DATA CORNER

Pediatric Genomic Data Inventory: A Valuable Resource of Pediatric Cancer Projects

Each year, more than 200,000 children and adolescents under the age of 20 are diagnosed with cancer worldwide.

CTD² GUEST EDITORIAL

Quantifying the Timing of Metastatic Progression from Patient Genomic Data: Spatial Computational Inference of MEtastatic Timing (SCIMET)

Metastasis, the systemic spread of malignancy to an organ other than the primary site of cancer, is the chief cause of cancer related death. Consequently, understanding and controlling metastasis is one of the most pressing issues in cancer medicine.

HCMI PROGRAM HIGHLIGHTS

Collecting and Harmonizing Clinical Data Across an International Initiative

The Human Cancer Models Initiative (HCMI) is an international consortium with the goal to produce and distribute up to 1,000 novel human cancer models.

CTD² GUEST EDITORIAL

FASMIC: An Integrated Bioinformatics Resource for Functional Annotation of Somatic Mutations in Cancer

Next-generation sequencing studies identified thousands of somatic mutations in tumor cells.

OCG PERSPECTIVE

How OCG’s Programs Support the Innovative Science that Will
The mission of the National Cancer Institute’s (NCI) Office of Cancer Genomics (OCG) is to enhance the understanding of the molecular mechanisms of cancer; advance and accelerate technology development; and efficiently translate high-throughput data to improve cancer research, prevention,

DATA CORNER

Pediatric Genomic Data Inventory: A Valuable Resource of Pediatric Cancer Projects

Pamela C. Birriel, Ph.D. and Jaime M. Guidry Auvil, Ph.D.

Each year, more than 200,000 children and adolescents under the age of 20 are diagnosed with cancer worldwide. Although the cancer mortality rate has declined, cancer remains an important cause of child mortality with an estimated 80,000 cancer-related deaths worldwide in children over the age of one. Cancer is a genetic disease in that pediatric malignancies of the same name can differ in an older adult population, thereby requiring potentially unique therapeutic strategies. The Office of Cancer Genomics at the National Cancer Institute (NCI) is committed to advancing the understanding of the biology, as well as improving therapeutic mortality and morbidity, of childhood cancers through its highly-successful Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative. The pediatric research community has benefitted from programs like TARGET that provide access to well-curated molecular, genomic, and clinical data from large numbers of children and young adults, alongside the computational resources to analyze it.

To fuel new discoveries and treatments specific to the field of pediatric oncology, OCG developed a dynamic resource called the Pediatric Genomic Data Inventory (PGDI) to allow investigators to more easily locate genomic datasets being generated globally. PGDI is a new open-access resource for identifying and locating genomic datasets that can be used to further the understanding of childhood cancers and develop better treatment protocols for sick children. This resource lists ongoing and completed molecular characterization projects of pediatric cancer cohorts from the United States and other countries, along with some basic details and reference metadata.

The newly-launched PGDI is an evolving catalog that will be continually updated as new information is deposited by the research community and is only valuable if people contribute. Researchers are urged to participate and submit information about any pediatric molecular characterization cohort(s) and/or project(s) by becoming a ‘PGDI Data Contributor’. The success of this resource requires ongoing participation by investigators around the world, and feedback is both welcomed and encouraged.

The PGDI Contributor Application registers a secure submitter account (Figure 1). Detailed login instructions are then emailed to the applicant once a secure account has been created by an OCG site administrator.
Once an account is secured, data contributors can log-in and add content to the PGDI by completing the online form registering a new pediatric cancer project. Contributors have the option to submit information via an interactive online form or by uploading a tab-delimited file (Figure 2). Contributors can edit their submitted information at any time using either method.
When adding content for PGDI, whatever available detailed information should be included for the following categories: genomic project information; disease descriptors; patient case information; sequencing available (e.g. whole exome, whole genome, DNA panel, transcriptome); additional molecular data types available; and project contact information. Instructions for Contributors [6] provide further detailed instructions on how to register a new pediatric cancer sequencing project.

