MARCH 3-5, 2023

In-person event hosted at the Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Table Of Contents

Background

Introduction

Susan G. Komen®'s Big Data for Breast Cancer (BD4BC) Initiative leverages big data to fuel scientific discoveries, advance research, improve breast cancer outcomes and accelerate the delivery of equitable, patient-focused care. We believe that big data can be used to identify breast cancer disparities and variances in care and to improve methods for aggregating and analyzing clinical, genomic and other sources of data to drive scientific progress and save lives.

Komen, the University of Texas Southwestern Medical Center (UTSW) Lyda Hill Department of Bioinformatics and the UTSW Harold C. Simmons Comprehensive Cancer Center organized the inaugural Breast Cancer Hackathon Challenge, a project funded by Lyda Hill Philanthropies. The Breast Cancer Hackathon Challenge brought together creative individuals and teams with diverse backgrounds to develop new data science approaches and tools to address immediate needs or outstanding challenges in the fight against breast cancer.

By The Numbers

176	Applicants
54	Participants
9	Teams
3	Projects

5 Countries
20 US States
34 Institutions

Projects

Challenge Questions

We issued a request for proposals (RFP) inviting researchers in the breast cancer community to submit challenge questions, along with sources of relevant publicly-available data, focused on significant issues in breast cancer that could be addressed with innovative, big data tools and solutions. The RFP (see Appendix A) was sent to current and former Komen Scholars (https://www.komen.org/breast-cancer-research/komen-scholars/), Komen's Career Catalyst Research grantees, and Komen's Scientific Advisory Board (https://www.komen.org/about-komen/our-people/scientific-advisory-board/).

We were especially interested in challenge questions that incorporated diverse disciplines and fields of study, new technologies, methods and/or leveraged existing resources in innovative ways. We were also interested in questions that were applicable to populations that experience disparities, either by focusing on specific populations or by demonstrating broad and inclusive applications across a diverse population. Below are the challenge questions we received.

Challenge Questions Submitted in Response to RFP
The diploid human genome carries two copies of alleles for each gene (except for the sex chromosomes). The maternal and paternal copies of genomes are both required for normal development and cell differentiation, and therefore are both expressed with the exception of genomic imprinting loci where transcription depends on parental origin. In breast cancer, patients can carry one (heterozygous) or two (homozygous) different mutations on the two alleles. Considerable amount of efforts have been made to study the function of single mutations by computational predictions or by high-throughput experimental approaches. However, systematic interrogation about the diploid functional effect of genomic mutations has been lacking. Many cancer-causal mutations are indeed neomorphic or 'gain-of-function' (GOF). Innovative computational algorithms taking the genome context into account will be critical to better understand breast cancer, aligning with Komen's areas of interest.
Although many studies have focused on identifying single drug targets, it remains enigmatic how to discover more effective and actionable drug combination targets in breast cancer. Combination therapy regimens for breast cancer are promising but have traditionally been non-specific with broad toxicity profiles and developed in an ad hoc manner. There is growing interest in identifying synergistic interactions for combination therapy, but large-scale experimental screening has been technically challenging and expensive given the large number of gene pairs one has to screen. As a result, existing RNA-interference and CRISPR-based screenings have been limited to only a few hundred genes, far from saturating all possible (~4 x 10^8) pairwise interactions in the human genome. To resolve this challenge, efficient in silico screens and prioritization of co-targetable pathways will be critical, enabling more powerful combinatorial therapeutics in difficult-to-cure subtypes of breast cancer.
One of the most critical challenges in breast cancer research is personalized medicine (n = 1): the lack of a reference cohort to categorize the patient and prioritize the best treatment for the malignancy. Precision medicine strives to match patients with the best pharmaceutical regimen for treatment. Existing approaches are limited, as only patients having characteristics that are comparable to others with a positive response to treatment are given the medicine. These methods aim to predict their impact on treatment outcome using DNA biomarkers, genomic and transcriptomic information, protein interaction networks, synthetic lethal interactions among others. However, implementing these models often requires data from a large number of patients and extensive knowledge of exact targets for each drug, limiting their applicability. Resolving this challenge is a critical need of our breast cancer patient community, and aligns well with Komen's mission to prevent and cure breast cancer.
Why is obesity a risk factor for postmenopausal women and a protective factor for premenopausal women? Understanding the endocrine or inflammatory mechanisms by which obesity alters risk in breast cancer could directly lead to behavioral, social or pharmacological interventions with enormous potential to alter risk and prevent breast cancer.

What if breast cancer patients had a digital platform to help them with medication adherence?

What if ALL breast cancer patients had equal access to treatment options and clinical trials?

What if patients with breast cancer truly had digital ownership & understanding of their medical records?

One approach to target advanced breast cancer is to identify and therapeutically target the metabolic vulnerabilities of the key players in the tumor microenvironment. However, this has been very difficult to study because of technological limitations. With recent big data of human breast cancer, especially single cell RNA-seq (scRNA-seq), and the tools to analyze such data, we propose a project to infer the comprehensive metabolic activities within and among major immune and non-immune cell populations in the microenvironment by applying new computational tools to scRNA-seq datasets of human breast cancer. Such tools include scFEA that can estimate metabolic flux and MEBOCOST that can assess metabolic cell-cell communications. This project allows us to answer the question: What are the hyperactivated or silenced metabolic pathways, both shared and uniquely regulated, by key immune and non-immune cell populations in human breast cancer.

Understanding prevalence of detectable ctDNA in early stage HR+ breast cancer (and potential clearance with adjuvant endocrine therapy, or emergence on endocrine therapy) at various time points after completion of primary therapy, and how it related to recurrence.

There are increasing numbers of non-coding RNA being identified, but databases documenting them are under-developed or non-existent. The ability to search for their prevalence, differential expression, synteny among humans and animal models, etc. are all foundational. Such databases would be a cornerstone of multiple research projects. Along with collaborators, we have already identified ancestry differences for ncRNA, which means this effort could also impact issue of racial disparities.

What are the differences in gene expression in metastases versus primary tumors? What are the differences in gene expression in different sites of metastasis? (i.e., potential explanations for organotropism) Metastasis is major cause of death and loss of quality of life. We need to understand what drives, inhibits and modifies metastasis efficiency.

My challenge question is related to reference genome used for genomic analyses of breast cancer genome. Recent studies assembling pan-genome after sequencing of 910 humans of African descent have identified ~300 MB extra genome in African ancestry, which affected 315 distinct protein-coding genes (Nature Genetics 51:30-35). The question then is why are we comparing breast tumors of women of African ancestry to reference genome to determine genomic aberrations in tumors and then determine targeted therapies. Reanalyzing genomic and transcriptomic data of breast tumors of African American with "normal" tissues of African Ancestry is required for us make difference in health disparity.

Can you create a tool that maps all functional impact germline variants and acquired somatic alterations including mutations, copy number changes, and epigenetic alterations present in a cancer (cell and issue) into a biological process map of normal breast cells and tissues? Cancer research focuses on a few dozen "cancer driver" genes and tends to ignore the thousands of other germline and somatic alterations present in every cancer which are assumed to be functionally irrelevant "passengers". We believe that this view is too simplistic. Every cancer has a unique clinical behavior the same way as every person has a unique face. The uniqueness of our faces, and our organs, is due to the combined effect of thousands of polymorphisms that were born with compounded by somatic epigenetic changes acquired during aging. The challenge that we pose is to develop a tool that captures the combined biologic effect of all germline and somatic alterations that a particular cancer harbors.

