Informatics Technology for Cancer Research (ITCR) 2022 Annual Meeting

September 12-15, 2022
Eric P. Newman Education Center (EPNEC)
Washington University School of Medicine
St. Louis, MO

GENERAL INFORMATION

Scientific Sessions (All registered attendees WUSM/ITCR)	**September 13-14** **Breakfast/registration**—Atrium lobby, 7:30am-8:40am daily **Presentations/Lighting talks**—Main auditorium **Posters**—Great rooms A & B
Workshops (WUSM/ITCR workshop registrants)	▪ **9/12**, 3-5pm, Imaging Data Commons (IDC) Workshop, Seminar Room A ▪ **9/15**, 9am-12pm, ITCR Training Network (ITN) Workshop, Seminar Room A
Social Activities (For ITCR registrants/awardees)	**9/12**, 7-9pm, Evening meet-up at Chase Club (located at the Chase hotel) **9/13**, 6:30-9:30pm, Moulin Lounge, Rhone Rum Room **9/14**, 6:30-9:30pm, Reception at Missouri Botanical Garden
Transportation Available September 13-14 (For ITCR registrants/awardees)	▪ From Chase hotel to EPNEC—shuttle starting at 7am, rotating until 10am. ▪ From EPNEC to Chase hotel—~5pm following poster sessions, rotating for two pick-ups. ▪ To/from the Chase hotel and social activity venues—bus will pick at the Chase hotel at ~6pm both days; depart to the Chase hotel at ~9:30pm **NOTE:** For those staying at AC Hotel, please use the Chase hotel stop. Shuttles/buses cannot stop at the AC hotel. Walk time between the Chase and the AC hotel is ~3-4min; between the Chase and EPNEC it is ~15min.
Sponsors	Siteman Cancer Center (SCC) Department of Radiology Department of Medicine, Division of Oncology
Host/Scientific Committees	Kooresh Shoghi, PhD, WUSM Obi Griffith, PhD, WUSM Malachi Griffith, PhD, WUSM Jennifer Beane, PhD, Boston U. Daniel Marcus, PhD, WUSM Jeremy Goecks, PhD, OHSU Harry Hochheiser, PhD, U. of Pitt. Lydia Kavraki, PhD, Rice University Juli Klemm, PhD, NCI
Administrative Support	Ms. Nancy Reidelberger Ms. Melinda Dioneda Ms. Jeanette Freant
Social Media	**Twitter:** Please use *#ITCR2022*, *#NCIITCR, and @NCIDataSci* when referring to the meeting.
Question & answer during the meeting	Virtual (and in-person) participants can ask questions via Slido.com, code #ITCR2022 or use the link: https://app.sli.do/event/4hH6ixgr8XtFXkZz9tBD77

Keynote Presentation—Sheila A Stewart, PhD

Introduction by Dr. Timothy Eberlein; Director, Alvin J. Siteman Cancer Center; Spencer T. and Ann W. Olin Distinguished Professor; Washington University School of Medicine

Dr. Stewart is the Gerty T. Cori Professor of Cell Biology and Physiology in the Department of Cell Biology and Physiology and Medicine at Washington University in St. Louis, the Vice Chair of Cell Biology and Physiology, and the Associate Director for Basic Science at the Siteman Cancer Center. She received her Ph.D. in Microbiology and Immunology from UCLA in 1997 where she studied HTLV and HIV biology and completed her postdoctoral fellowship in Cancer Biology at the Whitehead Institute at MIT in Robert Weinberg's laboratory. In her own laboratory, Dr. Stewart has studied telomere dynamics and spent the majority of her time trying to understand how age-related changes in the tumor microenvironment impact tumorigenesis. Her laboratory has shown that aged stromal cells, similar to cancer associated fibroblasts, express a plethora of p38MAPK/MK2-dependent pro-tumorigenic factors and has developed murine models to explore the role senescent stromal cells play in the preneoplastic and premetastatic niches. More recently, her group has begun to explore how some of these same changes contribute to therapy-induced comorbidities. In addition, the laboratory is examining how age-related changes in the premetastatic niche facilitate tumor cell seeding, dormancy and outgrowth and how these changes alter the local immune response to facilitate tumor cell proliferation.

