OFFICE OF CANCER CLINICAL PROTEOMICS RESEARCH

## Spring 2020 Young Investigator Spotlight Part I: PCC and PTRC

*In this special 2-part series we highlight young investigators in the CPTAC program! In Part I we feature scientists from the Proteomic Characterization Center (PCC) and Proteomic Translational Research Center (PTRC). Read on for how these young investigators are developing and implementing their unique tools! Responses edited for length and clarity.*

*Chia-Feng Tsai, Ph.D.*, is a post-doctoral fellow at the Pacific Northwest National Laboratory with Tao Liu's group. He received his Ph.D. in Chemistry from the National Taiwan University and works on nanoscale proteomics analysis methods.

### How are nanoscale proteomics analysis tools used in your current projects?

I've always had a genuine interest in developing new methods and technologies for nanoscale proteomics analysis and their potential in clinical applications. Clinical samples, however, are often small which limits assessment by current methods that need more input. I aim to develop new approaches to reduce sample input and loss while increasing analytical throughput. First, we developed an integrated sample processing approach for high-throughput and comprehensive phosphoproteomic analysis. In addition, we developed **BASIL** (Boosting to Amplify Signal with Isobaric Labeling) to enable comprehensive analysis of nanoscale samples. We routinely apply these methods to the analysis of human tissues including biopsies and laser micro-dissected tissues.



### Analysis at the single cell level is often desired but can be difficult in research. What tools have you developed to accomplish this?

Mass spectrometry (MS)-based proteomics has great potential for antibody-independent, quantitative proteomic analysis of single cells but ineffective sample processing and MS sensitivity limit exploration. We have demonstrated that the addition of an MS-compatible detergent can significantly minimize sample loss during processing. More recently, we incorporated our **improved BASIL** (iBASIL) strategy into the PNNL-developed platform nanoPOTS, which enables reliable clustering of hundreds of single cells, revealing important biological differences in their resulting protein profiles.

### A hobby of yours is baseball. Which team and player are your favorite?

Baseball is very popular in Taiwan. A former Taiwanese pitcher, Chien-Ming Wang, played for the New York Yankees. He was one of the best starting pitchers for the Yankees in 2006 and 2007, wining 19 games in both seasons because of his hard sinker! Back then I was a fan of the Yankees and I recently became a fan of the Seattle Mariners after I moved to Washington State. *Upper right: Chia-Feng Tsai, Ph.D., with his family (provided by C-F.T). Lower right: Chien-Ming Wang of the 06'-07' New York Yankees.*



---

*David Clark, Ph.D.*, is a Postdoctoral Fellow at the Johns Hopkins University in Hui Zhang's group. He holds a Ph.D. in Molecular Medicine from the University of Maryland with a focus on cancer biology using mass spectrometry-based proteomics. He works on applying proteomics methods to CPTAC cancer types.

### You were the lead author on the ccRCC study published in **Cell**. What was your approach to the project and what did you learn during the study?

The overall goal of our study was to link the genomic alterations that are observed in clear cell renal cell carcinoma (ccRCC) to the functional modules that drive the oncogenic pathobiology. This involved a multi-omics approach for the molecular characterization of ccRCC tumor tissues in order to delineate the dysregulated cellular pathways that contribute to pathogenesis and identify potential novel therapeutic targets. Our study showed that chromosomal translocation is a frequent event in ccRCC, potentially explaining the hallmark chromosome 3p loss and positing that ccRCC may be defined by such rearrangements. Our phosphoproteomic characterization results also identified several drug targets, WEE1, ERK1 and EGFR, that are understudied in ccRCC.

Finally, we found that ccRCC could be classified into four major immune subtypes: CD8+ inflamed, CD8-inflamed, VEGF immune desert and Metabolic Immune Desert. These subtypes were defined by distinct features and predicted to have disparate responses to immune checkpoint and anti-VEGF therapies which would have direct translational application in the clinical setting.

