OCG PERSPECTIVE

Contributions to Cancer Research: Finding a Niche in Communication

This past July, I started a journey into the fields of communications and cancer research when I joined the Office of Cancer Genomics (OCG) as a fellow in the National Cancer Institute (NCI) Health Communications Internship Program (HCIP).

FEATURED RESEARCHERS

Demystifying Clear Cell Sarcoma of the Kidney: An Interview with Dr. Saskia Gooskens

Saskia Gooskens, M.D., is a Ph.D. candidate in the Department of Pediatric Hematology and Oncology at Erasmus University Medical Center (Erasmus MC) – Sophia Children’s Hospital in Rotterdam, The Netherlands. While in medical school, Dr.

CONNECTING THE DATA

Tipping the Balance: Epigenetic Dysregulation in Non-Hodgkin Lymphomas

In 2008, the respective laboratories of Dr. Marco Marra and Dr.

EXPLORING CANCER GENOMES

The International Cancer Genome Consortium

The completion of the human genome sequence in 2003 sparked a global revolution in human disease research. For cancer research, the complete sequence meant that scientists could compare the DNA and RNA sequences between cancer and case-matched normal tissues to determine the genetic features of these diseases.

GUEST EDITORIAL
In the process of malignancy, cells acquire multiple genetic alterations, such as focal mutations, translocations, amplifications, or deletions, which allow them to survive and divide uncontrollably.

**NCI GENOMIC PROGRAM HIGHLIGHTS**

**Translating Cancer Targets into Nanotechnology-based Therapeutics: Accelerating Research Through Collaborations**

Dr. Stephanie Morris is a Program Manager in the NCI Office of Cancer Nanotechnology Research (OCNR), which supports the Alliance for Nanotechnology in Cancer Program.

---

**OCG PERSPECTIVE**

**Contributions to Cancer Research: Finding a Niche in Communication**

Jessica Mazerik, Ph.D.

This past July, I started a journey into the fields of communications and cancer research when I joined the Office of Cancer Genomics (OCG) as a fellow in the National Cancer Institute (NCI) Health Communications Internship Program (HCIP). Cancer genomics and working in an office were new and uncharted territory for me: before I came to OCG, I was finishing a Ph.D. in cell biology at Vanderbilt University in Dr. Matthew Tyska’s laboratory. I moved away from the bench to explore a career where I could use my background in biology to share with a broader audience stories about scientific discoveries and my excitement for basic research.

This is an exhilarating time to be embedded in genomics-related communications, especially in the field of cancer research. Collaborative team science, like that being done through OCG, and advances in functional genomics – linking genomic alterations in tumors to biological consequences – are spurring discoveries that will hopefully lead to new therapeutic developments. Medical advances for cancer treatments remain in demand: the American Cancer Society projected that in 2014, over 1.6 million people will be diagnosed with cancer and over 585,000 will die from the disease\(^1\). These numbers emphasize the need for the research community to keep working toward understanding the molecular basis of this disease and keep open lines of communication about discoveries. They also underscore the need to keep members of the public, many of whom will experience cancer personally or through a loved one or relative, informed of advances in basic cancer research and treatment options.

OCG fosters collaborations that promote discoveries, facilitate data sharing among cancer researchers, and enhance possibilities for investigators to work together and build on each
other's findings. OCG currently supports three team science programs, Therapeutically Active Research to Generate Effective Treatments [2] (TARGET), Cancer Genome Characterization Initiative [3] (CGCI), and Cancer Target Discovery and Development [4] (CTD²). TARGET and CGCI are aimed at defining the genomic underpinnings of childhood cancers and rare cancers, respectively. CTD² focuses on validating and characterizing the biological significance of the genomic alterations revealed by large-scale genomic initiatives, such as The Cancer Genome Atlas [5] (TCGA), TARGET, and CGCI. In OCG communications, our audience members are primarily investigators supported by the office as well as the broader scientific community. Our goals are to help researchers understand cancer genomics research (largely related to OCG research initiatives), and make data generated by OCG investigators widely available and easily accessible to the scientific community.

One of my first contributions to the office's communications effort was to design an OCG poster to present at NIH's annual Research Festival [6], which brings together scientists from around campus, as well as trainees from other academic institutions. This is a special event; it allows scientists to interact with each other and learn about research happening across NIH institutions. Although the majority of attendees at this event are scientists, some non-scientists, such as vendors and family members of patients at the NIH Clinical Center, also attend. While presenting the OCG poster at the Research Festival this past November, I had a unique opportunity to interact with some of these individuals. One conversation was especially memorable.

A woman pushing an empty baby stroller stopped in front of the OCG poster and stood back, reading and taking in the words and images. She was quiet with a soft demeanor. After a moment, I asked her if I could talk her through the poster content, and she replied softly that she would appreciate the additional explanation. I briefly summarized the work supported by OCG: comparing the genetic make-up of tumors to normal tissue, and validating that information in hopes of providing insight to develop new therapies. Then, I started to highlight specific aspects of each OCG program. When I started talking about TARGET and its goal to molecularly characterize childhood cancers, she revealed that she was a visitor at the NIH Clinical Center with her child, who was suffering from a brain tumor. They had visited the Clinical Center several months before to undergo treatments, which had not been successful, so they were back again for a clinical trial. I immediately expressed sympathy and hope, saying, “You are in good hands at the NIH. Hopefully you will have some positive results from the clinical trial.” Her response was selfless and inspiring: “Even if the trial doesn’t work for us, at least it will help others down the line.”

I was taken aback by her altruism, and struggled to respond. Her words are a reminder that our goal as a cancer research community is to strive to improve and lengthen the lives of patients so they can enjoy their lives and loved ones to the fullest. It is easy to lose sight of this human perspective of research as we go about our daily experiments, routines, and projects, but we should try to keep this big picture goal in mind.

