MalignanciesOpens in a New Tab [3] initiated the HIV+ Tumor Molecular Characterization Project (H+TMCPOpens in a New Tab [4]) with the goal of understanding the effect of HIV-infection on the development of certain cancers.

The advent of highly active anti-retroviral therapy (HAART) has considerably slowed disease progression from HIV to full-blown AIDS, thereby increasing the number of people living with HIV. It is not known why the incidence of certain cancers, but not others, increases in patients with HIV infection. Among the cancers with elevated prevalence is aggressive B-cell Non-Hodgkin lymphoma (NHL) and late-stage lung cancer. Even though some cancers have a viral origin, and immunodeficiency could more readily lead to the development of tumors caused by viruses, many questions remain.

**Opportunity:**

Second- and third-generation sequencing technologies and informatics tools allow nearly complete characterization of tumor transcriptomes and tumor genomes, together with identification of sequence alterations. Initial results support the concept that many of the chromosome, transcriptome and epigenome changes in cancer cells are not directly related to cancer development. Therefore, different analytical and experimental approaches need to be used to distinguish between “driver” and “passenger” alterations, including the characterization of similar tumors that arise in the context of the complex biological setting.

NCI already supports tumor genome characterization initiatives for both NHL and lung malignancies occurring in individuals with competent immune systems (immunocompetent) through the Cancer Genome Characterization InitiativeOpens in a New Tab [5] and The Cancer Genome AtlasOpens in a New Tab [6] program. Therefore, by using systems biology analytical methods in H+TMCP to compare cancers with identical histologic subtypes that develop in both immunocompetent *and* immunodeficient people, we should be able to identify the mechanisms that either:

- allow tumors to counteract one’s innate cancer immunity (immune surveillance), *or*
- become altered only in the presence of external cancer-causing factors, such as oncogenic viruses

Specifically, H+TMCP will identify the potentially causative mutations and their associated pathways occurring in HIV-associated cancers which may respond to new or existing treatments. H+TMCP is just one example of how OCG integrates various trans-NCI projects and empowers the research community with the goal of developing better treatments for patients with cancer. All data generated will be made available to the scientific community for further analysis and use.

Sincerely,

Daniela S. Gerhard, Ph.D.
Director, NCI Office of Cancer Genomics

Back to Top

# GUEST EDITORIAL

## Linking Tumor Genomic Features to Cancer Therapeutics

**Stuart L. Schreiber, Ph.D.**

Stuart L. Schreiber, Ph.D., a collaborator for several NCI initiatives, has developed systematic ways to explore the field of chemical biology by using small molecules. Most recently Dr. Schreiber was awarded the *Fourth Annual AACR Award for Outstanding Achievement in Chemistry in Cancer Research*. Schreiber's efforts have led to the development of two new anti-cancer drugs and have transformed the way in which researchers develop targeted cancer therapeutics. In this guest editorial, Dr. Schreiber elaborates on small-molecule science, personalized medicine and what is needed to reduce the cancer burden on patients.

The ability to understand cancer genomes and the advances in small-molecule science provide a radically new foundation for creating the medicines we've only imagined since declaring the war on cancer decades earlier — the ones needed to take out this disease. We've learned the power of linking genetic features of cancers to drug efficacies — and that the extraordinary consequences of exemplars like imatinib/Gleevec are not restricted to this drug and its genetically matched leukemia, chronic myelogenous leukemia (CML). Recent studies show, for example, high response rates with genetically matched drugs targeting extremely challenging cancers such as melanoma. These advances are encouraging, but they still only affect less than 1% of patients suffering today from cancer. And the most dramatic cases are ones rationalized during or even after the discovery of the actual small-molecule therapeutic (e.g., the compound later named imatinib). So where do we go from here; how do we exploit our new foundation and insights *comprehensively* and *prospectively* so that all cancer patients can benefit?

I would like to review one simple idea — an idea I find attractive since it addresses the challenge comprehensively and it is on a direct path to cancer patients, even in cases where we have not yet gained all the insights into the disease that we seek. In fact, it's an idea that starts with patients.