**What are the next steps after this work? What other projects are you working on?**
Currently we seek to validate our results in an independent cohort to identify which protein features are maintained using orthogonal proteomic approaches. Furthermore, we are focusing on evaluating the functional consequences of select kinase inhibitors on cell growth and survival using *in vitro* and *in vivo* ccRCC models. Lastly we are using protein panels based on the four immune types identified with the ultimate goal of pinpointing protein markers for stratifying patients for therapeutic intervention.
I am also working on the role of extracellular vesicles (EVs) in ccRCC by profiling the protein cargo of circulating EVs in ccRCC patient blood samples. This study aims to identify proteoforms associated with ccRCC and the four immune subtypes.
Ultimately, we want to address several areas in ccRCC/extracellular tumor biology such as linking tissue protein profiles to circulating EV protein profiles and further support the diagnostic utility of EVs in the clinical setting.

**From your perspective, why is methods development so important in proteomics research?**
Current proteomics approaches, paired with advanced mass spectrometry instrumentation, has enabled a deep interrogation of the proteome associated with several cancer types. The additional insight provided by examination of post-translational modifications (PTMs) further elaborates on the complexity of protein function in disease. Identifying the aberrant protein profiles associated with cancer is only the first step and we need to understand how these aberrations in protein expression impact the larger cellular system.
Proteomic methodologies that elucidate spatial information, protein-protein interactions, and enable detection of minute amounts of sample will greatly expand our understanding of how dysregulated protein abundance can impact cellular function and possibly elucidate therapeutic or diagnostic targets not detected with current technologies. *Upper right: David Clark (provided by D.C.)*

---

*Shayan Avanessian, M.A.,* is a Research Associate at the Broad Institute with Michael Gillette's group. He earned an M.A. in Biotechnology from Boston University and works on data generation and data integrity throughout processing.

**What does your role in proteomic and phosphoproteomic data generation involve?**
Generating high-quality proteomic and phosphoproteomic data begins by using patient samples in a scheme that minimizes bias and variability. From there, we process the samples for data acquisition. This involves lysing and digesting all patient tissue samples, enriching for post-translational modification and running the samples on a mass spectrometry instrument. We keep strict quality control measures throughout these procedures to ensure high-quality data for our analyses and future investigations.
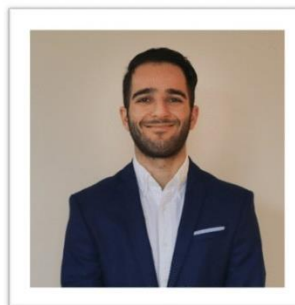
**A goal of yours is to 'minimize bias and pre-analytical variability'. Why is this important and how do you achieve this?**
In experimental design factors like sex, smoking status, country of origin, age and tumor stage influence individual proteomic signature. Therefore, we strategically randomize the initial steps of processing and sample order during a mass spectrometry run so that similar patient variables are not grouped closely together. Our methods ensure that batch to batch variabilities do not propagate to particular patient variables and allow for accurate analysis of patient features not artificially affected.

**You've been in the Boston area for most of your professional career. What would you tell a tourist to check out first?**
I love Boston. For one of the biggest cities in the country, it has a certain hometown charm that is unrivaled (as are our sports teams!). Walking through the Commons on a nice day, sitting by the harbor with a gelato from the North End or appreciating the skyline on the Charles River from Massachusetts Avenue or

Longfellow Bridge will never get old for me. *Upper right: Shayan Avanessian, M.A. (provided by S.A.). Lower right: Boston Commons.*



---

*Shrabanti Chowdhury, Ph.D., is a Postdoctoral Fellow in the Department of Genetics and Genomic Sciences at the Ichan School of Medicine in Pei Wang's group. She earned her Ph.D. in Applied Statistics from the University of California, Riverside in 2016 and works on methodology development in the CPTAC program.*

### What are your main projects within the CPTAC program?

My main project is on predicting platinum refractoriness in high grade serous ovarian cancer (HGSOC), an often lethal gynecological malignancy diagnosed at an advanced stage. The goals of this project include understanding the mechanism behind platinum resistance, developing new methods to predict platinum refractoriness/resistance before the treatment, and detecting novel therapeutic approaches to overcome this resistance for HGSOC patients. Our team performed extensive proteogenomic analysis on HGSOC cell lines, xenograft models and treatment naïve tumors with well-characterized platinum response status. Together with other team members, I used sophisticated statistical/bioinformatics tools to study the molecular signatures driving platinum response. I have also developed a new tool to build high dimensional graphic model based on proteogenomic data to facilitate system-learning and to screen for protein/gene markers driving platinum response in HGSOC.