There are many ways to contribute to this goal and the endeavor of learning about, preventing, and treating cancer. As its mission statement suggests, OCG does so by supporting research that strives to enhance the understanding of the molecular mechanisms of cancer, advance genome science and technology development, and accelerate the translation of findings into the clinic. Other offices, institutions, and individuals at NCI and elsewhere contribute in their own unique ways, too, through research, education, advocacy, and communication.

Communication within and outside the cancer community plays important roles in fulfilling this
goal of improving and lengthening patients’ lives. Within the scientific community, communication can spark collaborations and promote data sharing, thereby streamlining the process of turning molecular discoveries into opportunities for therapeutic development. Broader communication efforts can raise public awareness about the most current research efforts, findings and medical advances and help alert patients to new clinical trials or recently developed therapeutics. This can empower patients to make the most informed decisions about consent or treatments. Spreading this information, thus, provides a critical link between those deeply involved in the cancer research field and those who are not involved but are affected by cancer through their own experiences or through the experiences of friends or family members.

Team science and technological developments are leading to unprecedented new knowledge about the molecular mechanisms of cancer. It is critical now more than ever to build a bridge between the scientific community and the public, including patients, through good education and communication. Although OCG primarily targets a scientific audience, my experiences in the office are helping me build my scientific knowledge and become a better communicator overall. I would like to channel these experiences for future endeavors of developing creative ways to inform and educate various audiences about cancer research. Patients and their families, like the mother I met at the NIH Research Festival, are counting on the cancer research community. Let's keep working together to fulfill their expectations.

References


FEATURED RESEARCHERS

Demystifying Clear Cell Sarcoma of the Kidney: An Interview with Dr. Saskia Gooskens

Edited by Jessica Mazerik, Ph.D., and Jaime Guidry Auvil, Ph.D.

Saskia Gooskens, M.D., is a Ph.D. candidate in the Department of Pediatric Hematology and Oncology at Erasmus University Medical Center (Erasmus MC) – Sophia Children’s Hospital in Rotterdam, The Netherlands. While in medical school, Dr. Gooskens became interested in pediatric oncology research and, thus, began studying treatment and outcomes in children with cancer. Several publications resulted from this work, as did an intense interest in clear cell sarcoma of the kidney (CCSK), a rare form of childhood renal cancer.

Inspired to pursue a Ph.D. to complement her medical school education, Dr. Gooskens now studies the clinical and biological facets of CCSK and other non-Wilms renal tumors. Drs. Marry van den Heuvel-Eibrink and Rob Pieters are her advisors. She recently completed a four-month fellowship in the United States (Chicago, IL) in the laboratory of Elizabeth Perlman, M.D., as part of her doctoral training. There she gained new experience in molecular characterization by performing integrated genomic analysis on data generated from CCSK tumors. Dr. Perlman, Head of the Department of Pathology and Laboratory Medicine at Ann & Robert H. Lurie Children’s
Hospital of Chicago, is the Principal Investigator for the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) Kidney Tumors project.

In this interview, Dr. Saskia Gooskens shares some of her knowledge about CCSK and her experiences participating in the TARGET Kidney Tumors project.

What are the clinical and epidemiological features of CCSK? What are the challenges in working with such a rare tumor?

CCSK is indeed rare, which presents a challenge in working with this type of tumor. To collect clinical data or tumor samples from a sufficient number of CCSK patients, intensive global collaboration between international pediatric renal tumor study groups is necessary. CCSK comprises an estimated 5% of all primary renal tumors in children and occurs in about 30 patients per year in the United States. The median age of patients diagnosed with CCSK is about 2.5 years and there is a male predominance for this cancer. Most patients present with stage I or II disease, and only 5-7% of patients present with metastatic disease, most commonly in the lungs or bone.

Currently, 5-year event-free survival is approximately 80% and 5-year overall survival is around 85% for CCSK patients. These survival rates are favorable, however, curing patients with advanced stage or recurrent disease remains challenging. Additionally, CCSK is treated very intensively with high doses of chemotherapy (three or four drugs, including anthracyclines) and radiotherapy in patients with stage II, III, or metastasized disease. These required intensive therapies cause serious toxicity, such as anthracycline-related cardiotoxicity, even for children with low stages of disease, and often result in other devastating complications like infertility and second malignancies. Due to lower incidence and subsequent lack of sample availability, clinical studies on CCSK are scarce and the molecular background of this cancer is poorly understood. This scarce availability of information makes clinical management in individual children difficult, especially when tumors behave aggressively. Working to increase knowledge of CCSK makes it valuable and rewarding to study this particular tumor type.

What genetic alterations have been previously identified in CCSK?

Not much is known about the molecular background of CCSK. Gene expression profiling studies have reported evidence for activation of Sonic Hedgehog and protein kinase B (or Akt) cell proliferation pathways. Over the last two decades, three independent studies have reported a clonal balanced translocation involving chromosomes 10 and 17 \( \text{t}(10;17)(q22;p13) \) in patients with CCSK. This prompted a study by O'Meara and others, who found a rearrangement of a gene with unknown function, \( \text{FAM22} \) (chromosome 10), and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon, \( \text{YWHAE} \) (chromosome 17), in 6 of 50 cases. Little and colleagues also tested a dozen CCSK samples and showed they were strongly immunoreactive for epidermal growth factor receptor, EGFR. Gene amplification and somatic mutations in \( \text{EGFR} \) were described in 2 of these cases.

How is CCSK distinguished from other renal tumors, such as Wilms tumor?

There are currently no reliable immunohistochemical or genetic markers specific to CCSK, so it must be distinguished from other pediatric renal tumors by histology alone. It is important that expert pathologists perform the histological review of pediatric renal tumors, as CCSK shows remarkable morphologic diversity and therefore is often difficult to distinguish from other pediatric
renal tumors.