The NCI's The Cancer Genome Atlas (TCGA)Opens in a New Tab [6], Therapeutically Applicable Research to Generate Effective Treatments (TARGET)Opens in a New Tab [7] and Cancer Genome Characterization Initiative (CGCI)Opens in a New Tab [5] programs are revealing genetic (copy number variants; mutations; expression) and epigenetic (DNA methylation) features of patient-derived cancers from a wide range of tissue origins (collectively labeled 'genomic features'). But to bring benefits to patients, we need: 1) models of cancers amenable to high-throughput science; 2) highly specific probes of nodal points in cancer circuitry; and 3) computational methods that correlate cancer genetic feature/drug efficacy relationships. All three are now available.

Although the tissues that were characterized are not amenable to testing for drug efficacy, many hundreds of cancer cell lines are being similarly characterized, and recent advances in high-throughput cell culturing and small-molecule screening now permit such testing using these cell lines. Studies in the past year suggest that, although not perfect models of cancers, cancer cell lines have great value as predictors of drug efficacies and resistance mechanisms. But this alone will only permit a better understanding of the patients best suited for the already-approved drugs. Although drugs do not exist for the vast number of candidate targets in cancer, small-molecule probes ('tool compounds') of these targets either exist or can be obtained using national resources such as the Molecular Libraries Probe Production Centers Network [8] (MLPCN)Opens in a New Tab. Lastly, computational methods have been developed that permit unbiased analyses of the genomic features of the cancer cell lines that correlate with the ability of the drug or probe to arrest growth or to induce death of the cells. Pilot experiments, including the NCI's Cancer Target Discovery and Development ($CTD^2$) NetworkOpens in a New Tab [9], prove that correlations of cancer genetic features and drug efficacy can be established that parallel the dramatic consequences seen with the small number of cancer patients currently taking genetically matched drugs.

Relating the genetic features of cancers to drug efficacy *comprehensively* will require a coordinated national effort, the type at which the NCI has proven adept. Examining a breadth of genomically characterized cancer cell lines and a comprehensive collection of highly specific small-molecule probes and drugs could yield an Internet-accessible look-up table that points to a drug or drug precursor that is tailored to the genetic features of distinct cancers, as well as to combinations of drugs that target both efficacy and resistance mechanisms. While many cases will point to a probe rather than a drug, the information will make clear what drugs need to be developed and their path to development will be facilitated by the publicly accessible information concerning probe development. A goal of this undertaking is to facilitate the *prospective* discovery of drugs that are genetically matched to patients and that yield high rates of durable responses. Here, the same genetic features used to create drug or probe efficacy correlations point to biomarkers for development and optimal patient populations for efficient clinical investigation.

We should aim to have the Food and Drug Administration's (FDA) approval of imatinib, which did not entail costly and time-consuming phase-3 clinical trials due to a 90% response rate in phase-2 trials, be the standard to which future cancer drug development efforts aspire. The overall cost of the clinical development of next-generation cancer drugs could easily be diminished substantially by greater response rates of selected patient populations, the ability to identify genomic-based biomarkers, and the ability to detect resistance mechanisms in cell-line models of cancers and therefore to select rationally drug combinations that yield durable responses.

This research path begins and ends with cancer patients. It is empirical in nature, driven by data and correlations, and does not necessarily require the deep understanding of cancer that is a laudable aspiration of the cancer research community (although every bit of understanding will nearly always be helpful). It is the most direct path of which I am aware to discovering the revolutionary cancer drugs that we imagine resulting from our investment in cancer research.

*The author declares conflicts of interest that are available online.* (Funded participant in CTD² and MLPCN; consultant to Forma Therapeutics and Ariad Pharmaceuticals.)

Back to Top

# TARGET PROGRAM HIGHLIGHTS

## TARGET Investigator Highlights the Importance of Translating Cancer Research to Patient Care

**Cheryl Willman, M.D. CEO University of New Mexico Cancer Center**

Cheryl Willman, M.D., is one of those clinical researchers who inspires cancer patients to hope for a cure. A two-time cancer survivor herself, Dr. Willman is passionate about changing the outlook for patients in her community by ensuring that everyone in New Mexico has world-class cancer care. Dr. Willman, now the Director and Chief Executive Officer (CEO) of the University of New Mexico Cancer Center, was one of the first physicians selected for the Physician-Scientist training program at the National Institutes of Health (NIH) — a program devoted to bringing more clinicians into cancer research. Now, almost three decades later, she has been consistently funded by the National Cancer Institute (NCI) to explore the biology and genetics of leukemia and lymphoma.