Submitted information is sent to a PGDI moderator within NCI to review for completeness and formatting prior to posting.

The OCG website provides the ability to view, sort, and search through the interactive inventory through the PGDI Display [7] in a customized way. Individuals can view the default display categories, select up to 10 categories among cancer types and data (Figure 3) to create a custom view, and/or download the complete tab-delimited inventory.

PGDI, the investigator-driven community resource, will continue to grow with data as data generation and participant interaction expand. For more details about PGDI and/or for help accessing and submitting data, please contact the NCI OCG at: ocg@mail.nih.gov [8].
CTD² GUEST EDITORIAL

Quantifying the Timing of Metastatic Progression from Patient Genomic Data: Spatial Computational Inference of MEtastatic Timing (SCIMET)

Christina Curtis, Ph.D.

Metastasis, the systemic spread of malignancy to an organ other than the primary site of cancer, is the chief cause of cancer related death. Consequently, understanding and controlling metastasis is one of the most pressing issues in cancer medicine. Despite its obvious clinical significance, many aspects of this lethal process remain elusive. In particular, when and how tumor cells become competent to disseminate and seed distant metastases is unknown. Indeed, to date, few functional determinants or ‘drivers’ of metastasis have been identified.

Our limited understanding of the metastatic process is due, in part, to the difficulty in obtaining paired primary cancers and distant metastases for genomic characterization. This task is practically challenging since relapses can occur many years apart such that patients are lost to clinical follow-up. Moreover, often only the first site of metastasis is biopsied. Further complicating matters, electronic medical records do not necessarily capture information on recurrence, making it challenging to retrospectively identify individuals with metastatic disease. For the aforementioned reasons, there remains considerable debate regarding the timing of metastasis although it has traditionally been assumed to that the ability to metastasize is acquired late during growth of the primary cancer, such that only a subset of cells have the capacity to disseminate and successfully seed metastases at foreign tissue sites.

Tumor progression and metastasis have long been appreciated to be somatic (clonal) evolutionary processes driven by somatic mutations (genetic and epigenetic), natural selection and neutral genetic drift. While the elements of clonal evolution are well defined, the dynamics that govern tumor initiation, subsequent growth of the founding tumor clone and progression are poorly understood. Fortunately, tumors covertly record their ancestry in the form of somatic genomic alterations that are stably inherited during cell divisions. Thus, mutational data can be used to reconstruct the lineage relationship amongst cells within a tumor using phylogenetic...
tools, thereby providing a window into this otherwise unobservable process. However, phylogenetic approaches cannot accurately time metastasis as they do not account for factors that can confound the interpretation of tumor evolutionary dynamics, including tumor spatial architecture, differences in growth and mutation rates and the extent of subclonal selection.

In order to estimate the timing of metastasis, the Curtis lab pioneered a new technique termed Spatial Computational Inference of METastatic Timing (SCIMET)\(^7\). This new method, developed by postdoctoral fellow Zheng Hu, leverages a spatial (3-D) computational model to simulate the spatial growth of realistically sized tumors (composed of $\sim 10^9$ cells or $10\,\text{cm}^3$) and a robust statistical inference framework to measure evolutionary parameters, such as the timing of metastatic dissemination and the mutation rate from genomic data in a patient-specific manner. SCIMET enables the efficient simulation of the dynamics of metastasis in thousands of virtual patients under a range of parameters while explicitly accounting for tumor spatial architecture and the ‘mode’ of tumor evolution, both of which critically influence clonal dynamics and can otherwise confound the interpretation of metastatic timing.

Figure: Inference of patient-specific evolutionary dynamics and the timing of metastasis from cancer genomic data using SCIMET.

In particular, the Curtis lab previously described a Big Bang model of colorectal tumor growth wherein after transformation, tumor growth is propagated by numerous equally fit subclones compatible with effectively neutral evolution or nearly neutral evolution\(^8\). This new model stands in contrast to the de facto sequential clonal evolution model of continuous positive selection during tumor progression and highlights the early origin of intra-tumor heterogeneity with several clinical implications. Building on the finding that distinct modes of evolution can be operative in different tumors and at different stages of disease progression, the Curtis lab developed a method to classify the two dominant ‘modes’ of tumor evolution, namely effective neutrality and subclonal selection from tumor sequencing data\(^5\).