**How did you first decide you wanted to do the fellowship in Dr. Perlman’s lab, and how did you arrange it?**

Over the last decade, members of the Renal Tumor Study Groups at the European International Society of Paediatric Oncology (SIOP) and the Children’s Oncology Group (COG) have established a harmonious collaboration. This partnership has already resulted in successful projects that help guide treatment of renal cancer in children. At the 2012 International SIOP pediatric oncology meeting in London, Dr. van den Heuvel-Eibrink introduced me to one of her many collaborators in this effort, Dr. Perlman. It was then that we first discussed the TARGET Kidney Tumors project and the possibility of me doing a fellowship with her in Chicago. This fellowship offered a unique opportunity to gain experience in molecular characterization, work with a major expert in the field of pediatric renal tumors (Dr. Perlman), and be part of the collaborative effort to link genomic research to clinical advances.

**Describe the research project you worked on while you were in Dr. Perlman’s lab.**

During my fellowship, I contributed to the comprehensive molecular characterization of CCSK. I analyzed and integrated whole genome sequencing data, along with array-based methylation, gene expression, and copy number data. These data were generated using primary tumor and matched normal kidney samples from 13 CCSK patients. The raw data were ready for analysis when I arrived in Chicago, and during my four months there, I applied bioinformatics tools and analytical techniques to analyze specific data and integrate the data across all platforms. Although I am back in The Netherlands, we are currently finalizing the manuscript that describes the findings of this extensive CCSK molecular characterization.

**What did you learn from the fellowship, and how will you apply that towards your current and future medical endeavors?**

It is impossible to explain everything I learned in this limited space! When I arrived in Chicago in April 2013, I had only experience with clinical and basic molecular biology research. During my four months with Dr. Perlman, I learned how to interpret, analyze, and integrate high-throughput genomic data. Dr. Perlman is such an expert in the field of pediatric tumors and was very patient in explaining to me all the analysis techniques and discussing the results.

I was also fortunate to meet and collaborate with other researchers in Dr. Perlman’s lab, the COG, and the National Cancer Institute (Office of Cancer Genomics, Cancer Therapy Evaluation Program and the NCI Center for Bioinformatics and Information Technology). Before returning to The Netherlands, I further had the opportunity to present my research at the annual TARGET Steering Committee meeting in Washington, DC.

The results of this TARGET CCSK project will hopefully contribute to ongoing efforts to improve and develop therapeutic interventions for children with CCSK. As for the future, I plan to combine molecular research with clinical work as I continue my medical training in pediatric oncology after I complete my PhD. What I learned in Dr. Perlman’s lab will definitely be valuable as I pursue this endeavor.

Apart from the science, the fellowship gave me an opportunity to live in a fascinating American metropolitan city. My free time was scarce, but I also had the opportunity to connect with the Chicago field hockey team (I am an active field hockey player in The Netherlands).
References


Back to Top

CONNECTING THE DATA

Tipping the Balance: Epigenetic Dysregulation in Non-Hodgkin Lymphomas

Shannon Behrman, Ph.D.
In 2008, the respective laboratories of Dr. Marco Marra and Dr. Randy Gascoyne at Canada’s Michael Smith Genome Sciences Centre (GSC) [9], British Columbia Cancer Agency (BCCA), partnered together to sequence the DNA and RNA of the two most common non-Hodgkin lymphomas (NHLs). Their objective was to diagnostically characterize the NHL tumors by using advanced sequencing techniques. Previous molecular studies had identified a few hallmark abnormalities in these cancers, but a more detailed picture of the genomic landscape was needed. Their comprehensive analysis, supported by the Cancer Genome Characterization Initiative [3] (CGCI), revealed a surprising array of novel alterations[^1][^2][^3]. Among the most interesting were alterations that pointed to the epigenome as an influential player in lymphomagenesis.

The two NHLs, follicular lymphoma (FL) and diffuse large B cell lymphoma (DLBCL), are cancers of mature B cells. FL is slow-growing and difficult to treat; DLBCL is often aggressive and clinically diverse. In both cancers, processes involved in B cell differentiation are frequently perturbed[^4]. Normal differentiation occurs when an antigen-activated B cell induces formation of a germinal center in the lymph nodes. The B cell localizes to the germinal center and then proliferates, while simultaneously undergoing antibody diversification. The resulting outcome is a diverse collection of B cells with antigen receptors of varying affinities. The mature B cells with the highest affinity receptors are selected to become memory or antibody-secreting plasma cells and exit the germinal center. Those not selected self-destruct through apoptosis.

Lymphomagenesis can occur when acquired alterations block completion of differentiation and cause B cells to divide uncontrollably and evade cell death. Consequently, FL and DLBCL may resemble B cells at a particular stage of differentiation[^4]. Gene expression profiling revealed FL tumors have expression signatures like those of germinal center B cells. Additionally, DLBCL...
tumors can have signatures like those of germinal center B cells or activated B cells (B cells that have exited the germinal center, but have not fully matured into memory or plasma cells). Accordingly, DLBCL tumors are classified into at least two subgroups: germinal center B cell-like (GCB) and activated B cell-like (ABC). These subgroups show differences in growth behavior and response to treatment, suggesting differences in pathogenesis as well.

Separate genome-wide characterization studies provided additional clues that the underlying pathogenesis differs for each NHL subtype: several genetic abnormalities are specific to lymphomas derived from either germinal center B cells (i.e. FL and GCB DLBCL) or activated B cells (i.e. ABC DLBCL). For example, overexpression of the BCL2 oncoprotein, a result of a translocation between chromosomes 14 and 18, is only found in GCB-derived lymphomas (89% FL and 30-40% GCB DLBCL)5,6. NF-κB signaling dysregulation, on the other hand, is only found in the ABC subtype7,8. These initial molecular profiles were very insightful, but limited in granularity. Dr. Marra and the CGCI-supported investigators at the GSC reasoned that exposing underlying genetic alterations to inform more precise treatments for FL and DLBCL would require higher-resolution, integrative “omic” analyses.