Should we propose adjuvant pembrolizumab to all patients with triple negative breast cancer following neoadjuvant pembrolizumab and chemotherapy?

What are the transcriptional and epigenetic reprogramming drivers in either tumor cells or in the surrounding/microenvironmental immune cells that cause the breast cancer metastatic progression and therapy-resistant progression? This question is related to the Komen's interest on how to eliminate the metastatic or therapy-resistant breast cancers. If we can identify the different drivers in either tumor cells themselves or in the immune cells surrounding the tumor cells, we can use them as biomarkers or as therapeutical targets.

Review Panel

The challenge questions submitted in response to the RFP were reviewed by a panel of breast cancer experts and patient advocates to evaluate their novelty and potential impact on the breast cancer field. Review panel members were assigned submissions to ensure there were no conflicts of interest during the review process. Listed below are the review panel members.

Sunil Badve
Vice Chair, Department of Pathology
Director, Pathology Cancer Program
Emory University

Gordon Mills
Professor, Department of Cell, Developmental and Cancer Biology
Oregon Health & Science University

Nancy Lin
Associate Professor, Department of Medicine
Harvard Medical School

Elizabeth Morris
Chair & Professor, Department of Radiology
UC Davis

Daniel Stover
Assistant Professor, Department of Medical Oncology
The Ohio State University

Charles Perou
Professor, Department of Genetics
University of North Carolina

David Boone
Assistant Professor, Department of Biomedical Informatics
University of Pittsburgh

Lajos Pusztai
Professor, Department of Medicine
Yale University

Jorge Reis-Filho
Chief, Experimental Pathology Service
Memorial Sloan Kettering Cancer Center

Carla Lloyd
Breast Cancer Patient Advocate
Komen Advocate in Science

Cheryl Jernigan
Breast Cancer Patient Advocate
Komen Advocate in Science

Marian Johnson-Thompson
Breast Cancer Patient Advocate
Komen Advocate in Science

Hanneke Leegwater from team 3B (Drug-a-Gene) presents the team's findings at the end of the Hackathon event. Hanneke traveled from Amsterdam to attend the Hackathon.

Challenge Question Score Sheet

Reviewers evaluated the submitted challenge questions using the quantitative scoring rubric shown below.

Score	Descriptor	Additional Guidance on Strengths/Weaknesses
1	**Outstanding**	Extremely strong with negligible weaknesses
2	**Very Good**	Strong but with numerous minor weaknesses
3	**Good**	Strong but with at least one moderate weakness
4	**Satisfactory**	Some strengths but also some moderate weaknesses
5	**Poor**	Very few strengths and numerous major weaknesses

Criteria	Score
How well does the challenge question relate to breast cancer and Komen's areas of interest (such as risk factors, services and interventions, early detection, new technologies, outcomes, and precision medicine as described on page 2 of the Request for Challenge Questions)?	
How well does the challenge question specifically relate to conquering metastatic and aggressive breast cancers, such as inflammatory breast cancer, and/or eliminating disparities in care and outcomes?	
How would you rate the impact of the projected outcomes on the breast cancer field and/or breast cancer care if the proposed challenge question was answered?	
How would you rate the public availability and appropriateness of the data proposed to address the challenge question?	

"I think I took a lot of hard skills away from this, but making those new connections, learning from one another, and even hearing from people like the patient advocates, has brought a whole new perspective into my research and a renewed excitement to continue working on my own projects at home."
-   Madison Sharp, research/clinical staff, team 2B

Challenge Question Winners

The top five scored challenge questions were selected as winning submissions, and each submitter received a cash prize ($200), concluding the challenge question phase of the Hackathon. The winning challenge questions are paraphrased below in no particular order.

Challenge Question Topic	Submitter
Use of scRNA-seq datasets to identify pathway activation in human breast cancer cell populations	**Xin Lu,** University of Notre Dame
Prevalence of detectable ctDNA after primary therapy and its relationship to recurrence in HR+ breast cancer	**Sara Tolaney,** Dana-Farber Cancer Institute
Differences in gene expression between primary and metastatic breast tumors	**Danny Welch,** Kansas University Medical Center
Development of a mapping tool for genomic variants in normal breast tissue	**Lajos Pusztai,** Yale University
Development of predictive models for combination therapies in breast cancer	**Stephen Yi,** University of Texas at Austin

Project Descriptions

The top five scored challenge questions were considered for development into Hackathon projects. The Komen Hackathon staff organized individual calls with those who submitted the winning challenge questions to further frame, define and prepare these challenge questions into projects. Three projects were ultimately chosen based on several factors, including the ability to define the unique barriers that made the question a challenge, the availability of public data to support working on the challenge question and the resources that were needed to develop creative data-driven solutions.

#PROJECT-1 – SUBMITTED BY LAJOS PUSZTAI, YALE UNIVERSITY
Predict and visualize combined functional effects of germline and somatic alterations in breast cancer

Cancer research focuses on a few dozen "cancer driver" genes and tends to ignore the thousands of other germline and somatic alterations present in every cancer, which are assumed to be functionally irrelevant "passengers". This view is too simplistic. Every cancer has a unique clinical behavior the same way as every person has a unique face. The uniqueness of one's face, and organs, is due to the combined effect of thousands of polymorphisms that one was born with compounded by somatic epigenetic changes acquired during aging. The challenge here is to develop a tool that captures the combined biologic effects of all germline and somatic alterations within a particular breast cancer.

#PROJECT-2 – SUBMITTED BY XIN LU, NOTRE DAME UNIVERSITY
Define and compare metabolic states of cell types found in the breast tumor microenvironment

One approach to target advanced breast cancer is to identify and therapeutically target the metabolic vulnerabilities of key players in the tumor microenvironment (cancer cells, fibroblasts, immune cells, etc.). However, this has been very difficult to study due to technological limitations. With the recent advent of big data in human breast cancer, especially single cell RNA-seq (scRNA-seq), the goal is to infer the comprehensive metabolic activities within and among major cell types in the tumor microenvironment using scRNA-seq datasets of human breast cancer.

#PROJECT-3 – SUBMITTED BY STEPHEN YI, UNIVERSITY OF TEXAS AT AUSTIN

Identify and prioritize personalized drug combinations based on the genomic landscape of breast cancer
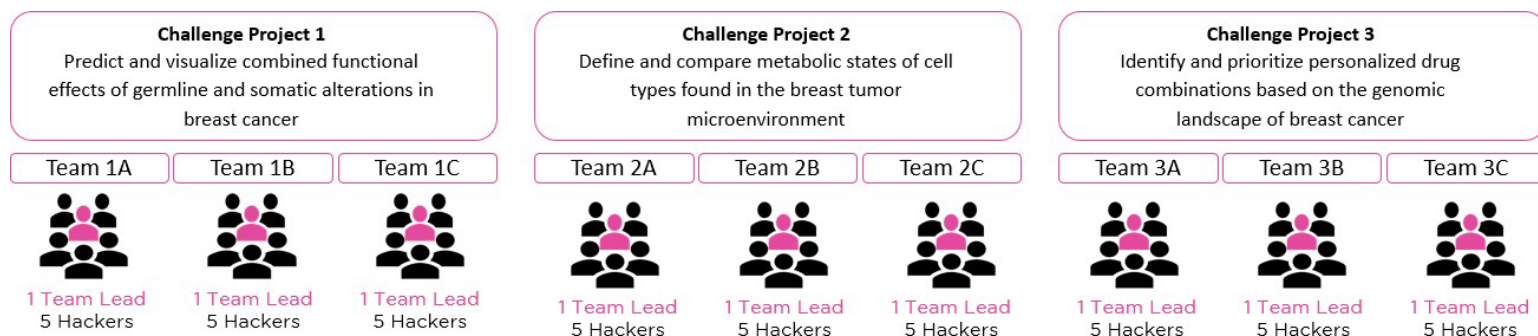
Breast cancer cells exploit multiple pathways to evade the selective pressure of single drugs, promoting therapeutic resistance and clinical relapse. More rational identification of new targets in breast cancer for combination drug regimens is an essential next step in providing long-term clinical benefits. The goal is to accelerate the discovery of combination therapies through integrative, systematic network-based identification of co-occurring genomic alterations in breast cancer patients. The accumulation of omics data from breast cancer increases the feasibility and chance of success of computational analyses to identify synergistic interactions for combination therapy, which has been technically challenging. Therefore, the development of efficient in silico screens and prioritization of co-targetable pathways will be critical, enabling more powerful combinatorial therapeutics for breast cancer.

