NATIONAL CANCER INSTITUTE
Office of Cancer Genomics

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

Issue 1 : March, 2010

DIRECTOR'S NOTE

Director's Note

Dear Colleague:

OCG PROGRAM HIGHLIGHTS

Recovery Act Funds Accelerate Cancer Genomics Research

The National Cancer Institute (NCI) has invested a significant portion of its American Recovery and Reinvestment Act (ARRA) funds in existing cancer genomics research programs, including The Cancer Genome Atlas (TCGA)

OCG PROGRAM HIGHLIGHTS

New Program Highlight: Cancer Target Discovery and Development (CTD²) Network

New Program Highlight: Cancer Target Discovery and Development (CTD²) Network

CONNECTING THE DATA

Enabling Cancer Genomics through Technology: An Overview of caBIG®

The cancer Biomedical Informatics Grid (caBIG®) is a virtual network of data, people, and organizations working toward the common goal of redefining—and improving—cancer care.

FEATURED RESEARCHERS

William Hahn, M.D., Ph.D.

“Don’t be afraid to ask important questions.”

DIRECTOR'S NOTE

Director's Note

Dear Colleague:

Welcome to the first National Cancer Institute (NCI) Office of Cancer Genomics (OCG) electronic newsletter. We are proud to launch this new communication tool to provide updates on ongoing projects, announce new projects, and highlight how OCG's efforts further the NCI mission to improve the lives of cancer patients by advancing the understanding of cancer's mechanisms at the molecular level.

NCI founded OCG in 1996 to examine molecular changes in cancer cells and to synthesize collection of data and reagents—ultimately working toward making this critical information publicly available. In an effort to also make this data “user-friendly,” OCG developed novel analytical methods for investigators to implement. This community approach to science continues to enhance global scientific efforts, and this free access policy underlies all of OCG's ongoing projects. Thanks to the rapid and easy access to data, researchers are able to integrate their own results with colleagues' findings, helping to accelerate the translation of basic data for the benefit of patients.

Among many diverse efforts to advance cancer research and treatment, the NCI supports three strategic projects in cancer genomic characterization:

- Therapeutically Applicable Research to Generate Effective Treatments (TARGET)Opens in a New Tab [2];
- The Cancer Genome Atlas (TCGA)Opens in a New Tab [3]; and
- The Cancer Genomic Characterization Initiative (CGCI)Opens in a New Tab [4].

These efforts aim to provide a comprehensive description of changes that occur in each cancer patient, determine the specific patterns that help to distinguish cancer subtypes with associated clinical criteria, and make all information available through an openly accessible data portal. The data generated by these projects promise to revolutionize our approach to cancer research in the next decade and are expected to dramatically accelerate our understanding of cancer. President Obama's recently-announced 2011 budget proposal underscores the significance of these projects: “The Budget will support the completion of a comprehensive catalog of cancer mutations for the 20 most common malignancies, setting the stage for complete genomic characterization of every cancer as part of medical care within 10 years.”

Furthermore, the translation of these large scale data “from bench to bedside” is under way, supported by the OCG initiative called the Cancer Target Discovery and Development Network (CTD^2 Network).

OCG projects enable and integrate ongoing research developed by other NCI divisions. Our overarching goal is to use public resources efficiently to benefit

patients in the U.S. and globally by developing therapies that individualize treatment with maximum efficacy and safety.

During President Obama's visit to NCI, National Institutes of Health (NIH) Director Dr. Francis Collins boldly said, “We are about to see a quantum leap in our understanding of cancer.” The OCG team enthusiastically agrees and appreciates your interest in following this transformative journey through the stories in our newsletters.

Sincerely,

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

Back to Top

OCG PROGRAM HIGHLIGHTS

Recovery Act Funds Accelerate Cancer Genomics Research

The National Cancer Institute (NCI) has invested a significant portion of its American Recovery and Reinvestment Act (ARRA) funds in existing cancer genomics research programs, including The Cancer Genome Atlas (TCGA)Opens in a New Tab [3] and Therapeutically Applicable Research to Generate Effective Targets (TARGET)Opens in a New Tab [2], and new programs like the Cancer Target Discovery and Development (CTD²) Network [5]. The NCI Office of Cancer Genomics (OCG) provides oversight and input to these and other programs, helping to ensure Recovery Act funds are used to accelerate genomics research.