By applying their expertise in sequencing and bioinformatics, the CGCI investigators examined the whole genomes, exomes, and transcriptomes of a large number of FL and DLBCL tissues. They published their results in three papers over the course of several years1,2,3. Integrating the data revealed an interesting trend in these lymphomas: genes encoding histone-modifying proteins and histone proteins10 themselves were frequently mutated. Of those genes, EZH2 and MLL2 had the most recurrent mutations. Both encode enzymes called methyltransferases that regulate gene expression by adding methyl groups to histones. By methylating histones, these enzymes modify the degree to which the DNA is packed and, thus, the degree to which genes and their promoter regions are accessible to transcriptional machinery. The EZH2 and MLL2 mutations are particularly compelling, because they are both predicted to function in malignancy by turning down gene expression.

**EZH2**

EZH2, an enzymatic subunit of the polycomb repressor complex 2, plays an important role in a variety of developmental processes. It tri-methylates lysine 27 on histone 3 (H3K27me3), which generally represses transcription. Expressed in a specific time window during B cell differentiation, EZH2 is required for the formation and functions of the germinal center9. Late-stage differentiating B cells must turn down EZH2 expression, so they can exit the germinal center to become memory or plasma cells. Before the publication of Morin et al. 2010, overexpression of EZH2 was linked to breast, prostate, and other cancers10,11, but mutations in the EZH2 gene itself had not yet been associated with any cancer type.

The CGCI investigators, including Dr. Marra’s then graduate student Ryan Morin, discovered a mutation hot spot in the EZH2 gene in their sequencing of mRNAs of DLBCL and FL cases1. The mutations recurrently affected tyrosine 641 (Y641) in the catalytic domain of EZH2. Interestingly, the Y641 mutations were restricted to GCB-derived lymphomas (21% GCB DLBCL and 7.2% FL), suggesting they are important in GCB-derived B cell malignancy.

To determine the effect of Y641 mutations on EZH2 function, the investigators performed *in vitro* experiments on several purified mutant proteins. The mutations disrupted the enzyme's ability to add the first methyl group onto a histone peptide, indicating they were loss of function (inactivating). However, they tested the EZH2 mutant proteins singularly, and not in the presence
of the wildtype EZH2. Y641 mutations were heterozygous in all cases identified, meaning each tumor had one copy each of the mutant and wildtype EZH2. Later experiments done by a different group showed that EZH2 mutations were actually gain of function (activating) when paired with the wildtype. Furthermore, EZH2-mutant cell lines have more tri-methylated H3K27 sites as compared to wildtype EZH2 cells, supporting the notion the mutations are indeed gain of function. Reconciling the two in vitro experimental results, researchers theorized that wildtype EZH2 efficiently adds the first methyl group to H3K27, which enables mutated EZH2 to add the second and third methyl groups more readily than the wildtype. The end result is more tri-methylated H3K27 sites and, thus, more repressive chromatin.

The question of precisely how activating mutations in EZH2 contribute to B cell lymphomagenesis has elicited great interest from different research groups and is currently being explored. One recent study in mouse models and human cell lines indicates that mutant EZH2 may continuously silence genes antagonistic to malignancy (e.g. germinal center exit and proliferation checkpoint genes), which are normally transiently repressed in germinal center B cells. As a result, mature B cells remain locked in a state of incomplete differentiation in the germinal center, where they rapidly grow and divide unhindered. Taking this into consideration, it would make sense that EZH2 mutations would not participate in the pathogenesis of ABC DLBCL, because EZH2 is no longer expressed in activated B cells that have left the germinal center.

**MLL2**

MLL2 methylates the fourth lysine of histone 3 (H3K4), which promotes transcription. MLL2 regulates a diverse array of cellular processes and signaling pathways, including the tumor suppressor p53 pathway, and is frequently mutated in certain cancers. For example, another group of CGCI investigators studying medulloblastoma identified recurrent mutations in MLL2 in 14% of all cases. The majority of MLL2 mutations identified in this study were predicted to be loss of function, hinting that it may be a tumor suppressor. However, little is known about how MLL2 participates in oncogenesis.

In their analysis of sequence from the genomes, exomes, and transcriptomes of 117 FL and DLBCL tumors, Morin and the CGCI-supported investigators identified MLL2 as one of 109 recurrently mutated genes. MLL2 showed the greatest mutational frequency, with the majority of identified mutations predicted to disrupt MLL2 function. These largely inactivating MLL2 mutations were found in 89% of FL cases and 32% of DLBCL cases, a prevalence on par with the t(14;18) translocation. Unlike EZH2, MLL2 did not show bias towards any DLBCL subtype. Because the role of MLL2 in B cell differentiation is not well-established, the function of MLL2 mutations in lymphomagenesis remains unclear. However, it is possible that impairment of its H3K4 methylating activity alters gene expression in favor of malignancy. This remains to be determined.

**So, What’s Next?**

Sequencing the DNA and RNA of FL and DLBCL tumors revealed many frequent alterations in histone-modifying genes, most notably in genes affecting H3K27 (EZH2) and H3K4 (MLL2) methylation. The investigators at the GSC, along with several other groups, are following up by studying how the EZH2 and MLL2 alterations specifically contribute to lymphoma biology and how they may be exploited to improve treatment strategies for these cancers. Because both genes might regulate chromatin marks in the same promoter regions and H3K4 methylation...
opposes H3K27 methylation, researchers speculate that both alterations produce the same outcome in differentiating B cells: increased H3K27-methylated chromatin and repressed transcription of genes antagonistic to malignancy. However, additional and/or alternative theories for their roles in lymphomagenesis are entirely plausible. Both EZH2 and MLL2 have widespread functions that are not fully characterized and their exact mechanisms in B cell differentiation are not completely or clearly defined.