Team 2A (the Oncology Outlaws) hard at work hacking. Pictured (clockwise from left): Jordan Miner, Connor Vessely, Ethan Moss, Furkan Ozmen, Haiqi Zhu, Xin Lu (Team Lead)

Teams

Based on the number of projects and eligible participants, it was determined that three teams would work on each project, for a total of nine teams. Each team would have six members: one team lead and five participants ("hackers"). Individuals that submitted the winning challenge questions were invited and agreed to serve as team leads. Additional team leads were chosen based on recommendations from these three team leads as well as from the pool of applicants. Team structures are shown below.

Challenge Project 1	Challenge Project 2	Challenge Project 3
Predict and visualize combined functional effects of germline and somatic alterations in breast cancer	Define and compare metabolic states of cell types found in the breast tumor microenvironment	Identify and prioritize personalized drug combinations based on the genomic landscape of breast cancer

Team 1A	Team 1B	Team 1C	Team 2A	Team 2B	Team 2C	Team 3A	Team 3B	Team 3C
1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers	1 Team Lead 5 Hackers

Participant Selection

The request for applications (RFA) to participate in the Hackathon (see Appendix B) was distributed to academic institutions, including those affiliated with former and current Komen Scholars, Komen's Scientific Advisory Board, and Komen's Career Catalyst Research grantees as well as NCI-designated cancer center directors and the NCI's Partnerships to Advance Cancer Health Equity (PACHE) program affiliates. The UTSW Lyda Hill Department of Bioinformatics advertised the opportunity to their past hackathon participants and UTSW students and faculty. While we were anticipating up to 90 applications, a total of 176 individuals applied to participate in the Hackathon. Komen conducted an internal review of the applications, ultimately selecting 45 hackers to participate in the event. Participants were ranked and selected based on the following criteria: coding experience, previous hackathon experience, personal connection to cancer, cancer specific research, computational experience, breast cancer career goals and project specific experience.

Team Formation

In their applications, hackers were able to rank the three projects selected for the Hackathon, based on which one they would prefer working on. That preference was considered during team formation whenever possible. Hackers were sorted into teams based on academic level and institution to ensure a balanced team composition with no institutional overlap. Then, each team was further balanced to maximize diversity among hackers across gender, race and ethnicity. Those selected to participate were formally invited, and upon acceptance, were connected via Discord, an instant messaging social platform, four weeks prior to the event. This allowed participants to connect early with the Hackathon staff and their teammates, build relationships and team comradery, and begin formulating approaches and strategies to tackle their projects. Small incentives ($5 Starbucks gift cards) were used to successfully encourage hackers to engage and communicate via the platform before the event. These incentives also contributed to the competitive spirit of the Hackathon.

Technical Facilitators

Nine technical or "tech" facilitators were recruited by the UTSW Lyda Hill Department of Bioinformatics to assist with the event. Tech facilitators underwent project management training hosted by the Texas Advanced Computing Center. Tech facilitators received an honorarium for their work on the Hackathon. One tech facilitator was assigned to each team and coordinated data and package preparation with USTW's BioHPC team prior to the in-person event. Additionally, Komen staff met with tech facilitators to set expectations and provide support and guidance. During the event, tech facilitators served as project managers for their teams, ironing out technical difficulties and keeping teams on track and focused. Tech facilitators were responsible for meeting with Komen staff as needed to provide updates and resolve issues.

Team Assignments

TEAM 1A: THE PAS FINDERS
Team Lead: Lajos Pusztai
Hackers: Saketh Katta, Mingyu Huang, Allison Yang, Ko-Han Lee, Ilana Schlam

TEAM 1B: THE BIOINFORMAGICIANS
Team Lead: Harikrishna Nakshatri
Hackers: German Corredor, Atharva Charuhas Bhagwat, Aurel Gyorgy Prosz, David Davalos-DeLosh, Tugba Yildiran Ozmen

TEAM 1C: THE DIGITAL TWINS
Team Lead: Mihaela Sardiu
Hackers: Antony Gitau, Lucas McCullum, Matthew Dean McCoy, Siri Kothapalli, Hung-Ching Chang

TEAM 2A: ONCOLOGY OUTLAWS
Team Lead: Xin Lu
Hackers: Ibrahim Furkan Ozmen, Connor James Vessely, Haiqi Zhu, Jordan Miner, Ethan Moss

TEAM 2B: 2B OR NOT 2B
Team Lead: Chi Zhang
Hackers: Vincent Wu, Olwyn Doyle, Madison Sharp, Zack Mudge, Basel White

TEAM 2C: THE METABOLITEWEIGHTS
Team Lead: Maria Chikina
Hackers: Arlen Larry Gyden, Kristoff Kovacs, Chiara Corti, Jeremy Juybari, Samuel Rysdyk

TEAM 3A: THE PINK PANTHERS
Team Lead: Stephen Yi
Hackers: Joshua Gruber, Md Tamzid Islam, Lucas Bartel, Xue Zong, Jingxin Fu

TEAM 3B: DRUG-A-GENE
Team Lead: Nidhi Sahni
Hackers: Tra Ngo, Jessica Kasamoto, Jiaxiang Tao, Siyang Peng, Hanneke Leegwater

TEAM 3C: ONCOMBOAI
Team Lead: Zhijie "Jason" Liu
Hackers: Boyoung Jeong, Daniel Kouchekinia, Jake Shannin, Shuying Li, Yi Hsiao

Breast cancer patient advocates were a key component of the challenge – bringing their unique perspectives and experiences, asking tough questions to get teams thinking and ultimately awarding one team the Advocates Choice Award. From left to right: Marian Johnson-Thompson, Jamie LaScala, Anne Meyn

Event

The Hackathon event was held at the Lyda Hill Department of Bioinformatics on the UTSW campus in Dallas, TX. Before the Hackathon, a two-part event focused on women in computational biology was hosted by the Lyda Hill Department of Bioinformatics. Hackathon participants were invited to travel to Dallas early to attend, and the majority opted to do so. The first part of the event consisted of talks, two given by scientists invited by Komen. Katherine Hoadley of UNC Chapel Hill spoke on team science in cancer genomics, and Minerva Cordero of UT Arlington (a Lyda Hill IF/THEN Ambassador) spoke on advocating for diversity in STEM. The second part of the event consisted of career huddles, hosted by academic and industry mentors. Komen invited several of these mentors, including Allissa Dillman, Aron Parekh, Danny Welch, Jerome Jourquin, Cheryl Jernigan, Katherine Hoadley, Minerva Cordero, and Laura Mydlarz.