### What are some of the challenges facing data science and how does your work address them?



The major challenges are the quality, quantity and complexity of the data. Critical decisions based on poor-quality data can have serious consequences where even high-performance algorithms fail to extract meaningful information from such sources. To mitigate this, a data scientist must first perform thorough quality check and rigorous pre-processing before downstream analysis. Additionally, a sufficient quantity of data is necessary to fit model parameters. When sample size is limited, one needs to carefully control false positives and properly interpret results. In the end, given the complexity of proteogenomic data, many simple models may fail to account for different data features, while more sophisticated methodologies are often needed to better characterize the data structures. Part of my research focuses on developing novel computational/statistical methods for integrating different types of -omics data to better understand the biological system underlying complex diseases. For example, the tool described in the previous question is based on statistical models with some unique features and fast computing ability to infer causal association in gene/protein regulatory networks.

In summary, statistical knowledge is fundamental to data science, and serves as the corner stones for study design, variable selection and measurement, data collection and analysis. Our contributions range from novel algorithm development to downstream analysis implementation. At the same time, a data scientist also needs expert domain knowledge of the field to bring new insights to the data visualization and analyses, as well as result interpretation.

### You work with a non-government organization (NGO) in India. What are the main goals of the group?

Education is a basic necessity that can eradicate poverty and promote equality. Yet in India, an aspiring global superpower, only 8% of the relevant age group can afford undergraduate education and above, largely due to financial constraints. **Padakshep** is a non- profit organization formed and powered by a group of like-minded people across the globe who believe in the role of education in shaping a society. Our primary goal is to provide basic educational guidance to a selected group of bright yet financially underprivileged students. We carry out several initiatives to reach a wide spectrum of students and help establish the funding, teaching and

resources to help them fulfill their dreams. *Upper right: Shrabanti Chowdhury, Ph.D. (provided by S.C.). Lower right: The Padakshep group provides resources for education in India.*



---

***Sunil Joshi*** *is an M.D./Ph.D. student at the Oregon Health Sciences University in Brian Druker's group. He earned his bachelor's at the University of California, Berkeley. Sunil has several awards including an NCI-NRSA, The Paul & Daisy Soros Fellowship, and AACR's Scholar-in-Training. Sunil's work focuses on the tumor microenvironment in acute myeloid leukemia.*

### How has being part of CPTAC helped your career development as a future medical scientist? What projects are you working on?

My experience as a member of our CPTAC team has reinforced the importance and value of teamwork. Working with my primary Ph.D. mentor Brian Druker, M.D. and co-mentor, Elie Traer, M.D., Ph.D. has enabled me to collaborate and learn from scientists, physician-scientists and bioinformaticians from across the Oregon Health & Science University and Pacific Northwest National Laboratory. As a budding physician-scientist I enjoy the opportunity to explore the science by actively participating in the planning, executing, and troubleshooting of experiments—essential skills for my future career. Moreover, I am constantly reminded that persistence is key in making discoveries. My project is focused on understanding how the tumor microenvironment influences drug resistance in acute myeloid leukemia (AML). I developed an *in vitro* two-step model of drug resistance to targeted inhibitors observed in patients with AML. Specifically, our model considers drug resistance mediated by the bone marrow microenvironment and acquired mutations leading to relapse. CPTAC has allowed us to leverage global and phospho-proteomics to prioritize signaling pathways conferring drug resistance which we've validated with genetic and biochemical techniques. From this work we designed a targeted proteomics panel to assess clinical samples from our ongoing AML clinical trial. We ultimately hope to identify combinatorial strategies that may improve the durability of currently available FDA inhibitors to circumvent drug resistance.

### How have you worked to increase diversity and inclusion in the STEM fields? What goals do you have for your community efforts?