For EZH2, the path to drug development is more straightforward, because the recurrent mutations result in increased activity, which can be counteracted by chemical inhibition. The biopharmaceutical company, Epizyme, recently reported tumor growth inhibition in NHL mouse models bearing the cancer-causing EZH2 mutations when treated with a small molecule inhibitor of EZH (EPZ-6438). They have since initiated a phase I/II clinical trial to test the safety and efficacy of using EPZ-6438 to treat lymphomas harboring those mutations. For MLL2, therapeutic development is trickier, because the recurrent mutations result in decreased activity, rendering chemical inhibition a non-starter. To circumvent this problem, Dr. Marra’s lab is working on identifying synthetic lethal interactions in MLL2 mutant tumors as potential alternative drug targets.

The Emerging Role of the Epigenome

In addition to EZH2 and MLL2 mutations, the CGCI-supported investigators and other groups detected frequent mutations in other genes encoding previously identified epigenetic modifiers (e.g. CREBBP/EP300). They also found novel mutations in genes encoding histone proteins (e.g. HIST1H1C) and transcription factors that recruit epigenetic modifiers (e.g. MEF2B), confirming the importance of epigenetic dysregulation in B cell lymphoma biology. Clearly, more research must be done to elucidate the epigenome’s specific mechanistic contributions.

Dr. Marra and colleagues are currently applying their savvy sequencing and bioinformatics approaches to examine the epigenome of NHL to gain deeper insight into its role in lymphomagenesis. Considering the mysteries that have been revealed in their previous studies of the genome and transcriptome, the NHL epigenome could prove equally illuminating.

References


EXPLORING CANCER GENOMES

The International Cancer Genome Consortium

Shannon Behrman, Ph.D., and Jessica Mazerik, Ph.D.

The completion of the human genome sequence in 2003 sparked a global revolution in human disease research. For cancer research, the complete sequence meant that scientists could compare the DNA and RNA sequences between cancer and case-matched normal tissues to determine the genetic features of these diseases. Large-scale cancer genome characterization projects were initiated at Johns Hopkins University (US)\(^1\), Wellcome Trust Sanger Institute (UK)\(^2\), and the National Cancer Institute/National Human Genome Research Institute (US)\(^3\). By leveraging advances in biomedical research, genomics technologies, and bioinformatics, each of these initiatives cataloged an array of genetic alterations in different tumor types. While the results were enlightening and exciting for the cancer research community, it called to attention the need for more open dialogue and data sharing between researchers participating in disparate cancer genome studies around the world.

To address this pressing need, cancer researchers and representatives of government institutions from 22 countries held a meeting in 2007 in Toronto, Canada. They discussed creating an international consortium that would serve as a hub for communication and exchange of “lessons learned,” as well as a venue to share results and provide opportunities for collaborations. The idea was well received by the meeting’s participants, and the International Cancer Genome Consortium \(^4\) (ICGC) was established. The ICGC formed a number of working groups to develop and refine policies for global cancer genome analysis and data sharing. These groups would draw from the insights gleaned from The Cancer Genome Atlas \(^5\) in the US, Cancer Genome Project \(^15\) in the UK, and other large-scale initiatives.

The ICGC Mission

The ICGC serves two main purposes. First, it serves as a centralized communications forum for the international scientific community, where researchers regularly share information and engage in helpful discussions about cancer genome research. Secondly, it deploys a bioinformatics database and portal (below) with the ultimate goal of warehousing the “genomic, transcriptomic and epigenomic changes in 50 different tumor types and/or subtypes which are of clinical and societal importance across the globe” \(^5\). It is an incredibly ambitious effort, with 17 countries currently participating (Figure 1) and 42 cancer genome projects included to date.

To encourage submission of high quality, reliable, and comprehensive data for each of these projects, ICGC provides guidelines for participation \(^16\). Strong recommendations are included for various aspects of genomic analysis, from tumor number and sample quality to data collection and generation. For example, data must be high quality and resolution (sequence-level is preferred but not required) with a minimum level of coverage. To allow for comprehensive analysis with clinically relevant discoveries, ICGC encourages researchers to provide a range of somatic mutations and alterations. This might include single nucleotide variants, indels, and chromosomal
rearrangements, along with clinical, histopathological, and, if any, environmental data for each tumor type or subtype.

**Introducing the ICGC Data Portal**

Compiling vast amounts of data from ICGC projects across the globe and disseminating them to the international research community make data storage and management a herculean challenge. The ICGC developed a portal, called the ICGC Data Portal to provide a single location where all analyzed data in standardized formats would reside. Member-country projects, which house their own raw data in respective local databases in disparate formats, send synthesized data in universal formats to the ICGC Data Portal[17]. The Portal, which was launched in April 2010, is an easy-to-use web platform where users can visualize, query, and download “open-access” ICGC project data.

To balance reliability with need for timeliness, the ICGC releases approved, universally formatted data on a quarterly basis. The most recent data release (#15) occurred on February 3, 2014. In future releases, the Portal will undergo restructuring to accommodate summary and other data. Previous versions reside in the ICGC Data Repository[18]. The ICGC Data Coordination Center (DCC), which is housed in the Ontario Institute for Cancer Research[19], manages the data portal and releases.

*Figure 1: A list of countries participating in ICGC as of February 2014*

**Protecting Patients: A Major Challenge in Global Data Sharing**

Another enormous challenge in genome research is generating and sharing data that result in impactful discoveries without compromising the confidentiality and rights of patients. For ICGC, this issue is further complicated because each country has its own laws for patient protection, informed consent, and institutional review board (IRB) approval processes. ICGC has taken into account these legal and regulatory differences and developed suggested guidelines for informed consent, data access, and ethical oversight that minimize the risk of individual patient identification without impeding important research opportunities. These guidelines evolve over
time to keep up with ever-changing laws of participating countries (read the latest version from 2013 [21]). Continued dissemination of such guidelines to the worldwide research community is a major objective of ICGC.