Since the mode and tempo of progression differ between tumors evolving neutrally and those
subject to stringent subclonal selection, it is critical to account for this when characterizing the dynamics of metastasis.

By simulating tumor progression and the underlying genomic profiles for thousands of virtual tumors, SCIMET can delineate the contribution of different factors to the metastatic process and define the parameter sets that best explain the observed patient genomic data. This is of critical importance since ground truth measurements in the human system are not feasible. SCIMET can be used to determine optimal study designs and tissue sampling schemes and generates testable predictions regarding the metastatic process that can be evaluated experimentally.

Application of SCIMET to primary cancers and paired metastases has already begun to reveal mechanistic insights into this otherwise unobservable process, including the temporal dynamics and routes of metastasis. SCIMET will be an important resource for the cancer genomics community in light of the increasing focus on profiling metastatic disease and its broad generalizability for studying diverse solid tumors. Indeed, the systematic characterization of tumor evolutionary dynamics and the determinants of metastasis in clinically annotated patient cohorts has the potential to identify new therapeutic targets, as well as biomarkers that may ultimately be used to inform patient stratification and predict disease progression.

References

The Human Cancer Models Initiative [20] (HCMI) is an international consortium with the goal to produce and distribute up to 1,000 novel human cancer models. These models will be developed from a wide range of patient tumor types using next-generation culturing techniques, such as organoids and conditional reprogramming. Not only will the sequencing data acquired from the models, parent tumor, and normal tissue be accessible to the wider research community, but the models will also be associated with clinical data. Data from the NCI-funded Cancer Model Development Centers (CMDCs) will be found at the National Cancer Institute's (NCI) Genomic Data Commons [21] (GDC), and European consortium member data will be at the European Genome-phenome Archive [22] (EGA). The founders of the HCMI at the NCI, Wellcome Sanger Institute, Cancer Research UK, and the foundation Hubrecht Organoid Technology recognize the importance of clinical data to the utility of HCMI models, as it provides context regarding the patient’s exposures and treatments, as well as vital diagnostic information.

In the process of developing a system to provide these data, participants in the initiative have learned about the trials and rewards of collecting clinical information, both from different institutions and distinct countries. During a face-to-face meeting in June 2018, the HCMI consortium members shared their experiences with clinical data collection, and this perspective piece aims to share challenges as well as resolutions and opportunities that have arisen in response to these challenges.

The first necessity to providing thorough and accurate clinical data is establishing collaborations in which there is a clinical team that is fully dedicated to the goals of the project. Surgeons, oncologists, nurses, clinical research coordinators, and the institution all need to buy-in to the goals of the program to ensure its success. The HCMI has been very fortunate to be working with hospitals around the world that recognize the benefits of these models. When challenges arise (and they do), tissue source sites address them. The first step for the HCMI was to determine the clinical data elements that would be collected for each tumor type. With models being developed from cancers ranging from pediatric tumors, such as osteosarcoma and Wilms tumor, to common and rare cancers including breast cancer, lung cancer, gastric cancer, and ampullary cancer, clinical data working groups have been convened to address the important clinical data to be collected for 19 distinct tumor types with more anticipated in the future.

Working groups are comprised of oncologists, surgeons, pathologists, and clinical research coordinators from tissue source sites in the US, UK, and Netherlands. Members are asked to identify important clinical data collected at their sites, while keeping in mind the goal of providing information that will be useful to downstream model users. Of critical import is ensuring that the data are standardized so that they can be compared regardless of their origin. The development of translation tables has been important for guaranteeing consistency, especially in terms of demographic information, such as ethnicity, which is collected using different standard vocabulary depending on the country in which it is collected. As an added assurance, after the clinical data working groups have met, it is the job of the NCI, in collaboration with the Cancer Data Standards Registry and Repository [23] (CaDSR) to ensure that the metadata and vocabulary of the clinical data elements enable longevity and consistency of the data across all NCI programs. The selected clinical data elements are then used by the NCI Clinical Data Center to produce Clinical Report Forms (CRFs) for each tumor type. The completed CRFs are posted on the HCMI Resources [24] page, with more to be added as new models are developed.