After the Hackathon welcome and opening presentations, teams began working on their projects. Throughout the event, experts and patient advocates observed teams, answered questions and provided suggestions. These experts and advocates also contributed to the overall score teams received, as detailed in the "Judging & Awards" section of this report.

To encourage participants to take breaks from coding, Komen organized talks during lunch from the scientific community. Danny Welch and Cheryl Jernigan (judges) gave a talk highlighting the importance for researchers and advocates to work together. David Boone (expert) presented a mentoring program he leads for trainees aimed at encouraging high school students to get into STEM fields. Finally, Gaudenz Danuser (Hackathon host and judge) spoke about novel findings coming from his laboratory as well as from other researchers within his department.

Attendees had various roles in the Hackathon, detailed below.

Role	Description
Participant	Hacker on a team
Team Lead	Faculty member who served as the leader of a team
Technical Facilitator	UTSW staff who assisted teams with data access, project management, and communication before, during, and after the event
Advocate	Patient advocates who visited teams and emphasized the importance of the patient voice in breast cancer research
Expert	Breast cancer research experts who visited teams to offer advice and expertise
Judge	Judges who attended final presentations and scored teams to determine the winners
BioHPC Team	UTSW staff who were responsible for the utilization of BioHPC (high performance computing) and GitHub data/code repository before and during the event

"I have people on my team in computer science, and then we have biologists, engineers, just kind of everyone coming together to create this community."
-   Jordan Miner, graduate student, team 2A

Technical Facilitators

Paul Acosta (team 3C)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Pavel Avdeyev (team 1B)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Stephan Daetwyler (team 2A)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Zach Marin (team 1A)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Hanieh Mazloom-Farsibaf (team 2C)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Ahmed Shalaby (team 3B)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Annie Wang (team 1C)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Sadik Yildiz (team 3A)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Felix Zhou (team 2B)
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Patient Advocates

Anne Meyn
Breast Cancer Patient Advocate
Komen Advocate in Science

Marian Johnson-Thompson
Breast Cancer Patient Advocate
Komen Advocate in Science

Jamie LaScala
Breast Cancer Patient Advocate
Komen Advocate in Science

Experts

David Boone
Assistant Professor, Department of Biomedical
Informatics
University of Pittsburgh

Katherine Hoadley
Assistant Professor, Genetics Computational
Medicine Program
University of North Carolina

BioHPC Team

Devin O'Kelly
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Peng Lian
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Paniz Karbasi
Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Judging & Awards

Judging Scoring and Criteria

Judges awarded the majority (about 60%) of possible scoring points based on the final presentations by teams. The BioHPC team, experts and advocates awarded about 40% of possible scoring points based on teams' performances prior to presentations as described below.

Score	Descriptor	Additional Guidance on Strengths/Weaknesses
5	Outstanding	Extremely strong with negligible weaknesses
4	Very Good	Strong but with numerous minor weaknesses
3	Good	Strong but with at least one moderate weakness
2	Satisfactory	Some strengths but also some moderate weaknesses
1	Poor	Very few strengths and numerous major weaknesses

Key Attributes	
Idea	Did the proposal address the challenge project? Was the proposal innovative?
Implementation and/or Execution	How much progress did the team make towards achieving their goals? Was the solution technically challenging, impressive, or interesting?
Impact	Does the proposed solution provide a meaningful solution for the challenge project? Does the proposed solution move the breast cancer research field forward? Is the solution easy to understand?
Potential	Is it realistic to develop the solution further? If the team was successful in fully developing their solution, how impactful would it be to the breast cancer community?
Presentation	Did the team present their work clearly to address the challenge project? Did the presentation convince you there is a viable product?
Technical Merit	Evaluation of the viability of the code produced by teams *(Provided by the BioHPC team)*
Expert Score	Evaluation of approach, novelty, and potential impact *(Provided by the experts)*
Advocate Score	Evaluation of team cooperation and relevance to breast cancer patients *(Provided by the advocates)*

Judging Panel Members

Gaudenz Danuser
Professor and Chair, Lyda Hill Department of Bioinformatics
UT Southwestern Medical Center

Ariella Hanker
Assistant Professor, Harold C. Simmons Comprehensive Cancer Center
UT Southwestern Medical Center

Cheryl Jernigan
Breast Cancer Patient Advocate
Komen Advocate in Science

Danny Welch
Professor and Chair, Departments of Cancer Biology and Molecular & Integrative Physiology
Kansas University Medical Center

Allissa Dillman
Workforce Development & Continuing Education Adjunct Instructor
Montgomery College

Awards and Prizes

AWARD	CRITERIA	PRIZE (per individual)
Overall Winner	Highest combined average score	$200 Gift Card HydroFlask TileMate Speks Rocketbook Trophy
1st Runner Up	Second highest combined average score	$150 Gift Card HydroFlask TileMate Speks Trophy
2nd Runner Up	Third highest combined average score	$50 Gift Card HydroFlask TileMate Trophy
Advocate Choice Winner	Advocate consensus	Trophy

Hackathon Outcomes

Overall Winner: Team 1B (Bioinformagicians)

The Challenge (#Project-1): Predict and visualize combined functional effects of germline and somatic alterations in breast cancer

The Solution: ShinyMagic (**M**utations **A**nalysis of **G**enetic **I**nformation for **C**ancer with R Shiny)

ShinyMagic App

We developed a web server for interactive analysis and visualization of genetic signatures