Dr. Willman leads the University of New Mexico team as part of the acute lymphoblastic leukemia (ALL)Opens in a New Tab [10] project for the NCI Therapeutically Applicable Research to Generate Effective Treatments (TARGET) [7] initiative. ALL is the most common cancer found in children. The TARGET High Risk (HR) ALL project is using comprehensive genomic approaches and high-throughput gene resequencing to identify new therapeutic targets for pediatric ALL. To Willman, TARGET represents precisely what has changed about her research over the past several decades. Instead of strictly concentrating on molecular mechanisms, her work is now clearly focused on how to translate research discoveries to patients.

Sobered by the reality that transferring clinical findings to the community hospital setting quickly or effectively, particularly in multi-cultural and more distant rural communities as in New Mexico and the American Southwest, is a big challenge, Dr. Willman has also become a staunch advocate for eliminating cancer health disparities. New Mexico represents a prime target for removing barriers in healthcare as a state with a “minority majority”, with a population comprised of more than 50% Hispanics and American Indians. Her work on ALL has been central to this fight. Recent research shows that two of the principal mutations identified in ALL (JAK2 and CRLF2) are found most often in children of Hispanic descent, particularly those with a significant degree of American Indian genetic ancestry. Indeed it seems that without including children of different ethnic backgrounds, these striking associations, which are already being used as the impetus for a clinical trial, might have never been made.

Dr. Willman is a true scientific leader and challenges all of us to think big. Speaking

to the next generation of scientists, Willman encourages future researchers to pick an important question without focusing on specific technologies or methods. That, she suggests, is the only way that those “Aha!” moments will come, leading to the clarity that changes science and, hopefully, lives.

Back to Top

# NCI GENOMIC PROGRAM HIGHLIGHTS

## TCGA Researchers Discover Novel Molecular Subtype in Patients with Distinct Clinical Features


A study conducted by The Cancer Genome AtlasOpens in a New Tab [6] Research Network and published in *Cancer Cell* used epigenomic profiling, which maps specific chemical changes or 'marks' to different areas of the genome, to reveal a new subtype of an aggressive form of brain cancer called Glioblastoma Multiforme (GBM). Most patients with GBM survive only 12-15 months after their initial diagnosis. However, patients with this specific subtype, called Glioma CpG Island Methylator Phenotype (G-CIMP), have a median survival of three years. Patients with this subtype are also younger at diagnosis with a median age of 36 compared to a median age of 56 for other GBM subtypes. While investigators are still determining which of these specific methylation 'marks' play a driving role in this survival difference, the results lay a foundation for the potential development of treatments targeted at this specific epigenomic profile. View Pubmed AbstractOpens in a New Tab [11].

Back to Top

# CONNECTING THE DATA

## OCG Data Portals: Opening the Doors to Cancer Genomics Discovery

Current lower-cost, high-throughput methods developed for genomic sequencing and genotyping have opened a literal floodgate for the generation of colossal amounts of genotypic information that further require advanced analyses, predominantly through bioinformatics means. The Inaugural Office of Cancer Genomics' *OCG e-News* described the caBIG® bioinformatics network of tools used by some OCG projects, which may have left you wondering: what entity manages this intricate process of supplying both the broader public and specialized researchers with data results to advance cancer care?

The answer . . . a Data Coordinating Center (DCC), specific to each project, residing within the NCI Center for Bioinformatics (NCICB). The NCICB DCCs track data generated through various genomics projects (Therapeutically Applicable Research to Generate Effective Treatments (TARGET)Opens in a New Tab [7], Cancer Genome

Characterization Initiative (CGCI) [5] and The Cancer Genome Atlas (TCGA)Opens in a New Tab [6] included), further administer quality control standards to the data and then make them available through publicly accessible databases supported by both the NCI (caBIG®) and the National Center for Biotechnology Information (NCBI). The concept of a project-specific DCC was created to connect quality genomics research results with the scientific and medical communities that will use the data to gain insight into potential causes and targets of various malignancies.