Patient Advocate—Johanna Kenney, EdD

Introduction by Dr. Richard Wahl; Director, Mallinckrodt Institute of Radiology (MIR); Chair, Department of Radiology; Elizabeth E. Mallinckrodt Professor of Radiology; Washington University School of Medicine

Johanna Kenney, EdD, has over twenty years of experience as an educator and learning and development specialist working for corporate, academic, and military institutions. She received her Doctor of Education in Education Technology from the University of Florida in 2014 through an online program designed for working adults. In 2013 she was diagnosed with breast cancer at the age of 40, which means she researched and wrote much of her dissertation in an infusion room. Since then, she has committed herself to several patient advocacy and patient support campaigns, including being a peer mentor for After Breast Cancer Diagnosis, completing Project LEAD advocacy training, and participating as a DoD BCRP Consumer Reviewer. She is currently the newest member of the NCI Technology Research Advocacy Partnership (NTRAP) team.

ITCR 2022 PROGRAM

PRE-MEETING (Monday, September 12th)

Start	End	Description/Location		Presenter
3:00 PM	5:00 PM	Workshop	Imaging Data Commons (IDC) Workshop, Seminar Room A	ANDREY FEDOROV

SCIENTIFIC SESSIONS

Tuesday, September 13th

Question & answer in-person or via Slido.com, enter code #ITCR2022 or use the link: https://app.sli.do/event/4hH6ixgr8XtFXkZz9tBD77

Start	End	Type	Description	Moderator/ Presenter
7:15	8:40		**Breakfast/Registration:** Atrium lobby, starting at 7:30am **Poster set-up:** Great Rooms A & B (see poster list for poster numbering)	
8:40	9:00	Welcome/ Introduction	KOORESH SHOGHI JULI KLEMM TIMOTHY EBERLEIN	
9:00	**9:55**	**Keynote presentation**	**Leveraging the tumor stroma to enhance immunotherapy** Sheila Stewart, PhD Gerty T. Cori Professor of Cell Biology and Physiology Washington University School of Medicine	
9:55	**10:40**	**Theme #1**	**Deep Learning Technologies in Computational Imaging (10min + 3min Q&A)**	**DANIEL MARCUS**
		Talk 1	STAN-CT: Standardization and Normalization of CT images for Lung Cancer Patients	JIN CHEN
		Talk 2	The Federated Tumor Segmentation (FeTS) platform: An intuitive tool facilitating secure multi-institutional collaboration	SPYRIDON BAKAS
		Talk 3 (Virtual)	Distributed Learning of Deep Learning Models for Cancer Research	DANIEL RUBIN and JAYASHREE KALPATHY-CRAMER
10:40	11:10	Break	Atrium lobby	
11:10	**12:05**	**Theme #2**	**Multi-Modal Data Integration (10min + 3min Q&A)**	**JULI KLEMM**
		Talk 1	Integrative genomic and epigenomic analysis of cancer using long read sequencing	MICHAEL SCHATZ
		Talk 2	Informatics for Functional Integration of Heterogeneous Cancer Genome and Transcriptome Sequencing Data	VAKUL MOHANTY and YUKUN TAN
		Talk 3	Enhance UCSC Xena: extend interactive visualization to ultra-large-scale multi-omics data and integrate with analysis resources	JING CHUN ZHU

		Talk 4	PROTEAN-CR: Proteomics Toolkit for Ensemble Analysis in Cancer Research	LYDIA KAVRAKI
12:05	**1:05**	**Lunch**	Atrium lobby	
1:05	**2:00**	**Theme #3**	**Cancer Genomics Analysis Platforms (10min + 3min Q&A)**	**OBI GRIFFITH**
		Talk 1	GenePattern and GenePattern Notebook: Integrative Omic Analysis for Cancer Research	MICHAEL REICH
		Talk 2 (Virtual)	Advanced development of the Cancer Dependency Map portal (DepMap.org)	MIKE BURGER
		Talk 3 (Virtual)	The Integrative Genomics Viewer (IGV) for Cancer Research	JAMES ROBINSON
		Talk 4 (Virtual)	Overture: A Multi-Scale Data Platform for Cancer Genomics Research	CHRISTINA YUNG
2:00	**2:55**	**Theme #4**	**Co-Clinical Data Portals (10min + 3min Q&A)**	**KOORESH SHOGHI**
		Talk 1	Advancing method benchmarking and data sharing through crowd-sourced competitions in cancer research	JAMES EDDY
		Talk 2	Advancing Our Understanding of Cancer and the Human Microbiome with QIIME 2	GREG CAPORASO
		Talk 3	Patient Derived Cancer Model (PDCM) Finder	ZINA PEROVA
		Talk 4	Development of an Open-Source Preclinical Imaging Informatics Platform for Cancer Research	ANDREW LASSITER
2:55	3:20	Break	Atrium lobby	
3:20	**3:30**	**Talk**	**ITCR Training Network (ITN) overview**	**CARRIE WRIGHT**
3:30	**4:05**	**Lightning talks**	1. Ujjwal Baid, 2. Hyemin Um/Thomas DeSilvio, 3. Intekhab Hossain, 4. Jeremy Goecks, 5. Alexander T. Wenzel, 6. Ino de Bruijn	**JULI KLEMM**
4:05	4:10	Poster logistics		**OBI GRIFFITH**
4:10	5:10	Poster session	Odd numbered posters, Great rooms A & B	
5:10		Bus will pick-up at shuttle circle adjacent to EPNEC		
6:30	9:30	Social Activity	Moulin Lounge, Rhone Rum Room (Bus pick-up at hotel at 6:00pm)	