The Bioinformagicians developed ShinyMagic, an interactive tool for analyzing and visualizing the link between deficiencies in DNA repair and mutations in breast cancer. The team layered gene expression, somatic mutation status, copy number variation, patient race and patient survival data from The Cancer Genome Atlas (TCGA) as well as PAM50 molecular subtypes, homologous recombination deficiency scores and PARPi7 scores into the ShinyMagic application with a goal of identifying new therapeutic vulnerabilities. ShinyMagic creates heatmap and UMAP visuals from this information and allows users to sort and filter the visuals dynamically. The team demonstrated the utility of ShinyMagic by identifying four homologous repair genes that were differentially expressed in breast cancer patients with PIK3CA mutations compared to those without PIK3CA mutations. Their preliminary survival analysis using a subset of genes identified by Cox multivariate regression suggested that high- and low-risk patients could be identified. The Bioinformagicians were eager to continue to develop their solution after the Hackathon.
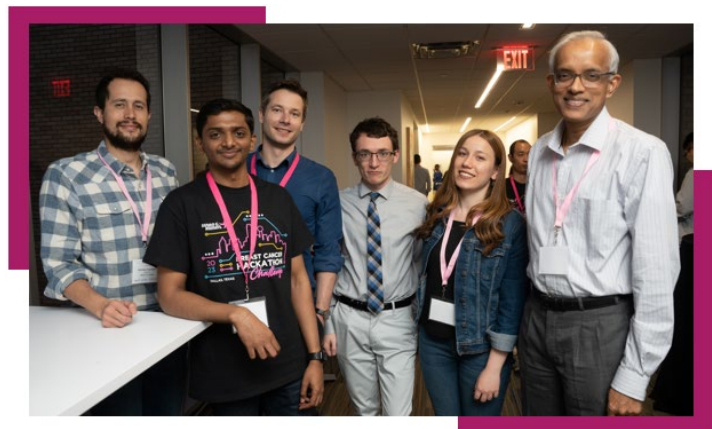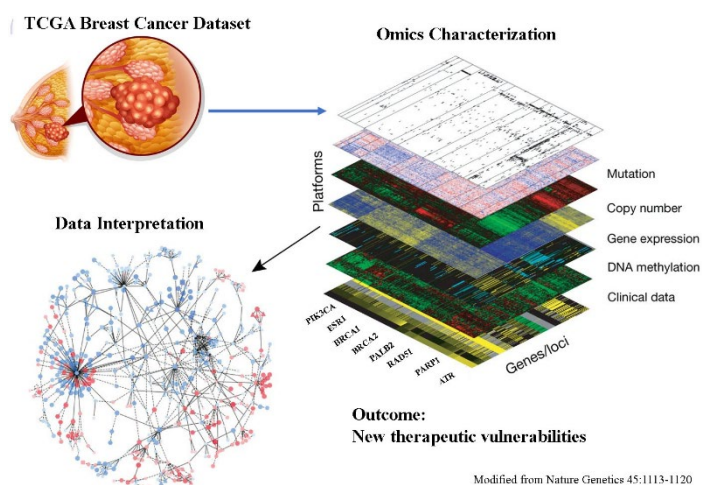
The Bioinformagicians (from left to right): Germán Corredor Prada, Atharva Charuhas Bhagwat, Aurél György Prósz, Davis Davalos-DeLosh, Tugba Yildiran Ozmen, Harikrishna Nakshatri (Team Lead)
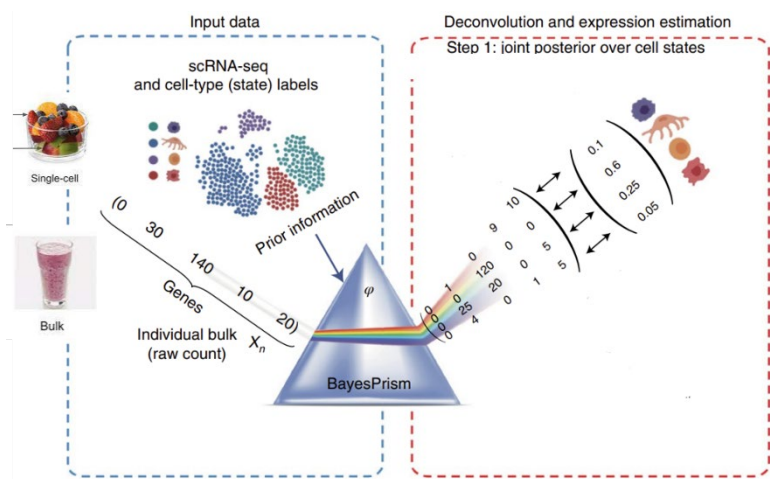
First Runner up: Team 2C (MetaboliteWeights)

The Challenge (#Project-2): Define and compare metabolic states of cell types found in the breast tumor microenvironment

The Solution: MicroFlux

The MetaboliteWeights developed a package, called MicroFlux, that combines bulk RNAseq deconvolution with metabolic flux prediction. MicroFlux accomplishes bulk RNAseq deconvolution using InstaPrism, a light version of the BayesPrism method. InstaPrism deconvolves bulk RNAseq tumor samples into cellular components using single cell RNAseq data as a reference. Essentially, this approach estimates the proportion of individual cell types present in bulk tumor data, similar to estimating the types of individual fruits used to make a fruit smoothie. Finally, MicroFlux uses

Adapted from Chu et al. *Nature Cancer* 2022

ElasticNetCV, a machine learning model, to predict the cell type-specific metabolic activity of each cellular component in a given tumor sample. The MetaboliteWeights demonstrated the clinical potential of their MicroFlux package by showing that high ADP pathway activity in myeloid cells was predictive of better outcomes in triple negative breast cancer.


The MetaboliteWeights (from left to right): Maria Chikina (Team Lead), Kristof Kovacs, Arlen Larry Gyden, Jeremy Juybari, Samuel Rysdyk, Hanieh Mazloom-Farsibaf, Chiara Corti

Second Runner up: Team 3A (Pink Panthers)

The Challenge (#Project-3): Identify and prioritize personalized drug combinations based on the genomic landscape of breast cancer

The Solution: Conditional Variational Autoencoder (CVAE) trained with drug screening data to predict individualized drug response
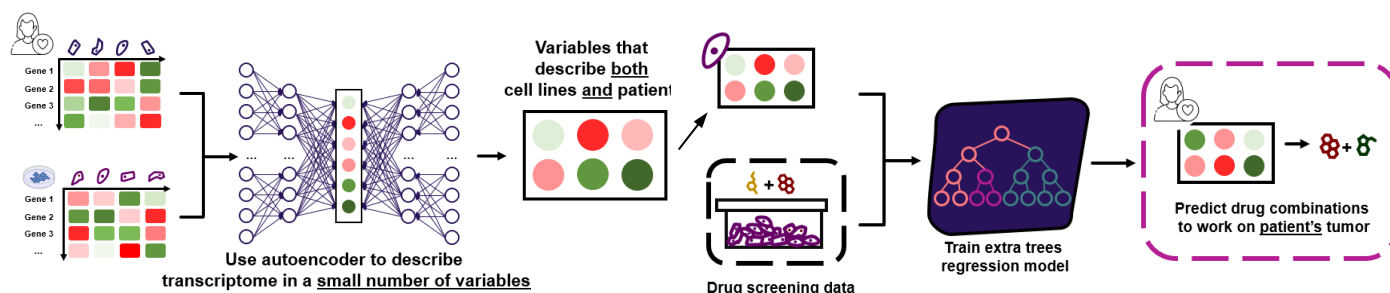
The Pink Panthers developed a framework that compresses massive amounts of single cell RNAseq data from breast cancer cell lines and patient tumors into a small number of variables, using a conditional variational autoencoder. The variables were then trained with drug screening data from breast cancer cell lines to predict individualized combinatorial drug response. The Pink Panthers visualized their results in two-dimensional space using UMAPs and showed that multiple drug combinations can be predicted for individual patients. Their results suggest that breast cancer heterogeneity could be detected and targeted on an individual basis.
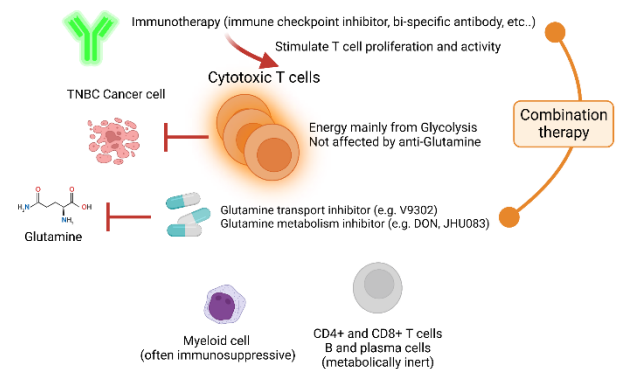

The Pink Panthers (from left to right): MD Tamzid Islam, Joshua Gruber, Jiaxiang Tao, Jingxin Fu, Lucas Bartel, Sadik Yildiz

Advocate Choice Winner: Team 2A (Oncology Outlaws)

The Challenge (#Project-2): Define and compare metabolic states of cell types found in the breast tumor microenvironment

The Solution: Breast cancer subtype- and cell lineage-specific metabolic profiles

The Oncology Outlaws were interested in identifying metabolic targets to improve response to immunotherapy in breast cancer patients. The team used two methods to identify metabolic targets using single cell RNAseq data from breast cancer patients: scFEA and MEBOCOST. Using scFEA, the Oncology Outlaws found an association between high glutamine/glutamate uptake and malignant cells. Building on the increased glutamine uptake in malignant cells, the team used MEBOCOST to predict the molecular sender and receiver of glutamine and predicted that myeloid cells sent glutamine to malignant cells. Based on these

Breast tumor treated by combining anti-glutamine and immunotherapy

findings, the Oncology Outlaws were interested in pursuing the combination of immunotherapy and glutamine transport or metabolism inhibitors in breast cancer. The Oncology Outlaws sought input from the advocates throughout the Hackathon and built developed working relationships that exemplified a true researcher-advocate partnership reflective of patient-centered research.