The NCICB DCCs maintain the OCG projects' Data Portals, which allow the broader public access to the biomedical and genomic information stored therein. Projects within OCG, including Cancer Genetic Markers of Susceptibility (CGEMS)Opens in a New Tab, CGCIOpens in a New Tab [12] and TARGETOpens in a New Tab [13], rely on this valuable resource to archive and distribute data resulting from these various initiatives in a user-friendly fashion for manipulation by the research community. The individual Data Portals use similar formats for the data generated, allowing for integration of multiple datasets across these OCG projects. Researchers have the added benefit of cross-comparing patient diseases, pathology and their genomic characteristics. The NCI works to vigorously protect the rights of all patients participating in these research projects, which is reflected in the multiple levels of approval required to gain access to any restricted data stored within the NCICB.

There are two levels of access for those wishing to obtain OCG project-generated data: open and controlled tiers. The open-access data tier requires no special permissions to retrieve publicly accessible data that cannot be aggregated to generate a dataset unique to an individual. Through this tier, researchers can investigate tissue pathology, de-identified clinical data, gene expression, tumor-specific copy number alterations and loss of heterozygosity, and somatic mutations. The controlled-access data requires specific approval from the NCI Data Access Committee after a requestor, and his/her Institutional Officer, have agreed to stringent terms for using the information acquired. The restricted datasets, while stripped of direct patient identifiers, are unique and are invaluable for research projects (such as target identification and therapeutics development) for which the open-access data are not sufficient. The controlled-access data tier contains data comprising additional demographic and clinical data, sample-to-trace relation (linking all sequence traces to a single participant), germline mutations, region-specific or genome-wide genotype data, and sequencing information for whole genome, exome and transcriptome.

OCG projects will contribute to the future of cancer care, and hopefully cure, by making project data available for broad use. The NCI DCCs enhance that effort through careful management of the ever-expanding databases that contain this wealth of genomic information generated by OCG initiatives. The Data Portals are meant to reflect the needs of researchers accessing the data and will therefore continually improve through user feedback. For instance, the TARGET and CGCI initiatives are introducing a tabular 'Data Matrix' approach to present data available for those respective projects. Providing the research community with the proper tools to access and navigate the data is critical to the overall success of cancer research by not only ensuring the accuracy of interpretation, analysis and subsequent translation of results, but also through securing the integrity of patient information gained from the generosity of study participants. Learn more about the resources and

tools available to you through an OCG Data Portal.

**OCG Data Portals:**

- TARGETOpens in a New Tab [13]
- CGCIOpens in a New Tab [14]
- CGEMSOpens in a New Tab

Back to Top

# FEATURED RESEARCHERS

## Perspectives from the Next Generation

**Ryan Morin, Ph.D.**

If analyzing next [2nd] generation sequence data is like trying to drink from a fire hose, then, as a graduate student focused on the application of next [2nd] generation sequencing to cancer research, I have swallowed a lot of water and have gotten very wet over the last few years.

In fact, the hose has been pumping faster and faster over time. Enhancements in sequencing throughput have meant a constantly increasing flow of data. Just when we think we have a handle on it, the speed of data generation increases exponentially and those of us responsible for analysis and pipeline development must continue consuming while attempting to maintain composure. I consistently find that code that was fast enough only a few months before has become intolerably slow. Just a couple of years ago, a few lanes of an Illumina sequencing run was an incredible dataset — enough to be the central focus of my masters thesis. As difficult and taxing as I found those data at the time, the depth and breadth of the data sets that I now routinely handle make that project appear quaint in comparison. I sometimes find myself reminiscing over the days when BLAST was still a viable aligner and when a few lanes of sequencing data was sufficient to yield a relatively high-impact publication. The more recent publication paradigm appears to be “one genome = one paper”, but alas, much like my code, that cannot last much longer either.