Wednesday, September 14th

Question & answer in-person or via Slido.com, enter code #ITCR2022 or use the link: https://app.sli.do/event/4hH6ixgr8XtFXkZz9tBD77

Start	End	Type	Description	Presenter
7:30	8:40	Breakfast	Atrium lobby	
8:40	8:50	Welcome/ Introduction	Welcome and introduction of patient advocate, Jo Kenney, by Dr. Richard Wahl	
8:50	**9:00**	**Talk (Virtual)**	**Presentation by patient advocate**	**JOHANNA KENNEY**
9:00	**9:55**	**Theme #5**	**Medical Informatics (10min + 3min Q&A)**	**HARRY HOCHHEISER**
		Talk 1 (Virtual)	Semi-supervised Algorithms for Risk Assessment with Noisy EHR Data	CHUAN HONG
		Talk 2	Optimizing the Population Representativeness of Older Adults in Cancer Trials	YI GUO
		Talk 3 (Virtual)	Extraction of Symptom Burden from Clinical Narratives of Cancer Patients using Natural Language Processing	MELIHA YETISGEN
		Talk 4	Cancer Deep Phenotype Extraction from Electronic Medical Records	HARRY HOCHHEISER
9:55	**10:50**	**Theme #6**	**Computational Genomics (10min + 3min Q&A)**	**JEREMY GOECKS**
		Talk 1 (Virtual)	Cancer-specific gene set testing	ROBERT FROST
		Talk 2	Utilizing Bayesian modeling to improve mutational signature inference in large-scale datasets	JOSHUA CAMPBELL
		Talk 3	Advanced development of Lancet, an emerging tool for complex variant calling in cancer genomics	GIUSEPPE NARZISI
		Talk 4	Software and algorithms for elucidating the structure, function, and evolution of extrachromosomal DNA	VINEET BAFNA
10:50	11:20	Break	Atrium lobby	
11:20	**12:15**	**Theme #7**	**Pharmacology and Liquid DNA (10min + 3min Q&A)**	**LYDIA KAVRAKI**
		Talk 1 (Virtual)	Data-driven QSP software for personalized colon cancer treatment	LEILI SHAHRIYARI
		Talk 2	An informatics bridge over the valley of death for cancer Phase I trials of drug-combination therapies	LANG LI
		Talk 3 (Virtual)	cfSNV: a software tool for the sensitive detection of somatic mutations from cell-free DNA	WENYUAN LI
		Talk 4	Predicting transcriptional signatures and tumor subtypes from circulating tumor DNA	GAVIN HA
12:15	12:20	Elevator Pitch	Winners of elevator pitch competition	BRITTANY MCKELVEY
12:20	1:20	Lunch	Atrium lobby	

1:20	**2:15**	**Theme #8**	**Single Cell and Spatial Genomics (10min + 3min Q&A)**	**JENNIFER BEANE**
		Talk 1	Identifying molecular signatures of tumor-immune interactions from spatial transcriptomics using latent space factorization	ATUL DESHPANDE
		Talk 2	Inferring cell state tumor microenvironment maps by integrating single-cell and spatial transcriptomics	ITAI YANAI
		Talk 3 (Virtual)	Informatics tools to analyze and model whole slide image data at the single cell level	GUANGHUA XIAO
		Talk 4	scDECO: A novel statistical framework to identify differential co-expression gene combinations systematically using single-cell RNA sequencing data	YEN-YI HO
2:15	**3:10**	**Theme #9**	**ImmunoOncology (10min + 3min Q&A)**	**MALACHI GRIFFITH**
		Talk 1	Informatics tools for identification, prioritization and clinical application of neoantigens	MALACHI GRIFFITH
		Talk 2	Curve-free phase I/II clinical trial designs for molecularly targeted agents and immunotherapy	YONG ZANG
		Talk 3	Computational approaches to unravel immune receptor sequencing for cancer immunotherapy	LI ZHANG
		Talk 4 (Virtual)	The cancer epitope database and analysis resource	BJOERN PETERS
3:10	**3:45**	**Lightning talks**	**1. Huiming Xia, 2. Alexander Getka, 3. Han Liang, 4. Guilherme Del Fiol, 5. Brian Neal, 6. Bassel Ghaddar**	**JULI KLEMM**
3:45	4:05	Break	Atrium lobby	
4:05	5:05	Poster session	Even numbered posters, Great rooms A & B	
5:05		Bus will pick-up at shuttle circle adjacent to EPNEC		
6:30	9:30	Social Activity	Reception at Missouri Botanical Garden (Bus pick-up at hotel at 6pm) Announce poster award winners	