To safeguard patient identities, only data that cannot be directly linked to an individual are available through the ICGC Data Portal. Such data are stripped of direct identifiers (e.g. names and social security numbers) and may include patient and tumor information, such as gender, age range, and histologic type/subtype, as well as interpreted data, such as normalized gene expression, computed copy number, and somatic variants. Nevertheless, individually unique, but not directly identifying, patient data, such as genotypes found in primary sequence data, do pose a theoretical risk of patient re-identification. Such data are considered “controlled access,” which means approval from corresponding funding institutions’ data access committees is required for use. There are two entities within ICGC that oversee controlled-access: the Data Access Compliance Office (DACO), which handles controlled data access requests, and the International Data Access Committee (IDAC), which helps establish data access guidelines and supervises DACO activities.

It is important to note that the regulations and policies of the different member countries vary, and some (including the US) may not submit controlled-access data to the ICGC Portal. In this case, researchers must apply to the country’s local agencies in order to obtain data directly from that project’s data repository. Examples of such repositories include the National Center for Biotechnology Information’s Database of Genotypes and Phenotypes, National Cancer Institute’s Data Coordinating Centers, and the Cancer Genomics Hub in the US or The European Bioinformatics Institute’s European Genome-phenome Archive in the UK.

OCG’s Therapeutically Applicable Research to Generate Effective Treatments [2] (TARGET) and Cancer Genome Characterization Initiative [3] (CGCI) are contributing interpreted data as it is published and verified. TARGET Acute lymphoblastic lymphoma phase I data, including clinical, somatic mutation from whole genome sequencing and RNA-seq, and gene expression data from 229 donors, were released on September 26, 2013 in release #14. Copy number and gene expression data generated by TARGET’s Neuroblastoma project were released February 3, 2014 with the most recent data release (#15).

**Inspiring a Global Shift Toward Open Communication**

The ICGC has a ten-year goal of providing the international research community access to data that molecularly characterize over 50 types of cancer. Already ahead of schedule, the ICGC continues to expand the number of available datasets by adding or initiating new projects, such as the Singapore biliary tract cancer project announced in late 2013 [6]. The ICGC is structured to dynamically respond to and evolve with changes that happen in technology development, patient consent, and data sharing. As the number of tumor types and amount of data grow, ICGC will undoubtedly continue to provide a unique global resource for cancer genome data.

The forward thinking that ICGC applies to international communication is also having a wider reaching influence over genomics data sharing across diseases. In June 2013, representatives from many countries met and signed a “global alliance letter of intent” that represented a commitment to responsibly share genomic and clinical data worldwide [7], while protecting patients. Representatives from both the National Cancer Institute and National Human Genome Research Institute, along with over 60 other institutions, signed this document. These efforts to communicate and share data internationally are steps in the right direction toward working as a
global community to advance the genomics field and capitalize on discoveries to help understand and treat diseases.

References


Back to Top

GUEST EDITORIAL

Cancer Therapeutics Response Portal: A CTD² Network Resource for Mining Candidate Cancer Dependencies


In the process of malignancy, cells acquire multiple genetic alterations, such as focal mutations, translocations, amplifications, or deletions, which allow them to survive and divide uncontrollably. In addition to these genetic changes, normal functioning proteins that play roles in innocuous cellular processes may become essential to cancer cell survival when they get co-opted into partnerships with oncogene proteins. The latter is a phenomenon known as non-oncogene addiction or oncogene-induced dependency. Insights into cancer genomes have led to novel clinical therapies that target oncogene proteins, and such targeted therapies have yielded high patient response rates. However, less than one percent of patients suffering from cancer benefit from these therapies, and patient-matched therapies that target essential non-oncogenic proteins or their related pathways have been difficult to determine. To accelerate discovery of patient-matched therapies, systematic approaches are needed to identify oncogene-induced dependencies that cancers acquire as a result of specific cellular or genetic features and small-molecule drugs that target them.

As part of the National Cancer Institute’s Cancer Target Discovery and Development (CTD²) Network, we profiled a large number of human cancer cell lines using a novel “Informer Set” of more than 350 small molecules to reveal such dependencies and their inhibitors. We deposited the dataset into a publicly accessible data portal, called the Cancer Therapeutics Response

Back to Top
Portal (CTRP; www.broadinstitute.org/ctrp \[26\]; Figure 1), so that other researchers can make connections between the genetic and lineage features of cancer cell lines and small-molecule sensitivities. The CTRP, along with findings from an initial analysis, are described in a recent publication (https://www.ncbi.nlm.nih.gov/pubmed/23993102 \[27\])\(^1\).

**Figure 1.** The Cancer Therapeutics Response Portal (CTRP; http://portals.broadinstitute.org/ctrp/ \[26\]) homepage (top left) allows users to access data through three entry points: small molecules, enriched features, or targets. Once inside the portal, users can mine this resource for novel and therapeutically exploitable vulnerabilities in different cancer types across ~185 small molecules. For example, a user might enter through the “Compounds” link then search for a molecule of interest, like navitoclax. Under the “General” tab, the navitoclax entry shows the chemical structure of the compound and provides other general information (top right). Under the “Enrichment Analysis” tab, users can select or exclude specific cell line subtypes and datasets. For example, CTNNB1-mutant cancer cell lines are sensitive to navitoclax (bottom).

Historically, human cancer cell lines derived from many different types of tumors have been used in high-throughput profiling efforts to reveal patterns of small-molecule sensitivities. These studies have been limited in the number, diversity, or level of molecular characterization of cancer
cell lines or small molecules used. One of the earliest profiling efforts, the NCI-60 probed a limited set of 59 cancer cell lines from various lineages with more than 100,000 diverse small molecules and identified mostly lineage-specific small-molecule sensitivities. More recently, the Cancer Cell Line Encyclopedia (CCLE), a joint effort between the Broad Institute and Novartis Institutes for BioMedical Research, profiled 479 cancer cell lines with significant genomic characterization using 24 anti-cancer drugs. Massachusetts General Hospital and the Wellcome Trust Sanger Institute profiled 350 cancer cell lines against 130 pre-clinical or clinical anti-cancer agents, though the set of genomic alterations correlated to sensitivity was limited to ~70 genes. The Lawrence Berkeley National Laboratory and MD Anderson Cancer Center profiled 77 therapeutic compounds against a panel of ~50 breast cancer cell lines and linked various subsets, pathways and genetic features of these lines to drug sensitivity.