Once the CRFs have been established, implementation in the form of clinical data collection is the goal of HCMI-associated clinical sites. As sites have begun the process of data collection and submission, they have identified mechanisms by which they can address challenges. For instance,
there are complexities to collection of data from patients that have switched hospitals; complications that are compounded by the fact that each hospital has a different mechanism for data collection and storage, as well as specific restricted vocabulary, making data merging and validation complicated. In addition, merging data from multiple sources often requires significant informatics support. Thus, tissue source sites rely on teams of knowledgeable and experienced data extractors and informaticians to bring disparate data into a format consistent with HCMI requirements. Development of these teams and standard operating procedures for data collection have been essential. Moreover, sites have begun to develop flexible and practical systems for data collection that can be applied at multiple institutions and allow for data entry using assorted mechanisms—from Excel spreadsheets and tab-delimited files to web-based submission systems. Finally, a robust quality control process must be in place, especially for data that are collected and/or entered into databases manually. HCMI contributors have established quality control pipelines that integrate not only the clinical sites but also the model developers and, in the case of the NCI-supported CMDCs, the Clinical Data Center which passes data through strict validators and checks for accuracy and completeness. Once the data have been collected, they then need to be submitted to a unified data repository that has been designed to enable data sharing. In the case of HCMI data, they will either be submitted to the aforementioned GDC (for US data) or EGA (for European data). At the conclusion of this process, the data are accessible to the research community.

The successful implementation of a clinical data collection-to-distribution process requires contributions from a range of people of varying expertise. From the NCI to the tissue source sites, the HCMI is addressing challenges to clinical data collection so the data that is ultimately distributed to the research community is of the highest quality. Please visit the HCMI website to see the fruits of the efforts of the many contributors to the initiative.

Many thanks go out to the following for providing tissues and clinical data and contributing their expertise and perspectives on clinical data collection to the HCMI: Office of Cancer Genomics external consultant Martin Ferguson; Dianne Reeves, Associate Director for Clinical Research Programs in CBIIT; Northwell Health and Feinstein Institute for Medical Research; Cambridge University Hospital; University Hospital Southampton NHS; University Hospital Birmingham NHS; Cancer Research UK Glasgow Centre; University Medical Center Utrecht; The Netherlands Cancer Institute; Meander Medisch Centrum; Ziekenhuis St. Antonius; Boston Children’s Hospital; Dana-Farber Cancer Institute; Brigham and Women’s Hospital; Boston Medical Center; and the Rare Cancer Research Foundation.

FASMIC: An Integrated Bioinformatics Resource for Functional Annotation of Somatic Mutations in Cancer

Jun Li, Ph.D. and Han Liang, Ph.D.
Next-generation sequencing studies identified thousands of somatic mutations in tumor cells. The availability of large-scale cancer molecular alteration data is used to identify "driver" somatic mutations\(^\text{[25]}\) that play a critical role in the initiation and development of human cancer. While different bioinformatic algorithms have been developed to computationally predict the functional effect of somatic mutations in protein coding regions, the accuracy of these methods is still limited and not only because somatic changes in non-coding regions are not evaluated. Some recent studies have systematically experimentally characterized the function of somatic mutations using pools of cDNAs. However, the interpretation of the results is confounded by the competition between cDNAs with different mutations in the pooled assays which is a critical concern. To overcome this issue, we developed a sensitive, efficient, and systematic approach to annotate the functional effect of >1,000 somatic genomic aberrations\(^1\).

To help the broader research community better use our results, we have built a user-friendly, open-access data portal, FASMIC (functional annotation of somatic mutations in cancer), which is publicly available at [http://ibl.mdanderson.org/fasmic/](http://ibl.mdanderson.org/fasmic/)[26].