The Oncology Outlaws (from left to right): Xin Lu (Team Lead), Connor James Vessely, Ibrahim Furkan Ozmen, Ethan Moss, Haiqi Zhu, Jordan Miner

Other Solutions

Team 1A (PAS Finders)

The Challenge (#Project-1): Predict and visualize combined functional effects of germline and somatic alterations in breast cancer

The Solution: Developed a tool to map acquired and inherited genomic alterations to biological pathways and assign a Pathway Alteration Score (PAS) that quantifies the severity of the combined alterations.

Team 1C (Digital Twins)

The Challenge (#Project-1): Predict and visualize combined functional effects of germline and somatic alterations in breast cancer

The Solution: Developed an integrative Digital Twins (iDT) tool to project a new breast cancer patient's features on an existing group to identify the digital twins and predict outcomes.

Team 2B (2B or not 2B)

The Challenge (#Project-2): Define and compare metabolic states of cell types found in the breast tumor microenvironment

The Solution: Developed a multi-step framework to cluster single cells based on metabolic profiles, predict metabolic flux modules unique to malignant tissues and correlate predicted metabolic flux modules with gene expression.

Team 3B (Drug-a-Gene)

The Challenge (#Project-3): Identify and prioritize personalized drug combinations based on the genomic landscape of breast cancer

The Solution: Developed a framework to identify new drug combinations for breast cancer patients by integrating drug combination screens, essential gene screens and patient expression profiles.

Team 3C (OncomboAI)

The Challenge (#Project-3): Identify and prioritize personalized drug combinations based on the genomic landscape of breast cancer
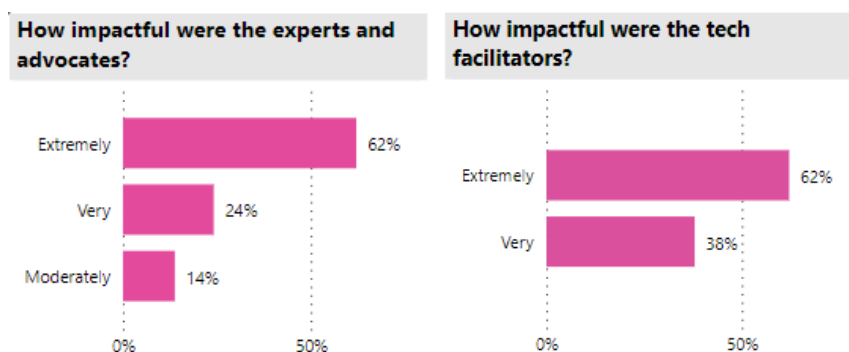
The Solution: Developed a tool based on mutational status, traditional machine learning techniques and a deep learning approach to calculate synergy scores for drug combinations, which could be used to predict combinatorial therapy response.

Post-Event Feedback

All Hackathon participants were given an opportunity to provide feedback on their experience. More than half of participants provided feedback, which is summarized below.

Overall, **97%** of respondents said they would participate in another Komen Hackathon, and **93%** of respondents gave the Hackathon good or excellent ratings. One respondent even commented *"I hope that many hackathons learn from Komen about how [to] organize hackathons which involves genomics/genetics data."* Indeed, **93%** of respondents agreed or strongly agreed that they had the resources they needed during the event.

The Breast Cancer Hackathon Challenge was unique in its incorporation of experts, including patient advocates, and tech facilitators throughout the event. Respondents strongly indicated that having these individuals at the event was impactful.

How impactful were the experts and advocates?

Extremely	62%
Very	24%
Moderately	14%

0%   50%

How impactful were the tech facilitators?

Extremely	62%
Very	38%

0%   50%

"I really enjoyed the atmosphere at the event it was great to interact with people who specialize in areas I was less familiar with. Furthermore, having the breast cancer advocates there really put things into perspective."

We also accomplished our goal of providing the participants with valuable content. The talks given throughout the Hackathon received positive feedback, with **93%** of respondents indicating that the speakers were very or extremely impactful.

"The talks were a great opportunity to take a little break and get inspiration. It was good to be reminded that data doesn't just grow on trees, it comes from actual people."

The positive impact of the Hackathon was also felt by the breast cancer patient advocates in attendance, with Cheryl Jernigan stating that a high point of the event for her was:

"Witnessing how these teams of undergraduate to research professors from (1) VERY different areas of specialties, (2) across the world (even one from Kenya!), and (3) diverse cultures, ages and backgrounds…formed into cohesive teams to develop possible data-enabled solutions to perplexing challenges in diagnosing and treating breast cancer. Their final 7-minute presentations were impressive and very creative."

Acknowledgements

The following people and businesses were integral to the success of the Hackathon.

Rebekah Craig
Department Administrator
UT Southwestern Medical Center

Erica Garza
Administrative Coordinator
UT Southwestern Medical Center

Sena Mevo
Department Project Specialist
UT Southwestern Medical Center

Kimberly Anderson
Financial Analyst II
UT Southwestern Medical Center

Jeon Lee
Assistant Professor, Lyda Hill Department
of Bioinformatics
UT Southwestern Medical Center

Je'aime Powell
Systems Administrator &
Technical Research Design Analyst
Texas Advanced Computing Center

Charlie Dey
Director, Training and Professional
Development
Texas Advanced Computing Center

Suzanne Patterson
Sr. Manager, Events
Susan G. Komen

Lori Young
Manager, Purchasing
Susan G. Komen

Brady Kazar
Director, Program Marketing
Susan G. Komen

Anjali Patel
Sr. Manager, Creative
Susan G. Komen

UT Southwestern Faculty Club
5323 Harry Hines Boulevard
Dallas, TX 75390

BBBop Seoul Kitchen
828 West Davis St.
Dallas, TX 75208

Social Pie Kitchen
5855 Maple Ave
Dallas, TX 75235

Xin_Lu_Notre_Dame 03/06/2023 6:42 PM
I would like to join everyone to celebrate the success of the Hackathon and send deepest appreciation to everyone at Komen and UTSW for organizing such a meaningful and memorable event. I always thought Komen is a beacon lighting up breast cancer research and guiding breast cancer researchers, and this time I am more convinced than ever. Equally warm appreciations to our best patient advocates! You are the reason for us cancer researchers to be passionate about our work everyday. Thank you, thank you, and thank you!

Concluding Statements

The Hackathon aimed to bring together creative individuals and teams with diverse backgrounds to develop new data science approaches/tools to address immediate needs or outstanding challenges in the fight against breast cancer. With this event, we accomplished our goal to utilize breast cancer data to uncover and address problems in breast cancer. We also contributed to the development of skills within the oncology workforce and partnerships with data scientists, and we are currently working to activate the next phase to seed future applications of big data approaches to breast cancer research.

The public response speaks for itself!

CompuMAINE Lab
340 followers
3w · 🌐

Congratulations to current lab members Basel White, Jeremy Juybari, Jordan Miner, and lab alumnus Zach Marin for their participation in the first Susan Komen Breast Cancer Hackathon at the UT Southwestern Medical Center!