POST-MEETING (Thursday, September 15th)

Start	End	Type	Description	Presenter
8:30 AM	9:00 AM	Breakfast	Outside Seminar Room A	
9:00 AM	12:00 PM	Workshop	ITCR Training Network (ITN) Workshop Seminar Room A	CARRIE WRIGHT

Poster List

Poster presenters—please hang your poster by the denoted poster number.

**denotes poster with lighting talk at the indicated day/presentation # (in parenthesis under the title heading).*

Poster #	Presenting author	Title
1	Prateek Prasanna / Chao Chen	Topology-Informed Tumor-Peripheral Vasculature Annotation and Analytics for Lung Cancer
2*	Hyemin Um / Thomas DeSilvio	RadxTools for assessing tumor treatment response on imaging (Lighting talk, Day 1, #2)
3	Mauricio Menegatti Rigo	PROTEAN-CR: A Proteomics Toolkit for Ensemble Analysis in Cancer Research
4	Romanos Fasoulis	APE-Gen2.0: Conformational ensemble generation of peptides bound to MHC receptors
5	Tracy Nolan	Adding a Quality Score to Data Collections in The Cancer Imaging Archive (TCIA)
6	Tahsin Kurc	ITCR Developed Enhancements to The Cancer Imaging Archive (TCIA)
7	Dexter Pratt	NDEx Pathway Analysis Embedded in cBioPortal
8	Michael Waters	Generative Adversarial Neural Networks and Population Data Enhance Pathway Analysis of Rare Treatment Resistant HPV+ Head and Neck Squamous Cell Carcinoma.
9*	Bassel Ghaddar	Reconstructing physical cell interaction networks from single-cell data using Neighbor-seq (Lighting talk, Day 2, #6)
10	Shuang Yang	Patients Adherence to Screening Recommendations for Lung Cancer
11	Jennifer Beane / Vijaya Kolachalama	Representation learning for histological profiling of lung squamous premalignant lesions and tumors
12	Yong Zang	Modified isotonic regression based phase I/II clinical trial designs identifying optimal biological dose
13	Matthew Inkman	Machine and Deep Learning Methods for Recurrence Prediction from Positron Emission Tomography Imaging in Cervical Cancer
14*	Brian Neal	Specialized Software for Network Analysis of the Immune Repertoire (Lighting talk, Day 2, #5)
15*	Huiming Xia	Computational prediction of MHC anchor locations guide neoantigen identification and prioritization (Lighting talk, Day 2, #1)
16	Megan Richters	pVACsplice: Predicting neoantigens from tumor-specific alternative splicing events derived from cis-acting regulatory mutations using whole exome and RNA sequencing data
17	Susanna Kiwala	Continuing improvements to pVACtools to meet the challenges of neoantigen identification and prioritization
18	Jacob Clark	Interactive Risk Maps for Cancer Control Effortsts
19*	Ujjwal Baid	The Federated Tumor Segmentation (FeTS) Initiative: The Largest To-Date Real-World Federation Focusing on NeuroOncology (Lighting talk, Day 1, #1)
20	Fusheng Wang / Jun Kong	Three-Dimensional Tissue Based Biomedical Research
21	Yukun Tan	novoBreak-rna: local assembly for novel splice junctions and fusion transcripts detection from RNA-seq data