Currently, FASMIC\(^{[27]}\) has curated >1,000 mutations including 923 missenses, 74 indel, 27 silent, and 25 nonsense mutations from 95 genes, among which 21 have >10 mutations for each gene. A few clinically actionable cancer genes, such as EGFR, BRAF, PIK3CA and ERBB2 have ~100 mutations per gene, which provide a good opportunity for researchers to study their allelic series. Overall, we have functionally annotated 1,049 mutations into different classes, including activating, inactivating, informative, non-inhibitory, inhibitory, and neutral mutations. The functional annotations were derived from two cell line models Ba/F3 and MCF10A based on the readout of the growth factor-independent cell-viability assays. In addition to the functional annotation, we also quantified protein expression using Reverse Phase Protein Arrays (RPPA) of >250 MCF10A cell lines that stably expressed different mutations. This protein profiling dataset is a valuable resource for elucidating the mechanisms underlying the phenotypic effects of the select mutations.

The best way to access our mutation functional annotation data is through FASMIC\(^{[27]}\), a web-based platform for researchers to query and visualize the data in an interactive manner. The main interface on the home page is a simple input box, in which users can search mutations of their interest by the corresponding gene symbol. Those queried mutations are displayed in a table view. All basic information for each mutation is displayed in different columns, including genomic position, amino acid change, functional annotation, and more importantly the details of availability status of the associated mutation data (green color indicates that the information is available). After selecting a mutation using the green icons under the column “Details”, the associated mutation data is presented under the table in six separated tabs: (i) “MS/Summary” shows detailed profile information about each selected mutation. The information includes genome build, genomic position, DNA change, variant type, variant classification, and functional annotations in different cell models, etc. (ii) “3D/3D View” is a 3D animation of the corresponding protein structure with the mutated residues highlighted in red color. One can rotate the 3D structure and find the 3D position of the selected mutation, which may provide important information for inferring the functional impact of the selected mutation based on the structural stability or the protein-protein interaction site, and further help better understand the function of the mutation in the system level. (iii) “MF/Mut.Freq.” shows the frequencies of each mutation across different cancer types in an interactive bar plot (downloadable using the menu icon). The mutation frequency data was obtained from The Cancer Genome Atlas project. (iv)
“FP/Func.Pred.” provides a table view showing the predicted functional effect. The functional scores were calculated based on 12 popular cutting-edge prediction algorithms (listed in the figure). This module helps researchers assess how well the predictions are aligned with the FASMIC annotation, and further provides more evidence to support the annotated functions. The color represents different functional levels. The red color means disease causing, pink means moderate, and white color means no effect. (v) “PE/RPPA” is a sorted scatter plot showing the differential protein expression level of a mutant relative to the wild-type gene across >300 protein markers. This plot shows the potential functional consequence of the selected mutation; and (vi) “PM/Literature” module lists the publications that are relevant to the queried mutation, which helps researchers easily find the previously published studies about the selected mutation.

**Figure:** Overview of the FASMIC web portal, which can be used for exploring functional effects of cancer somatic mutations.

In summary, FASMIC is a comprehensive database for functional impact of somatic mutations in cancer. It provides the functional annotation along with protein expression, mutation frequency, 3D structure, function prediction, and literature study to help researchers explore every detail about each mutation. This web resource will help identify potential driver mutations, discover novel biomarkers, improve prediction algorithm, and develop new drugs. We expect FASMIC to be a valuable resource to advance precision cancer medicine and develop novel cancer therapies.

**Reference**

11
The mission of the National Cancer Institute's (NCI) Office of Cancer Genomics (OCG) is to enhance the understanding of the molecular mechanisms of cancer; advance and accelerate technology development; and efficiently translate high-throughput data to improve cancer research, prevention, early detection, diagnosis, and treatment. OCG provides resources that bolster research aimed at reducing the burden of cancer; thus, aligning its mission with this year's theme from the American Association for Cancer Research's (AACR) annual meeting: “Driving Innovative Cancer Science to Patient Care”. Program managers for the current OCG initiatives share their perspectives on the important themes highlighted throughout the conference, along with how their respective programs contribute to those and further generate excitement and inspiration for the future.