Being one of the few bioinformatics graduate students in Dr. Marco Marra’s lab, I have the opportunity to participate in the analysis of many of the tumor transcriptome (RNA-seq) libraries sequenced at our center. I was also given the privilege of assisting in the analysis of some of the first whole tumor genomes for the presence of somatic mutations and other potential cancer drivers. As a student in a genome center, I have also become accustomed to side projects and other non-thesis work consuming significant fractions of my time. These opportunities are difficult to ignore, and some eventually result in exciting discoveries that are rewarded by high-profile publications. However, it is often not possible to predict which projects will be successful.

As I find myself focused on writing and running code that sifts through biological data sets of unprecedented size and complexity, I feel fortunate to be involved in the

challenges of this work. The opportunities and experience that come with the territory are well worth the challenges. The fire hose may be daunting at times, but, in the end, I welcome the deluge of data for what is carried within it — clues that will ultimately make a difference in the lives of cancer patients.

Back to Top

# PATIENT PERSPECTIVE

## What Do Patients Know About Cancer Genomics and Personalized Medicine?

**Judith Ebbert-Syfrett, M.P.H., B.S., R.N.**

"The goal of getting your genome done is not to tell you what you will die from . . . it's how to take action to prevent disease."
– George M. Church, M.D., Ph.D. (in "On a Mission to Sequence the Genomes of 100,000 People" by David Duncan, 6/7/10, *New York Times*.)

Although few studies have measured patients' knowledge of cancer genomics, it may be safe to assume that most people need to learn more about cancer genomics and personalized medicine. A study of nearly 2,000 persons with chronic disease in the Netherlands by Morrena et al. (2006)[1] found that although most participants felt genetics would lead to advances, the majority, mostly older and less-educated persons, believed they had insufficient knowledge. Since cancer risk increases with age, the fact that older people may be the least informed about cancer genomics is cause for concern.

Everyone should make an effort to learn about the potential for genetics (the study of genes and heredity) and cancer genomics (the study of the human cancer genome) to improve outcomes related to the prevention and treatment of cancer. People need to understand that their own genetic design is unique, and that their individualized gene pattern can help predict risk for cancer as well as potential response to treatment. They need to understand that expecting the same drug to produce the same result in every patient is an oversimplified approach that is evolving toward genomics-guided personalized treatment decisions. They need to know that cancer genomics is exciting for its fast-growing potential to help oncologists make individualized treatment decisions with better outcomes.

If patients are uninformed about genomics, what should they know and where should they go to learn about it? An Internet search for a simple definition of genomics clarifies just how challenging it is to answer this question. We have much to do not only to simplify cancer genomics resources on the internet, but also to create resources for people who don't use computers.

What can we tell patients who worry about confidentiality and safety? We can honestly say that laboratory advances are reinforced by policies that protect patients from employer and insurer discrimination via passage of the Genetic Information Nondiscrimination Act in 2008. We can assure patients that the Health Insurance

Portability and Accountability Act (HIPAA) of 1996 continues to ensure the privacy of medical records, test results, and risk assessments.

As one who supports cancer genomics research on a daily basis at a comprehensive cancer center, I see how far-reaching these advances can and will be. I also see clearly that cancer genomics advances and protective policies are outpacing public awareness, compelling us to focus energy and funds on lay education. As science soars ahead, so too must our commitment to the beneficiaries of this bold innovation. We owe patients the education that will empower them to make informed, potentially life-changing decisions.

*Judith Ebbert-Syfrett, M.P.H, B.S., R.N., a doctoral student in the College of Public Health at the University of South Florida, is the Women's Oncology Research Project Manager at H. Lee Moffitt Cancer Center & Research Institute in Tampa, Fla.*

---

The NCI Office of Cancer Genomics (OCG) is a useful resource for patients who seek information on the genomic science that is being performed in cancer laboratories worldwide. The *OCG e-News* online newsletter highlights valuable patient resources that include the perspectives of patient advocates and cancer genetics experts, information on accessible genomics data, and several NCI genomics projects.

---

[1] Morrena, M., Rijkena, M., Baandersab, A.N., & Bensigna, J. (2006). Perceived genetic knowledge, attitudes towards genetic testing, and the relationship between these among patients with a chronic disease. Patient Education and Counseling, 65(2): 197-204.

Back to Top

*/

---