I am passionate about supporting the success of students from under-served and disadvantaged communities as my parents immigrated to the US with the simple goal of providing their children a better future. At OHSU I have served on the M.D. and M.D./Ph.D. admissions committees and on the medical curriculum development committee. I use these as an opportunity to initiate dialogue that builds on the idea of inclusive excellence. I've also helped establish Oregon's first Minority Association of Pre-Medical Students Chapter, an initiative aimed at providing undergraduate pre-medical students mentorship and resources to craft successful applications. I currently mentor students from through our **Wy'east Post Baccalaureate Pathway Program**, a program dedicated to recruitment and success of trainees identifying as Native American.Through the **NIH BUILD-EXITO Program**, which aims to recruit diverse trainees into the biomedical sciences I worked with Letisha Wyatt, Ph.D. of OHSU to cohost workshops for students. Topics include dealing with micro-aggressions, the "diversity tax", negotiating/advocating for oneself, and preparing successful applications for graduate school. I've also worked with the OHSU Knight Cancer Center to organize the **CURE & Equity Summer Research Internship Program** dedicated to providing underrepresented students laboratory research exposure where for the past two years I've mentored a student in our laboratory. My goals are to create an inclusive learning environment and to raise awareness on how everyone from trainees to faculty members can be involved. I believe my endeavors are a 'first attempt' to address many of the structural barriers that limit access to opportunities for trainees from different walks of life.

**What advice would you give students interested in a medical scientist career?**

I believe that genuine curiosity, commitment as well as a passion for working at the translational interface are much needed. After college, I was a research associate at UCSF for four years which gave me a glimpse of a dual career in medicine and science. I would also recommend seeking mentors for guidance and honest feedback. At OHSU for my Ph.D. I have a mentorship team model with one primary mentor (Brian Druker) and two co-mentors (Cristina Tognon and Elie Traer). Having a team of strong and supportive mentors has been invaluable for my scientific acumen and interpersonal growth. Finally, I believe that taking the time to do multiple research rotations to find your 'research home' is very important during your Ph.D. studies. *Upper right: Sunil Joshi (provided by S.J.). Lower right: Sunil Joshi (left) with CURE high school intern (center) and co-mentor Cristina Tognon (provided by S.J.).*



---

*Shankha Satpathy, Ph.D., is a Scientist at the Broad Institute of MIT and Harvard in Steven Carr's group. He earned his Ph.D. from The Novo Nordisk Foundation Center for Cancer Research, Copenhagen and works on data generation and data analysis in CPTAC projects.*

**You work both on the data generation and analysis aspect of CPTAC projects. What projects are you working on and what has been your approach?**

I have been working on several projects with the PCC and PTRC arms of the CPTAC program. These include leading the non-small cell lung cancer study and prospective breast cancer study with the Baylor College of Medicine team led by Matthew Ellis. Additionally, as part of the Broad-Baylor joint PTRC program, I'm working on developing and applying micro-scaled proteogenomic methods to cancer samples from clinical trials.

My approach to has been to employ state-of-the-art mass spectrometry technology to generate high-quality proteomics data sets and perform integrative proteogenomic analysis using contemporary bioinformatics tools and utilize tools from the PGDAC arm of the CPTAC program. Given the interdisciplinary nature of these projects, I strive to work in a collaborative environment with the ultimate goal of generating resources for the cancer community.

**You helped develop BioTExt and MiProt to perform multi-omics analysis. How do these tools work?**

BioTExt, or Biopsy Trifecta Extractor, is a sample processing method designed to minimize the influence of tumor heterogeneity. In this method an alternate sectioning strategy is applied to single core needle patient biopsies, where breast tissue is removed with a hollow needle. The sections are then pooled together and subsequently used for genomic, transcriptomic and proteomic analyses. MiProt, an optimized microscale workflow, was then developed to circumvent the typically low yields from these biopsies and enables us to produce high-quality and deep-scale proteomics datasets from these samples. The Broad-Baylor team continues to develop and optimize these tools for proteogenomic characterization and apply them to interesting cancer biology questions. Both tools have been published in Nature Communications. *Note: These tools were highlighted in a recent issue of* **[NCI's Cancer Currents](#)**.



**What are your future goals as a proteomics and computational scientist?**

My future goal would be to continue working in the field of clinical proteomics and proteogenomics in an

innovative scientific institution. My ultimate goal is to continue doing meaningful research relevant to human health and disease. *Right: Shankha Satpathy, Ph.D. (provided by S.S.).*

Stay tuned for Part II of this series where we highlight PGDAC Scientists!