Cancer Genome Characterization Initiative (CGCI)

Nicholas Griner, Ph.D.

The Cancer Genome Characterization Initiative (CGCI) projects, such as the Burkitt Lymphoma Genome Sequencing Project (BLGSP) and the HIV+ Tumor Molecular Characterization Project (HTMCP), represent rare cancer studies where available research can be sparse. Large annual meetings, like this year's AACR conference, allow investigators from around the world to congregate and share scientific knowledge. This is very beneficial for many involved in the study of common cancers, but it is especially useful for rare cancer type research. Large conferences like AACR allow rare cancer studies to be shared with a larger audience of the scientific community, which can spur new ideas that may be absent at tailored meetings. In addition, since tissue accrual can be difficult for rare cancer studies, large conferences allow greater networking opportunities leading to unexpected sources for tissues.

As scientific advancement continues to grow at an unprecedented rate, so do technologies for more efficient and precise analysis. The 2018 AACR annual meeting had a great number of technologies in the spotlight, including those from vendors and research labs demonstrating the efficacy of such technology. This is very beneficial for rare cancer studies as the quality of tissue accrual, from such areas as third world countries, can present a great challenge to the quality and quantity of tissue accrued. Attending seminars and posters unrelated to rare cancers still provides a glimpse as to how these technologies can be applied resulting in more accurate data sets for rare cancer studies.
Cancer Target Discovery and Development (CTD\textsuperscript{2}) Network

Subhashini Jagu, Ph.D.

The current rate of accumulation of genomic data has far outpaced the rate of interpretation of this data. The next phase of translational cancer research is functional genomics, using varied methods of experimentation to determine the function of observed genomic alterations in cancerous cells. Cancer is a constellation of diseases with common characteristics. It is critical to develop a better understanding of the cause of initiation, progression, and metastasis by bringing together multidisciplinary research teams. Hence, the Cancer Target Discovery and Development (CTD\textsuperscript{2}) Network, a functional genomics initiative, was established to bridge the knowledge gap between cancer genomics and biology for the development of human therapeutics. The CTD\textsuperscript{2} Network aims to understand the tumor heterogeneity and drug resistance for the development of efficient strategies to identify optimal combinations of small-molecules or immunotherapy with small molecules. The Network Centers utilize a distinct array of advanced computational and systems biology methods, functional genomics and immunological approaches, small molecule and genetic screens to translate existing genomic data into biomarkers and therapeutic targets in cancer. They share data and resources (e.g. analytical tools\textsuperscript{[31]}, reagents\textsuperscript{[32]}, etc.) across the Network and with the research community.

The CTD\textsuperscript{2} Network demonstrated how OCG supported “Driving Innovative Cancer Science to Patient Care” at AACR's 2018 annual meeting. The meeting hosted several talks and posters by CTD\textsuperscript{2} Network investigators, focusing on combining state-of-the-art high-throughput informatic\textsuperscript{[33]} and experimental approaches to functionally validate discoveries from genomic studies and high-throughput and high content small molecule and genetic screens and advance them toward precision oncology. For instance, in the “Translational Applications of Systems Biology” session, CTD\textsuperscript{2} Network investigators discussed how systems biology approaches could be used to target adoptive responses; identify cancer dependencies, driver mutations, and tumor checkpoints; overcome treatment resistance; and rationalize combinatorial therapies.

Human Cancer Models Initiative (HCMI)

Caitlyn Barrett, Ph.D.

The field of patient-derived models has exploded in terms of the tissues from which models can be derived and the potential for model use. Culturing techniques, such as organoids and conditionally reprogrammed cells, offer a unique opportunity for the scientific community to study individual human tumors \textit{in vitro} that closely represent the heterogeneity of the primary tumor from which they were derived. The NCI and founding institutions of the Human Cancer Models Initiative (HCMI) at the Wellcome Sanger Institute, Cancer Research UK, and the foundation Hubrecht Organoid Technology have developed large scale programs that are contributing to the development of patient-derived cancer models. Sequencing data from the model, normal tissue, and parent tumor; patient clinical data; and the models will be available to researchers worldwide as a community resource. It was thrilling to note that at this year’s AACR annual meeting, patient-derived cancer models were incorporated extensively into the science presented.