Funding Enables TCGA Program Expansion, Pursuit of 20 New Tumor Types
NCI's Recovery Act investment in TCGAOpens in a New Tab [3], under the leadership of Joseph G. Vockley, Ph.D., director of the Office of TCGA Research, has facilitated the move from a pilot project to a full program that will take on the challenge of pursuing approximately 20 new tumor types over the next five years. This funding will support TCGA program areas for a period of two years.

TCGA will concentrate Recovery Act funds in two areas: building a more robust pipeline for accruing tissues and performing large-scale DNA sequencing on samples. Specifically, TCGA is committed to performing genomic analysis on 500 cases of each cancer type over five years. Each case consists of a sample of tumor tissue and normal tissue, often blood, from the same patient. Given its own rigorous standards and past experience, TCGA will need to process a minimum of 40,000 tumor and normal samples to gather the high-quality data needed to extensively profile all the tumor types at the depth of coverage and with the breadth of technologies that make TCGA so unique.

The TCGA expansion builds on the success of the pilot project, including the development of the infrastructure necessary to systematically characterize the

genomic changes in hundreds of tumors, the team science approach, and the broad use of the publicly accessible data sets that are enabling innovation. Most recently, TCGA researchers discovered that glioblastoma multiforme (GBM) is not a single disease but four distinct molecular subtypes (Verhaak, Hoadley, Purdom, et al, 2010Opens in a New Tab [3]). Researchers also found that response to aggressive chemotherapy and radiation differed by subtype, a discovery which may ultimately lead to more effective, targeted treatment strategies to combat GBM.

The expansion of TCGA is expected to lead to the most comprehensive understanding of cancer genomes in history. Data generated by the TCGA Network is made available via the TCGA Data PortalOpens in a New Tab [3], enabling researchers to advance their efforts and improve the prevention, diagnosis, and treatment of cancer.

Recovery Act Investment Enhances TARGET Pediatric Tumor Research Initiatives

The TARGETOpens in a New Tab initiative focuses on the systematic genomic characterization of pediatric tumors. Prior to the Recovery Act investment, TARGET investigators concentrated on two primary areas of research: acute lymphoblastic leukemia (ALL)Opens in a New Tab [6] and neuroblastomaOpens in a New Tab [7]. This new funding has expanded the TARGET initiative to include acute myeloid leukemia (AML)Opens in a New Tab [6], osteosarcomaOpens in a New Tab [8], and Wilms TumorOpens in a New Tab [9]. Additionally, the collaborators working on ALL and neuroblastoma have expanded the types of characterization being performed on these tumors and have included cases with recurrent disease. As a result of the changes, TARGET can provide a more comprehensive view of the extent of genomic alteration in these tumors.

Funding Supports Efforts of New CTD² Network

Over the past several years, genomic technologies have improved greatly and different platforms—such as those that detect copy number alterations, gene expression profiles, and methylation status—can now all be performed on DNA or RNA isolated from the same tissue. As these datasets become more abundant, researchers will need to distill these data down to the handful of driver mutations responsible for tumorigenesis, as well as potential targets for pharmacologic treatments. Accordingly, the CTD² [5] Network was created to develop a community of laboratories with experience in performing large-scale screens, optimizing biological assays, and developing informatics approaches that together can apply a systems approach to understanding the underlying genetic make-up of a specific tumor.

The following organizations have been selected as molecular CTD² Centers:

- **University of Texas Southwestern Medical Center Dallas, Texas**
  Michael Roth, Ph.D.
- **Cold Spring Harbor Laboratory, Long Island, N.Y.**
  Scott Powers, Ph.D.
- **Dana-Farber Cancer Institute, Boston, Mass.**
  William Hahn, M.D., Ph.D.; Lynda Chin, M.D.; and Ronald DePinho, M.D.
- **Columbia University, New York, N.Y.**
  Andrea Califano, Ph.D.
- **Broad Institute, Cambridge, Mass.**

Stuart Schreiber, Ph.D.

Back to Top

OCG PROGRAM HIGHLIGHTS

New Program Highlight: Cancer Target Discovery and Development (CTD²) Network

New Program Highlight: Cancer Target Discovery and Development (CTD²) Network [5]

OCG manages a diverse portfolio of initiatives that aim to enhance our understanding of the molecular mechanisms of cancer. Ultimately the goal of each of these programs is to improve the lives and outcomes of cancer patients, and where possible, prevent cancer before it becomes a tumor. The CTD² Network is one of these important initiatives. A recent investment from the Recovery Act spurred the formation of this five-team network of investigators who will leverage their respective expertise to accelerate the research pipeline from multi-dimensional cancer genomics datasets to target discovery and biological validation.