22	Lei Wang	A critical data source survey for cancer drug combination phase I trial design: physician and biostatistician perspectives
23	Lei Wang	An informatics bridge to improve the design and efficiency of Phase I clinical trials for anti-cancer drug combinations
24	Lei Wang	Drug Combo: An Informatics Bridge for Cancer Drug Combination Research
25	Anja Conev	3pHLA-score: improved structure-based peptide-HLA binding affinity prediction
26	Andrey Fedorov	NCI Imaging Data Commons
27	Deepa Krishnaswamy	Prediction Of Body Part Regions Utilizing the NCI Imaging Data Commons Platform
28	Mikhail Milchenko	Accurate segmentation of vestibular schwannoma on post-contrast MRI
29	Joseph Perl	The TOPAS TOol for PArticle Simulation - providing a maximally flexible software tool to users with vastly varying levels of computing expertise
30	Helga Thorvaldsdóttir	The Molecular Signatures Database Revisited: Extending Support for Mouse Data
31*	Jeremy Goecks	Analyzing Single-cell Spatial and Multimodal Tumor Atlases with Galaxy (Lighting talk, Day 1, #4)
32	Mary Goldman	Visualization and analysis of cancer genomics data using UCSC Xena
33	Jens Luebeck	AmpliconClassifier detects the mechanisms of focal genome amplifications in cancer
34*	Alexander Getka	Cancer Imaging Phenomics Toolkit (CaPTk): A Software Platform Leveraging Quantitative Radio(geno)mic Analytics for Computational Oncology (Lighting talk, Day 2, #2)
35*	Guilherme Del Fiol	GARDE: A Standards-Based Clinical Decision Support Platform for Identifying and Managing Patients who are Eligible for Genetic Testing for Hereditary Cancers (Lighting talk, Day 2, #4)
36	David A Hanauer	EMERSE text processing tool for clincal documents
37	Alexej Abyzov	Analysis of somatic mutations in 131 human brains reveals aging-associated hypermutability
38*	Alexander T. Wenzel	Data driven refinement of gene signatures for enrichment analysis and cell state characterization (Lighting talk, Day 1, #5)
39*	Ino de Bruijn	cBioPortal for Cancer Genomics (Lighting talk, Day 1, #6)
40*	Intekhab Hossain	Prior-informed NeuralODEs to discover sparse regulatory dynamics from time-course gene expression data (Lighting talk, Day 1, #3)
41	Cliff Meyer	MIRA: Joint regulatory modeling of multimodal expression and chromatin accessibility in single cells
42	Arpad Danos	Addition of multi variant Molecular Profiles to the CIViC data model
43	Arpad Danos	Development of Oncogenicity Assertions and incorporation of novel guidelines into CIViC
44	Ajay Venigalla	Aggregation of FGFR variant data and functional evidence into CIViC to aid the modification of FGFR specific oncogeniciy classification rules by the ClinGen FGFR ClinGen Somatic Cancer Variant Curation Expert Panel (SC-VCEP)
45	Jason Saliba	Reimagining and Enhancing the Clinical Genome Resource (ClinGen) Somatic Cancer Clinical Domain Working Group

46	Mariam Khanfar	Curating Cancer Variants in the CIViC 2.0 Interface
47	Cong Ma	Copy number and tumor clone inference in single-cell and spatially resolved omics data
48	Harry Hochheiser	Cancer Deep Phenotype Extraction from Clinical Medical Records (DeepPhe)
49	Andrew Lassiter	Development of an Open-Source Preclinical Imaging Informatics Platform for Cancer Research
50	Kaushik Dutta	Deep Learning based framework for quantitative estimation of Standard-Count [18F]-FDG PET from in-vivo preclinical Low-Count [18F]-FDG PET images
51	Hu Chen	DrBioRight: a next-generation analytics platform for analyzing omics data
52*	Han Liang	An Expanded Quantitative Protein Expression Atlas of Human Cancers (Lighting talk, Day 2, #3)
53	Wei Liu	TCPA: An open access resource for cancer functional proteomics data
54	Siavash Raeisi Dehkordi	OM2BFB: Detecting and elucidating Breakage Fusion Bridge structures in cancer genomes using Optical Mapping data

We are grateful to our sponsors—Siteman Cancer Center (SCC), Mallinckrodt Institute of Radiology (MIR), and the Department of Medicine, Division of Oncology—for their generous support of the meeting.