Models developed by the HCMI have great potential, and AACR was an excellent opportunity to see how scientists are beginning to use patient-derived models in their research. For example, to test targeted therapies, laboratories are developing new culturing methods, such as microfluidic
devices for 3D culturing of cancer cells, which enable better penetrance of chemotherapeutics for screening. To assess mechanisms of drug resistance, models are being developed from resistant disease and tested for new drug susceptibilities. Importantly, genomic and proteomic assessment of patient-derived models are offering new insight into mechanisms of tumor evolution, progression, and drug response. The HCMI will be producing and distributing up to 1,000 models into a scientific field that will benefit from new cancer models, and the next couple of years will clarify the potential of patient-derived models. The HCMI is a work in progress; for up-to-date information visit the HCMI Program page.

Therapeutically Applicable Research to Generate Effective Treatments (TARGET)

Jaime M. Guidry Auvil, Ph.D.

Childhood cancer is emerging as its own critical field in the research community, separate from adult disease counterparts by more than merely age. While many patients can be cured of pediatric malignancy, cancer remains the leading cause of death in children over one year old in the United States. Childhood cancers devastate families, both those who lose children to the disease and those who support their children as they endure a lifetime of hardship resulting from toxic harsh treatments. Once thought to be similar diseases, it is now widely recognized that cancers arising earlier in life are molecularly and behaviorally distinct from those that develop in the aged. Consequently, new research directed toward advancing precision oncology is being pursued to improve clinical outcomes and subsequent quality of life for these young patients, particularly in genomics, model systems, and immunotherapies.

This year’s annual AACR meeting focused largely on pediatric areas of investigation and the “big data” volumes that funded programs are generating in their studies. Of interest for all research efforts is the critical need for broad, open, and ethical data sharing policies and procedures to advance discovery. Much of the pediatric data sessions, in line with the rest of AACR generally, focused strongly on data sharing perspectives and processes. An international panel of experts convened by the AACR Pediatric Cancer Working Group discussed this key topic in depth, including what data types are stored and where, barriers to access, need for standardization of data and metadata formats and terminology, consistent consent language and processes, and all on a global scale. It is clear that scientific advances have been compromised by intended and inadvertent data silos, and the pediatric community is coming together to find solutions for childhood cancers. OCG initiatives, including its large-scale pediatric molecular characterization program known as Therapeutically Applicable Research to Generate Effective Treatments (TARGET), have always promoted broad and equitable data sharing practices to support the NCI mission.

Bringing it All into Focus

Pamela Birriel, Ph.D.

NCI had a large presence at this year’s AACR annual meeting in covering the latest basic, translational, clinical, and prevention-focused research in the field. OCG program managers presented during an NCI-sponsored session focusing on how the initiatives support “driving innovative cancer science to patient care”. OCG, as part of NCI, is innovating across the field of cancer genomics and translational science: from basic cancer genomics research to improving data accessibility via bioinformatics infrastructure development; identifying therapeutic targets, perturbagens, and biomarkers; exploring cancer heterogeneities and developing strategies to
overcome treatment resistance; as well as generating new cancer models to validate new therapeutic targets and diagnostic/prognostic markers. OCG unifies these activities by integrating research in different fields of cancer genomics—structural, functional, and computational [34]—to advance precision oncology and improve clinical outcomes.

Figure: OCG programs support innovative science that advances patient care.

Source URL: https://www.ocg.cancer.gov/news-publications/e-newsletter-issue/issue-19

Links
[8] mailto:ocg@mail.nih.gov
[12] https://ncit.nci.nih.gov/ncitbrowser/pages/home.jsf;jsessionid=495BA946CA8E50E08AF8B23C9F396106
[22] https://www.ebi.ac.uk/ega/home
[23] https://wiki.nci.nih.gov/display/cadsr/cadsr+wiki