Why is now the right time?
Investigators in the CTD² Network agree that now is the perfect time to optimize such a pipeline. Programs like The Cancer Genome Atlas (TCGA)Opens in a New Tab [3], Therapeutically Applicable Research to Generate Effective Treatments (TARGET)Opens in a New Tab [2], and the Cancer Genome Characterization Initiative (CGCI)Opens in a New Tab [4] are assembling the “parts lists” of different cancers. TCGA is developing a catalog of all the genomic changes in at least 20 different cancer types. Already, copy number data, alterations in miRNA and gene expression, changes in methylation, and targeted and whole genome sequencing have been completed in hundreds of ovarian serous cystadenocarcinoma and glioblastoma samples. At the same time, analytical approaches to interrogate this data are also rapidly improving. This convergence of data and bioinformatics makes the present the right time for this network approach.

The CTD² Network brings together some of the world's leading experts in high-throughput small molecule and RNAi screening, animal modeling, and computational modeling. Over the next two years, the five CTD² projects will provide standardized protocols and data to the entire cancer research community. You can learn more about the CTD² Network in future issues of OCG e-News and on the OCG Web site [10].

The following groups have been selected as molecular CTD² Centers:

Organization	Project Title	Principal Investigator
Broad Institute, Cambridge, Mass.	Targeting Causal Cancer Genes with Small Molecules	Stuart Schreiber, Ph.D.
Cold Spring Harbor Laboratory, Long Island, N.Y.	Cold Spring Harbor Laboratory Molecular Target Discovery and Development Center	Scott Powers, Ph.D.

Columbia University, New York, N.Y.	Systems Biology of Tumor Progression and Drug Resistance	Andrea Califano, Ph.D.
Dana-Farber Cancer Institute, Boston, Mass.	Functional Annotation of Cancer Genomes: TCGA, Glioblastoma and Ovarian Cancer	William Hahn, M.D., Ph.D., Lynda Chin, M.D., and Ronald DePinho, M.D.
University of Texas Southwestern Medical Center, Dallas, Texas	A Concerted Attack on Patient Specific Oncogenic Vulnerabilities in Lung Cancer	Michael Roth, Ph.D.

New Program Highlight: Cancer Genome Characterization InitiativeOpens in a New Tab [4]

The Cancer Genome Characterization Initiative (CGCI) is another program in OCG's portfolio that is actively applying genomic technologies to human cancers. Currently CGCI is focusing on medulloblastoma primarily in pediatric cases and diffuse large B-cell lymphoma (DLBCL) in adults.

Medulloblastoma is the most common malignant brain tumor of childhood, accounting for approximately 20 percent of pediatric brain tumor diagnoses. Investigators at Johns Hopkins University are sequencing more than 20,000 genes (the protein coding exome from Ensembl, RefSeq, and CCDS databases) in 23 cases in the discovery phase. The data were submitted to the National Center for Biotechnology Information (NCBI) trace archive. The next phase of the project will include the validation of the mutated genes and estimation of mutation frequencies in a larger number of cases.

Diffuse large B-cell lymphoma is the most common type of non-Hodgkin lymphoma. The Genome Sciences Centre at the British Columbia Cancer Agency has recently released data on 51 cases of DLBCL and follicular lymphoma. These data are allowing investigators to discover mutations in expressed genes and evidence of other genomic alterations including, but not limited to, translocations, insertions, and deletions. Most recently, CGCI researchers discovered a subset of lymphoma with somatic mutations in EZH2 proteins. The research (Morin, Johnson, Severson, et al, 2010), published in January in *Nature GeneticsOpens in a New Tab* [11], also demonstrated that altered EZH2 proteins, corresponding to the most frequent mutations found in human lymphomas, have reduced activity using in vitro histone methylation assays. For more information on CGCI, the cancers being studied, and data access, visit the CGCI Web siteOpens in a New Tab [4]. Controlled access data is

available to approved users who apply through an official data access requestOpens in a New Tab [12].