SITEMAN CANCER CENTER

Washington University in St. Louis School of Medicine | BJC HealthCare

Washington University in St. Louis School of Medicine

MIR Mallinckrodt Institute of Radiology

Washington University in St. Louis School of Medicine

John T. Milliken Department of Medicine
Division of Oncology

Elevator Pitches

Submitter	Pitch	Themed session or poster
Daniel Rubin	We developed computer software to enable leveraging images across many hospitals to build artificial intelligence models that can help doctors better evaluate cancer patients. Our software performs computations at each hospital locally and generates a composite model, without sharing any patient data. We applied our methods for training models to evaluate images of breast cancer and brain cancer patients. Our software is expected to enable developing more broadly applicable artificial intelligence models that will improve the care of cancer patients	Theme 1: Deep Learning Technologies for Computational Imaging
Jin Chen	We develop computer software to harmonize CT images of lung cancer patients. Lung cancer studies require analyzing CT images and measuring texture features of tumors in the images. But noises often interfere with tumor texture features, which are mainly caused by what CT machines are used and how they are set. Our software automatically reduces discrepancies caused by the noise while preserving the texture feature of tumors. Our software improves CT image quality, minimizes image discrepancy, and accelerates cancer research.	Theme 1: Deep Learning Technologies for Computational Imaging
Mary Goldman	UCSC Xena is an online tool to explore different types of cancer genomic and clinical data. Researchers often use us to find a new pattern in the data that they can validate in the lab and vice versa. These data patterns can vary considerably; an example would be noticing a particular mutation in a patient's tumor that is associated with a higher chance of patient survival. Researchers can explore large public datasets, their own datasets, or both.	Theme 2: Multi-Modal Data Integration; Poster 32
Yukun Tan	We developed novoBreak-rna, a method to detect novel splice junctions and fusion transcripts from RNA-seq data. Our method is based on the local assembly model, which combines the efficiency of alignment-based approach with the sensitivity of de novo whole transcriptome assembly approach. The feature of high sensitivity of our method will improve our understanding of cancer mechanisms, particularly due to splicing-factor alterations or gene fusions, and development of RNA-based therapies.	Theme 2: Multi-Modal Data Integration; Poster 21

Helga Thorvaldsdottir	The Integrative Genomics Viewer (IGV) displays visual representations of genomic data, allowing users to identify underlying patterns and potential anomalies. Given the enormous amount of data generated by research studies and in the clinic today, visualization is often a key element in gaining insight into the genomic basis and mechanisms of disease. These insights can help to develop hypotheses for further study, guide treatment, and point the way to new therapeutic targets.	Theme 3: Cancer Genomics Analysis Platforms
Helga Thorvaldsdottir	GenePattern is a bioinformatics application that helps biomedical researchers as they tackle complex problems at the forefront of cancer research, including patient diagnosis and prognosis, identification of new drug targets, and understanding disease mechanisms to improve patient care. In particular, GenePattern provides a simple web-based interface to hundreds of analysis methods and provides tools to help users access and run analyses, and keep track of their work and manage the results.	Theme 3: Cancer Genomics Analysis Platforms
Greg Caporaso	The human microbiome, the trillions of microorganisms living in and on our bodies, is a new frontier in cancer research, and QIIME 2 is the most widely used software platform in human microbiome research. Enabled by the technologies used to sequence the human genome, we're just beginning to understand how our microorganisms impact cancer development, treatment, and recovery. QIIME 2 enables scientists to learn from the vast quantity of data involved in a cancer microbiome study.	Theme 4: Co-Clinical Data Portals
James Eddy	We want to enable the scientific community to develop and stress-test cutting edge cancer research tools through bake-off style competitions. These competitions, or challenges, inspire and attract different skill sets to tackle tough biomedical questions using computational methods and algorithms — including those funded by ITCR. We are building frameworks for researchers to easily discover, securely host, and successfully participate in challenges. In addition to these frameworks, we will provide training and best-practice guidelines to ensure challenges remain rigorous and impactful.	Theme 4: Co-Clinical Data Portals

Harry Hochheiser	Researchers often review cancer patient records in the hopes of understanding the practical impact of treatments and procedures on outcomes. As much of this information is found only in clinical notes, this review is labor-intensive and expensive. Our DeepPhe tool uses models of written language to extract information about patients and their treatments, summarizing details from many notes to build a history of the individual's treatment. Results are stored in a database and can be visualized in our web-based tool.	Theme 5: Medical Informatics; Poster 47
Giuseppe Narzisi	We developed a computer software, named Lancet, to detect somatic mutations (DNA alterations) in cancer patients. These (very rare) alterations can (but not always) cause cancer and must be identified with high accuracy. Differently from most methods, Lancet jointly analyzes the genetic data of the tumor (and matched normal) sample using "local graphs", where (similarly to subway maps) mutations are automatically identified as unique paths to the tumor. This approach substantially increases the ability to discover DNA insertion and deletion	Theme 6: Computational Genomics
Vineet Bafna	Nearly a third of all cancer patients, in a multitude of tumor types, carry large, circular DNA, that have rearranged and separated from chromosomes to form extrachromosomal DNA (ecDNA) structures. Tumor promoting genes on ecDNA have increased copy numbers and are unusually active, resulting in increased pathogenicity and poor outcomes. We are developing software tools for detection and characterization of ecDNA. These tools will enable scientists to better understand ecDNA biology and expose vulnerabilities that can be exploited for therapy.	Theme 6: Computational Genomics; Poster 33, 54