By leveraging work from previous projects and recent advances in high-throughput technologies, we addressed previous shortcomings of high-throughput cancer profiling studies. We used 242 of ~1000 cancer cell lines from the CCLE, which had been genomically characterized for gene expression, amplification/deletion, and somatic mutation in 1,645 cancer genes, and had lineage or histology annotations. These cell lines were established from many different types of tumors. Furthermore, we used an Informer Set of 354 small molecules comprising 35 FDA-approved drugs, 54 clinical candidates, and 266 probes – tool compounds used in biological research to help illuminate pathways and mechanisms of protein function. These probes and drugs each selectively target distinct nodes in cancer cell circuitry and collectively modulate a broad array of cellular processes. Each cell line was grown in its preferred media, and each compound tested at eight different concentrations, with the starting concentration of each compound selected based on literature review. We generated concentration-response curves that determined each compound's potency and efficacy for each cell line. The area under the concentration-response curve was used to calculate compound (i.e., small-molecule) sensitivities. We input the resulting data as well as previously determined molecular characterizations/annotations into one database and created the CTRP for access, query, and visualization.

The current version of CTRP (Figure 1) uses pre-computed visualizations to display more than 75,000 statistically significant connections between 185 small molecules and cancer cell lines of a particular lineage or with particular mutations. The portal also allows users to customize the query by eliminating cancer cell line subtypes, such as frequently sensitive or highly mutated ones, from analyses. Specific details on how to make novel connections between molecular features of cancer cell lines and the sensitivity profiles are found in the corresponding publication (https://www.ncbi.nlm.nih.gov/pubmed/23993102). Of note, researchers may use the CTRP to identify drugs that could rapidly be tested clinically, because FDA-approved drugs and clinical candidates are included the Informer Set. This critical new resource may aid in advancing discovery of potential cancer drugs matched to the patient populations most likely to benefit from them.

The CTRP is a living resource that will grow over time. We recently completed a second phase of new small-molecule sensitivity data collection. This phase expanded the number of genetically characterized cancer cell lines to ~900 and the Informer Set to ~545 molecules. The new, enlarged Informer Set not only targets a larger swath of proteins, but also includes novel chemical entities, such as stapled-helical peptides, more FDA-approved drugs and clinical candidates, and rational combinations of small molecules. We are starting to plan a new generation of CTRP to accompany our new cancer cell-line profiling datasets. In particular, we anticipate three functionalities in a future CTRP version: (1) application of single-feature lineage or mutation enrichment analyses to new datasets with improved computational methods; (2) single-gene
correlation analyses between small-molecule sensitivity and gene-expression or copy-number features; and (3) compound and cell-line correlation and clustering analyses using small-molecule sensitivity data alone. We are also developing additional methods to analyze and visualize the data. For example, we are incorporating the ability to correlate small-molecule sensitivity to combinations of features or to genetic vulnerabilities uncovered by Project Achilles$^6$, another CTD$^2$-sponsored initiative that uses a genome-wide shRNA library to silence individual genes and identify those genes that affect cell survival.

We strongly encourage the scientific community to use the CTRP to mine for novel and therapeutically exploitable vulnerabilities in different cancer types. Many have already started using this hypothesis-generating tool$^1$. Preliminary user-trend analysis for the month of January indicates about 32 users log in per day, with each user viewing an average of >170 results. We hope to attract more users as we increase the inputs, features, and functionality of this resource. As the volume and diversity of data in publicly available resources like the CTRP continue to grow, the catalog of cancer genes and related dependencies as well as the drugs that target them will likewise grow. Thus, we anticipate the ability to match patients with potentially effective drugs and that improved patient outcomes will dramatically increase.

References

Dr. Stephanie Morris is a Program Manager in the NCI Office of Cancer Nanotechnology Research (OCNR), which supports the Alliance for Nanotechnology in Cancer Program. The Alliance focuses on the development and translation of nanotechnology-based cancer diagnostics and therapeutics. In the following article, Dr. Morris describes a recent funding opportunity developed to support collaborations between OCNR’s Alliance and OCG’s Cancer Target Discovery and Development (CTD²) Network. This opportunity leverages the strengths of both programs to accelerate the in vivo validation of cancer genes and the development of highly effective therapies that target them.

Traversing biological barriers such as the blood-brain barrier and cell membranes to deliver drugs to specific targets without affecting healthy tissues and cells is a regularly encountered obstacle. Nanotechnology has emerged as a unique approach that can bypass these barriers and deliver anticancer agents to tumor tissues in a directed and specific manner. Nanotechnology is the engineering and manipulation of materials to design tools and devices, primarily in the range of 1 to 100 nanometers. For cancer therapies, nanoparticles (e.g., liposomes, dendrimers, and hydrophilic polymeric nanoparticles, Figure 1, top) can be adapted to include specialized biomolecules, drugs such as existing or newly developed chemotherapeutic molecules, and other molecular scale components. Molecules that ensure targeting specificity can also be conjugated to nanoparticles. A major benefit of nanotechnology-based delivery platforms is that they are capable of directing therapeutics to their intended target, while protecting them from degradation and premature release (Figure 1; bottom).