Back to Top

CONNECTING THE DATA

Enabling Cancer Genomics through Technology: An Overview of caBIG®

The cancer Biomedical Informatics Grid (caBIG®) is a virtual network of data, people, and organizations working toward the common goal of redefining—and improving—cancer care. caBIG® is facilitating team science, where researchers collaborate and use standards-based resources, terminology, and technologies instead of performing research in isolation. Three goals define the caBIG® approach to changing scientific research from an independent, “siloed” method to a more integrative and collaborative effort:

- Adapt or build software that enables more collaborative use of cancer data;
- Develop an infrastructure that connects the entire cancer research community across the world, in real-time; and
- Develop and implement a standard vocabulary to improve communication and enable data integration.

Dozens of caBIG® applications are now in use by hundreds of researchers across the world, and their capabilities and reach continues to grow. The caBIG® community is structured into four domain workspaces that focus on specific areas of cancer research including clinical trials management; tissue banks and pathology; integrative cancer research; and *in vivo* imaging. In addition to these content-specific domain workspaces are cross-cutting and strategic level workspaces that focus on the underlying IT grid infrastructure and common vocabulary for this network, and a workspace focused on developing documentation and training for caBIG®.

Two OCG projects, The Cancer Genome Atlas (TCGA)Opens in a New Tab [3] and Therapeutically Applicable Research to Generate Effective Targets (TARGET)Opens in a New Tab [2], utilize caBIG®-compatible tools to make their datasets available to the scientific community. These tools include the Cancer Genome Workbench and the Cancer Molecular Analysis Portal. Providing the research community with standards-based formats and data types allows for easier integration with other projects and simplifies multi-disciplinary data analysis. As a result, it is easier for investigators to collaborate, using the same resources to make new discoveries.

Back to Top

FEATURED RESEARCHERS

William Hahn, M.D., Ph.D.

William Hahn, M.D., Ph.D. Associate Professor at the Dana Farber Cancer Institute Senior Associate at the Broad Institute

“Don’t be afraid to ask important questions.”

Dr. William (Bill) Hahn is a passionate oncologist making strides in the laboratory and in the clinic. Dr. Hahn knew early on that he wanted to be a doctor. He also knew that he would pursue oncology after watching several family members die from cancer. It was not until a research experience early on in his training at the Dana Farber Cancer Institute that Dr. Hahn realized that cancer research laboratories were contributing important, fundamental knowledge to a field where there was much more to learn. It was this convergence of a desire to pursue a career in oncology with his desire to make progress in improving treatments and cancer care through a fundamental understanding of the biology of cancer cells that made him pursue a dual degree.

Dr. Hahn notes there have been major paradigm shifts recently in our approach to cancer research. With new genomics technologies, we are able to ask comprehensive questions instead of focusing on one small element of a system. This comprehensive approach has condensed the timeline from target discovery to preclinical tests and clinical trials from a decades long process to something that might be accomplished over the training of one postdoctoral fellow (4-5 years).

Today Dr. Hahn is using multiple model systems, including tumor samples, to explore fundamental questions about cellular lifespan and how cancer cells use proteins, like telomerase, to become immortal. Hahn’s laboratory is interested in performing unbiased screens using RNAi technology to identify novel genes that may affect tumorigenesis. Together, this expertise is why Dr. Hahn is leading one of the five molecular Cancer Target Discovery and Development (CTD²) Network [5] centers funded by OCG. Together with Dr. Lynda Chin and Dr. Ron DePinho, Dr. Hahn will lead an effort to provide functional data to complement the cancer genomics work ongoing in their laboratories and public efforts like the Cancer Genome Atlas (TCGA) program. Using a variety of different methods ranging from assays in cell lines to mouse models, these investigators will determine which of the thousands of alterations found in cancer genomes are functionally important. Their goal is to find the real drivers and try to identify which ones may be the best targets for a therapeutic intervention.

The team approach is a powerful approach and the best way to tackle these complex problems, Hahn suggests. When asked about his philosophy for designing and implementing novel approaches that address tough, important questions like the ones he asks in his lab, he suggests: “Just do it.”

Back to Top

*/

Links
[1] https://ocg.cancer.gov/printpdf/taxonomy/term/4 [2] https://ocg.cancer.gov/programs/target [3] https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga [4] https://ocg.cancer.gov/programs/cgci [5] https://ocg.cancer.gov/programs/ctd2 [6]

https://www.cancer.gov/types/leukemia?redirect=true [7] https://www.cancer.gov/types/neuroblastoma?redirect=true [8] https://www.cancer.gov/types/bone?redirect=true [9] https://www.cancer.gov/types/kidney?redirect=true [10] https://ocg.cancer.gov/ [11] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2850970/ [12] https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login