Jeremy Juybari and his team placed 2nd in the event, and Jordan Miner and her team won the patient advocate award.

#breastcancer #hackathon

Chiara Corti @CCortiMD · Mar 6
On my way back after @SusanGKomen's 2023 Breast Cancer #Hackathon. Incredible experience! The outcome exceeded every expectation, as our team 2C was awarded the 1st runner-up prize! Thx also to @LHBioinf_UTSW, Lyda Hill and to @stolaney1 for suggesting the application! #BD4BC

Hanieh Mazloom-Farsibaf and 2 others

💬   ↻ 1   ♡ 16   📊 1,390   ⬆️

Jason Liu 03/05/2023 8:18 PM
Dear All, I am waiting for my return flight at the Love Field Airport. I need to thank Paul first for his whole-heart contribution to bring our whole new team to move on! I also feel so happy to know you all. 48 hours are not long but are long enough for us to build up a strong relationship! I only want to see I feel everyone in our team is great and the winner! Please keep in touch and let me know if I can be of any help!
❤️ 4

Davis Davalos-DeLosh 03/03/2023 6:08 PM
We are the BioinfoMagicians 🧙 We hack with precision and vision 🎯 We make your data disappear 🔮 And reappear with insights so clear 💡 We cast a spell on the competition 🌟 And win the hackathon with our ambition 💯

Arlen Larry Gyden · 1st
Data Science Student at University of Florida | Aspiring Biosta...
17h · 🌐

I just got back home from Susan G. Komen's 2023 Breast Cancer Hackathon, and I have to say, what an amazing experience!
#hackathon #BD4BC

In 48 hours, our team developed a tool to quantitatively assess the breast cancer microenvironment by cell type using accessible RNA sequence data with the future goal of advancing this product to predict a patient's response to tumor microenvironment-directed treatment. Our team impressed the judges and was awarded the 1st runner-up prize!

One key takeaway I am excited to apply in my #bioinformatics #research here at the University of Florida's Biomedical Engineering department with Dr. Xiao Fan is the importance of patient advocacy in scientific exploration. This #challenge taught me the importance of questioning my work and ensuring that my research goals have the patient in mind.

Lastly, I would like to thank Susan G. Komen, Lyda Hill Philanthropies, UT Southwestern Medical Center, and the BioHPC team for organizing this event. I am glad I got to work with my awesome teammates Chiara Corti, Hanieh Mazloom-Farsibaf, Jeremy Juybari, Kristof Kovacs, Maria Chikina, and Sam Rysdyk as they truly made this a memorable experience.

Contact Information

Grace Nicholl
Program
Administrator
gnicholl@komen.org

Jerome Jourquin
Sr. Director, Data
Science
jjourquin@komen.org

Aron Parekh
Scientific Programs
Manager
aparekh@komen.org

Jess Epps
Clinical Research Data
Manager
jepps@komen.org

Appendix A – Challenge Question Request for Proposals

BREAST CANCER HACKATHON CHALLENGE 2022

REQUEST FOR CHALLENGE QUESTIONS

Susan G. Komen®, the University of Texas Southwestern Medical Center (UTSW)'s Lyda Hill Department of Bioinformatics and the UTSW Harold C. Simmons Comprehensive Cancer Center are excited to announce the inaugural Breast Cancer Hackathon Challenge, a project funded by Lyda Hill Philanthropies.

Komen's Big Data for Breast Cancer (BD4BC) Initiative leverages big data to fuel scientific discoveries, advance research, improve breast cancer outcomes, and accelerate the delivery of equitable, patient-focused care. We believe that big data can be used to identify breast cancer disparities and variances in care and to improve methods for aggregating and analyzing clinical, genomic, and other sources of data to drive scientific progress and save lives.

The Hackathon aims to bring together creative individuals and teams with diverse backgrounds to develop new data science approaches/tools to address immediate needs or outstanding challenges in the fight against breast cancer.

YOU CAN HELP US FIGHT BREAST CANCER *and* WIN PRIZES!

We invite you to submit challenge questions—and suggestions for what data can be used to answer them—focused on significant issues in breast cancer that could be addressed with innovative, big data tools and solutions. Your ideas will be used to determine a set of challenge questions for the Hackathon. If you submit one of the selected questions, you will receive a prize (up to $250!) and the opportunity to lead a multidisciplinary team to tackle your question during the event.

Go ahead and submit challenge questions by following the link or QR code.
Due by 11:59 pm ET on June 24, 2022.
Click here to submit a challenge question!

What makes a good breast cancer hackathon challenge question?

The goal of Komen's research programs is to advance personalized medicine and improve health outcomes for everyone diagnosed with breast cancer. We are interested in hackathon challenge questions related to conquering metastatic and aggressive breast cancers and/or eliminating disparities in breast cancer care and outcomes. Questions can cover any area(s) of breast cancer research or any stage of the continuum of care.

We are especially interested in Hackathon challenge questions that incorporate diverse disciplines and fields of study (including those outside of the traditional life sciences), new technologies, methods, and/or leveraging existing resources in innovative ways. We are also interested in questions that are applicable to populations that experience disparities, either by focusing on specific populations or by demonstrating broad and inclusive applications across a diverse population.

Examples of areas of interest for proposed hackathon challenge questions:

- Genomic, biological, environmental, economic, lifestyle, and social factors that impact the burden of breast cancer for patients.
- Services and interventions to monitor care and improve outcomes across population(s). *(Special interest in population groups affected by breast cancer disparities.)*
- Detection of breast cancers, including those that cannot be detected due to current limitations in technologies, methods, treatments, etc. *(Special interest in inflammatory breast cancer and new or recurrent metastatic breast cancers.)*
- New technologies and methods, such as liquid biopsy, -omics, artificial intelligence, etc., in clinical practice that will provide better outcomes for patients.
- Patient outcomes during and after treatment (such as response, quality of life, recurrence, etc.).
- Precision medicine to identify the most effective and appropriate strategies to treat, detect, diagnose, and prevent breast cancer.

What type of data is needed to work on hackathon challenge question?

When submitting your Hackathon challenge question application, identify data sources that can be made accessible to all participants throughout the event. Data must be provided in a de-identified format to comply with regulatory standards to protect confidentiality and privacy of individuals if applicable. **No data containing personally identifiable information will be accepted for the Hackathon.**

While information about what data to use to answer the challenge question is not required for submission, we **strongly encourage** you to identify data sources that can be used to answer your challenge question. You can also provide suggestions for publicly available and/or published data sources that can be used to answer the challenge question.

Please note that all data, algorithms, tools, etc., created through the Breast Cancer Hackathon Challenge will be made available to the research community through a public data repository.

Susan G. Komen.

UT Southwestern
Lyda Hill Department of Bioinformatics

UT Southwestern
Harold C. Simmons
Comprehensive Cancer Center

LYDA HILL
PHILANTHROPIES

Who will participate in the hackathon to solve challenge questions?

Many types of people can participate in the in-person hackathon. We want you—oncologists, breast cancer researchers, research/patient advocates, and data scientists! Once specific Hackathon challenge questions have been selected for the event, people from diverse backgrounds from around the nation interested in participating in the in-person Hackathon will be invited to apply. Multidisciplinary teams will be organized and tasked with developing creative tools and solutions to answer one of the challenge questions. Multiple teams may work on the same challenge question – this is a (friendly) competition!

How do I submit my idea for a challenge question?