Leili Shahriyari	We have developed a data driven mathematical model to analyze the dynamics of the important players such as immune and cancer cells in osteosarcoma tumors with distinct immune patterns in the absence and presence of the most common chemotherapy drugs. Osteosarcoma is a rare type of cancer with poor prognoses. Our results indicate that the optimal treatments' start times and dosages depend on each tumor's unique microenvironment, implying the necessity of personalized medicine. We will extend our model to build data driven software that can recommend individual-specific optimal treatments.	Theme 7: Pharmacology and Liquid DNA
Malachi Griffith	We developed a free software package that helps doctors design personalized vaccines to treat cancer. Cancer is driven by DNA mutations that result in some proteins on the surface of cancer cells being different from normal cells. Personalized vaccines can be used to train the immune system of cancer patients to identify these cancer cells and destroy them. Our software helps to identify the proteins that are unique to tumor cells and pick the best ones to include in each vaccine, tailored to each patient's unique tumor.	Theme 9: ImmunoOncology; Poster 14, 15, 16, 17
Nina Blazeska	There is considerable interest in tumor-specific immunotherapies with cancer vaccines, alone or in combination with checkpoint inhibitors or other modalities. However, the field is hindered by the lack of a common repository, to access data regarding the targets of immune responses against cancer, and the specific receptors expressed by immune cells fighting cancer. The Cancer Epitope Database and Analysis Resource (CEDAR) provides such a repository; it also provides tools to analyze cancer-specific immune responses and predict novel immune response targets.	Theme 9: ImmunoOncology

Prateek Prasanna	We develop software to visualize and analyze blood vessels surrounding lung tumors; the morphology of these vessels is closely relevant to cancer progression and response toimmunotherapy. However, characterizing the rich geometry can be highly nontrivial and has not been done before. By combining advanced mathematics with powerful machine learning techniques, we will develop predictive biomarkers from complex vessel structures. Our low-cost non-invasive tools will identify which lung cancer patients could maximally benefit from immunotherapy and improve treatment planning.	Poster 1
Fred Prior	The Cancer Imaging Archive (TCIA) is a high value data publishing resource for NCI. The PRISM project's goal was enhancement of TCIA to meet the requirements of cancer Precision Medicine research. Work focused on two topic areas: digital pathology images, annotations and analysis results and semantic integration of non-image data. We increased TCIA's ability to support open science and cancer research by adding new types of data and the ability to define cohorts of subjects drawn from across TCIA collections.	Posters 5,6
Jennifer Beane-Ebel	Lung squamous cell carcinoma is the second most common subtype of lung cancer. Long before the invasive cancer develops, some cells of the respiratory tract develop changes in their genes that result in morphologic abnormalities. We are using computational approaches known as deep learning to identify, based on digitized images of biopsy tissue, which abnormal regions are likely to progress to cancer. This information will help stratify patients that need intervention to prevent the development of lung cancer.	Poster 11
Jacob Oleson	It is challenging to calculate stable cancer rates for small areas. We developed statistical methods allowing us to show age-adjusted cancer rates for every ZIP-code in a state and how those rates change over time. These rates, as well as probability of risk, are displayed in interactive maps accessible on the web. Such maps allow users to compare risk for cancer where they live to other cities in the state and identify areas with high risk in need of interventions.	Poster 18

Jun Kong	Pathologists use a microscope to examine a human tissue for diagnosis. Exploring tissues manually can be time consuming and subjective. Additionally, some tissue objects can only be observed partially due to arbitrary tissue cutting directions. We create a 3D digital pathology framework to reconstruct a complete 3D tissue volume, and develop software to extract and analyze information from the 3D tissue image data to understand disease patterns. This will allow efficient and quantitative disease diagnosis and predictions	Poster 20
Joseph Perl	We developed a computer program that makes it easier for research scientists to test exciting new ideas for radiation therapy. Through over 100 years of incremental advances, researchers have made this therapy increasingly effective, accurate and gentle for cancer patients. To unlock the creativity of today's researchers, our program lets them simulate treatment ideas on the computer without them needing to know computer programming - like how SimCity or Roblox lets any inspired, creative kid build a whole simulated world.	Poster 29
Jeremy Goecks	We created software that allows any cancer scientist to do complex data analyses on the web. Using our website, cancer scientists can analyze their own data or public datasets using thousands of analysis tools. These tools can be applied to identify distinct molecular features in tumors. Distinct tumor features can then be used to (1) select therapies that a patient’s tumor will respond to or (2) understand how a tumor changes when it becomes resistant to a patient’s current therapy.	Poster 31
David Hanauer	EMERSE (Electronic MEdical Record Search Engine) is software that helps researchers accurately find and extract important patient details like diagnoses, medications, procedures, and complications from clinicians’ clinical notes. EMERSE is easy for non-technical people to use and has similarities to search engines like Google. EMERSE enables health research that was previously inefficient or nearly impossible. As one user recently remarked, "The functionality is very friendly and it saved hours of time during the data collection process. Thank you."	Poster 36