Figure 1: Size scale for nanoparticles (e.g. liposomes, dendrimers, gold nanoshells, quantum dots, and fullerenes) as compared to other materials (top). Nanoparticles can be adapted to include biomolecules, drugs, or targeting and imaging molecules to form nanotechnology-based drug delivery platforms (bottom).
By acting as targeted carriers for drugs, nanodelivery platforms are capable of reducing toxicity and improving efficacy of chemotherapies. The first examples of successful nanotechnology-based medicines were DOXIL and Abraxane®. DOXIL® is liposome-encapsulated doxorubicin, a DNA intercalating compound used to treat a large number of cancers. Abraxane® is albumin-bound paclitaxel, a mitotic inhibitor used to treat breast, lung, and pancreatic cancers. Both nano-associated therapeutics have led to significant improvements in clinical outcomes for patients.

In 2005, NCI launched the Alliance for Nanotechnology in Cancer program [30]. The goal of the Alliance program is to rapidly develop novel cancer diagnostics and treatments from nanomaterials. Examples include highly sensitive, multiplexed in vitro diagnostic devices for early disease detection, and multifunctional imaging and therapeutic solutions that are progressing to clinical trials. The Alliance is able to achieve its goal by facilitating collaborations between researchers across disciplines, such as cancer genomics and proteomics.

Recently, the Alliance and the CTD² Network partnered to identify novel cancer targets and therapeutics that can be more rapidly advanced by nanotechnology. Cancer genome projects like The Cancer Genome Atlas [5] (TCGA) and others have generated vast amounts of molecular information on different tumor types and subtypes, revealing countless recurrent alterations. These alterations must be sufficiently tested in in vivo models to ascertain if they are functionally relevant in malignancy, and, if so, whether they can be targeted therapeutically. The CTD² Network has the expertise and dedicated high-throughput platforms needed to perform studies of functional validation and small molecule inhibition of cancer targets. Their research can greatly accelerate the Alliance’s development of successful nanotechnology-based medicines by providing candidate therapies (e.g. small molecule inhibitors, small interfering RNAs) to be tested in nanodelivery systems. In turn, the Alliance can provide the CTD² Network with enhanced mechanisms to overcome the challenges of delivering candidate drugs in vivo and evaluate their effectiveness.

Seizing the opportunity for a mutually beneficial collaboration between the two programs, OCNR awarded research supplements in September 2013 to three Alliance investigators partnering with members of the CTD² Network. To qualify for this award, Alliance investigators had to propose preclinical evaluation and validation of novel cancer target therapeutics delivered by
nanotechnology-based systems. The nanodelivery platforms described in their proposals had to be sufficiently mature, as demonstrated by prior *in vivo*/*in vitro* characterizations and successful delivery of a therapeutic or imaging agent *in vivo*. Genomic targets needed to be strongly justified, as did the rationale for selection of the partner CTD\(^2\) investigator.

### Table 1: Awarded Pilot Projects

<table>
<thead>
<tr>
<th>Project Title</th>
<th>Alliance Investigator and Affiliation</th>
<th>CTD(^2) Investigator and Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targeted siRNA nanodelivery to target ARID1A mutant ovarian cancers</td>
<td>Sangeeta Bhatia, M.D., Ph.D. MIT-Harvard Center of Cancer Nanotechnology Excellence</td>
<td>William Hahn, M.D., Ph.D. Dana-Farber Cancer Institute</td>
</tr>
<tr>
<td>Targeting oncogenic protein-protein interactions with antagonist peptides delivered by theranostic nanoparticles</td>
<td>Lily Yang, M.D., Ph.D. Emory University Cancer Nanotechnology Platform Partnership</td>
<td>Haian Fu, Ph.D. Emory University School of Medicine</td>
</tr>
<tr>
<td>Biological validation of lethality targets using nanotechnology</td>
<td>Anil Sood, M.D. Texas Center for Cancer Nanomedicine</td>
<td>Michael White, Ph.D. University of Texas Southwestern Medical Center</td>
</tr>
</tbody>
</table>

The awarded pilot projects are listed in the above table (Table 1). Briefly, Drs. Bhatia and Hahn are working together to develop novel treatment strategies for ovarian and other cancers. They are using tumor-penetrating nanoparticles functionalized with small interfering RNAs (siRNAs) to target candidate cancer genes revealed by functional analyses performed by the Hahn group. One promising target is *AT-rich interactive domain-containing protein 1A* (*ARID1A*), which is recurrently mutated in ovarian, breast, and other cancers.

Drs. Yang and Fu propose to treat pancreatic cancer by using receptor-targeted nanoparticles to deliver antagonistic peptides to inhibit the function of KRAS, a protein that is mutated in 95% of pancreatic tumors. Developed by Yang's group, these specialized nanoparticles are designed to overcome therapeutic resistance caused by tumor-stroma interactions within the tumor microenvironment, which is often characteristic of pancreatic cancer. They also have imaging capabilities that allow visualization of tumor response.

Drs. Sood and White are also working toward new treatments for ovarian cancer. They have mined the TCGA ovarian cancer genomic data and identified variations in the levels of several key microRNAs (miRNAs). After they complete the biological characterization of these miRNAs, they plan to test the efficacy of miRNA delivery using nanoliposomes. Their nanodelivery platform will be utilized and evaluated in a mouse model of ovarian cancer, where they can determine the contribution of these miRNAs to tumor growth.

The Alliance-CTD\(^2\) pilot collaborations aim to improve therapies for difficult to treat, aggressive cancers using novel therapeutic approaches. Less than a year into the award period, the successes of these pilot projects are difficult to evaluate. However, by merging tools and expertise, these collaborative efforts will likely accelerate the translational efforts of both programs. More importantly, the collaborations are exposing investigators from both the Alliance
and the CTD$^2$ Network to their respective programs' research interests, opening a window to future partnerships.

Source URL: https://ocg.cancer.gov/news-publications/e-newsletter-issue/issue-11

Links
[9] https://www.bcgsc.ca/
[15] https://www.sanger.ac.uk/science/groups
[25] https://ocg.cancer.gov/about-ocg/glossary#T
[29] https://ocg.cancer.gov/about-ocg/glossary#C