Submissions will require the following information:

- Applicant information (name, title, affiliation, email).
- Challenge question.
- Projected impact of answering the challenge question.
- Description of data (type of data, format of data, data fields, and potential sources of datasets).

Submissions are due by 11:59 pm ET on June 24, 2022.

Click here to submit a challenge question!

Frequently Asked Questions

Can I submit multiple challenge questions?
Yes, you may submit as many challenge questions as you like! Each question must be submitted separately.

Do I need to submit information about what datasets could be used to answer the challenge question I'm submitting?
No! But while data is not required to submit a challenge question, we strongly encourage you to identify data sources that can be used to answer your challenge question.

Will only one challenge question be chosen for the hackathon?
No! We will select several challenge questions for the event.

Can I submit a challenge question but not participate in the hackathon?
Yes! If your challenge question is selected for the hackathon, you will still receive a prize. (We will miss you at the event!)

Will I be able to participate in the hackathon if I do not submit a challenge question?
Yes! You will need to apply to participate in the event – stay tuned for more details!

If my challenge question is selected for the hackathon, will I have to apply to participate in the event?
No! If your challenge question is selected, and you want to participate in the hackathon, you are guaranteed a spot at the event and the opportunity to lead one of the multidisciplinary teams. We may contact you for additional information.

What will happen during the hackathon?
Multidisciplinary teams of data scientists, breast cancer researchers, research advocates, and oncologists will compete during the in-person event to ideate, implement, and present their best solutions to the challenge questions. A panel of judges will evaluate the solutions, and we will award prizes to the winning team(s)!

What will happen after the hackathon?
We will offer an opportunity for the winning team(s) at the in-person event to apply for funding to allow further testing, development, and validation of the approach developed during the hackathon.

For questions, please contact ssprograms@komen.org

The Breast Cancer Hackathon Challenge 2022 is a partnership between **Susan G. Komen** (the world's largest breast cancer organization), **UTSW Lyda Hill Department of Bioinformatics** (expert faculties in data science and with the infrastructure to share data with the research community) and **UTSW Harold C. Simmons Comprehensive Cancer Center** (the only NCI-designated Comprehensive Cancer Center in North Texas) with the generous support of **Lyda Hill Philanthropies**.

Appendix B – Hackathon Request for Participation

BREAST CANCER HACKATHON CHALLENGE 2023

REQUEST FOR PARTICIPATION

Susan G. Komen®, the University of Texas Southwestern Medical Center (UTSW) Lyda Hill Department of Bioinformatics and the UTSW Harold C. Simmons Comprehensive Cancer Center are excited to invite individual students, oncologists, breast cancer researchers and data scientists to apply to attend the inaugural Breast Cancer Hackathon Challenge, a project funded by Lyda Hill Philanthropies.

Where: UTSW Campus, Dallas TX
When: March 3-5, 2023

YOU CAN JOIN OTHERS TO FIGHT BREAST CANCER *and* WIN PRIZES!

Join us over an extended weekend (Friday afternoon to Sunday afternoon) to participate in the inaugural Breast Cancer Hackathon Challenge. This is a unique opportunity for individuals to come together in teams to develop new data science approaches and tools to address key challenges in the fight against breast cancer.

What can you expect from this event?

- Meet experts in oncology, mingle with patient advocates and make new friends across fields of expertise!
- Join a multidisciplinary team to work on one of three projects, led by a breast cancer expert!
- Compete to win prizes and the opportunity to validate the winning solution after the hackathon!
- We've got you covered! Travel, lodging, food, beverages and snacks will be provided for the event!

Go ahead and submit your application to attend by following the link or QR code.

Submissions are due by 11:59 pm ET on November 14, 2022.

Click here to submit your application!

BREAST CANCER HACKATHON CHALLENGE 2023

Who will participate in the Hackathon to work on projects?

People of diverse backgrounds are invited to participate in the in-person Hackathon. We want you—students, oncologists, breast cancer researchers, and data scientists—at any academic level! Once participants are selected, multidisciplinary teams will be assembled by the Hackathon's organizing committee and tasked with developing creative tools and solutions to one of the projects below.

Teams will be comprised of 5-8 individuals, with an expert in breast cancer serving as a team lead. Additionally, there will be other breast cancer experts, patient advocates, and technical leads on site to assist all teams. Multiple teams may work on the same project – remember, this is a (friendly) competition!

What are the projects I can work on?

Multidisciplinary teams will be assigned to one of the following projects (click each title to read the project brief):

#project-1://Predict/visualize combined functional effects of germline and somatic alterations in breast cancer.

> *'One size fits all' shouldn't be how we approach breast cancer treatment. Current pathway models are complex but do not account for one's individuality. Help us bring precision medicine to all people.*

#project-2://Define and compare metabolic states of cell types found in the breast tumor microenvironment.

> *It has been known for a long time that 'cancer' is not just a matter of cancer cells. Interaction with other cells often determines the onset and progression of cancer. Distinguishing the contribution of different cell types is limited when leveraging whole tumor data.*

#project-3://Identify/prioritize personalized drug combinations based on the genomic landscape of breast cancer.

> *Oncologists regularly get to expand their toolbox with new drugs, and one only hopes to determine the optimal treatments available to patients. Identification of combination therapies is limited by the number of possible drug combinations and genomic alterations.*

How do I submit my application to attend?

Applications will require the following information:

- Applicant information (name, affiliation, email, etc.).
- Area(s) of expertise/research (limit 200 words).
- Statement of interest (limit 200 words).
- Project preference.

Click here to submit your application!

Submissions are due by 11:59 pm ET on November 14, 2022

Susan G. Komen®

UTSouthwestern
Lyda Hill Department of Bioinformatics

UTSouthwestern
Harold C. Simmons
Comprehensive Cancer Center

LYDA HILL
PHILANTHROPIES

Frequently Asked Questions

Can I choose which project to work on?

You will be able to provide your preference for which project you would prefer to work on. We will do our best to accommodate this preference when teams are assigned to ensure balance in expertise.

Where can I get more information about the projects?

Background information and useful data sources can be found here.

How will applicants be selected?

Applicants will be selected based on several factors, including expertise or field of study, strength of the submitted statement of interest, and previous hackathon experience.

How will teams be formed?

Participants will be assigned to teams to ensure a multidisciplinary approach to each project. Team members will be encouraged to share expertise – so don't worry if you're not an expert in breast cancer or don't know how to code!

How will I communicate with my team before the Hackathon event?

You will be connected with your teammates prior to the event via a Discord server. More details will be provided if your application is accepted.

Will data be provided to work on the projects?

Yes! You will have access to the data needed for these projects.

Will technical or breast cancer experts be available to assist teams?

Yes! Breast cancer experts, patient advocates, and technical leads will be on site to assist any team in need.

Will computational resources be provided?

Bring your own laptop! We will provide additional data storage and computational resources as needed.

Will travel and accommodations be provided?

Yes! We will cover your travel, hotel, and meals, according to Komen's travel policy. If you are selected, we will provide instructions to use our travel agency to book your flights.

For questions, please contact SSPrograms@Komen.org

The Breast Cancer Hackathon Challenge 2023 is a partnership between Susan G. Komen (the world's largest breast cancer organization), UTSW Lyda Hill Department of Bioinformatics (expert faculties in data science and with the infrastructure to share data with the research community) and UTSW Harold C. Simmons Comprehensive Cancer Center (the only NCI-designated Comprehensive Cancer Center in North Texas) with the generous support of Lyda Hill Philanthropies.