Alexej Abyzov	We have developed a method and software, CNVpytor, for detection of low frequency deletions and duplications in cancer genomes using sequencing data. Discovering such genomic alterations has an important role in cancer research, clinical diagnosis and decision for choosing optimal treatment. CNVpytor can also be used to detect these genomic alterations in the normal tissues and samples, such as blood and saliva. This makes CNVpytor a useful tool for early detection of harmful mutations associated with cancer risk.	Poster 37
Cliff Meyer	Aberrant regulation of gene expression is the underlying cause of cancer. To study gene regulation, cancer research laboratories generate genome-wide maps of regulatory elements. Once these data have been published in focused studies, they are deposited in repositories where they can be difficult to discover and reuse. The Cistrome Data Browser extracts meaningful information about gene regulation from these repositories, then provides analysis and visualization tools to facilitate the effective reuse of globally generated experimental data, and accelerate cancer research.	Poster 41
Helga Thorvaldsdottir	Cancer outcomes are determined, in part, by the specific molecular processes that are driving the growth and survival of a tumor. Our Gene Set Enrichment Analysis (GSEA) software is an easy-to-use analysis tool that allows cancer researchers to determine the processes driving changes in biological samples at a molecular level. The insights enabled by GSEA support better understanding of the biology underlying tumor progression and ultimately help to identify candidate treatments.	Posters 30, 38
Obi Griffith	We developed a free, expert-curated web resource that helps doctors identify the best treatment options for cancer patients based on their unique tumor DNA. Cancer is driven by DNA mutations. Using these mutations to guide patient care requires experts to actively review and summarize the latest scientific advances. Our easily searchable resource coordinates these efforts. The free and open design of our database shares expert cancer mutation knowledge with everyone.	Posters 42, 43, 44, 45, 46

Han Liang	Proteins are the basic biological units and represent most therapeutic targets and biomarkers. We developed a web portal, TCPA, which contains rich information about protein expression from thousands of patient tumors. Through TCPA, researchers only need to click a few buttons to visualize protein expression levels in tumor cells in different formats and compare their patterns with those in normal tissues or across cancer types. It also helps identify which proteins most correlate with treatment response or patient survival time.	Posters 51, 52, 53
Brian Haas	We developed a software toolkit identify various types of mutations related to cancer from expressed genes in cancerous tumors. Depending on the mutations found, different treatment options could be available to target cancer in a personalized way. From transcriptome sequencing data derived from whole tumor tissues or even individual cells, our methods find evidence of genetic mutations, fusions, tumor viruses, and other aberrations. We make our tools easy to run and highly scalable for clinical research applications and diagnostics.	
Brigitte Raumann	Globus software enhances the potential for cancer research breakthroughs by simplifying moving and sharing research data. Globus eliminates barriers encountered by investigators when moving large amounts of data generated by tumor genomics or imaging studies, for example frustrations with corrupted data, slow transfer rates, poor or insecure network connections, and unwieldy cloud storage interfaces. Cancer researchers also rely on Globus to easily share large amounts of data with their collaborators and with the broader research community, all with the appropriate security and patient privacy protections.	

Rachel Karchin	Cancer sequencing projects have identified a very large number of DNA mutations whose importance in cancer is not yet understood. To better understand the impact of these mutations, our team has produced software for computational analysis of cancer mutations that can analyze either a few or up to millions of mutations. This tool is designed to act as a funnel to help researchers to find the small number of mutations whose molecular impact contributes to tumorigenesis, prognosis and treatment selection.	
Vakul Mohanty	We have developed Texomer, a statistical method that performs allele specific integration of bulk DNA (genomic) and RNA (transcriptomic) sequencing data from the same sample. The approach quantifies the degree of concordance between the copy number and expression of each mutation. Texomer also quantifies differences in regulation of variant and wild-type allele to prioritize functional mutations. Integrating information from the genome and transcriptome also allows improved categorization of patient tumors and exploration of their underlying biology.	