EVALUATION REPORT FOR THE NATIONAL CANCER INSTITUTE CANCER SYSTEMS BIOLOGY CONSORTIUM

Hannah Dueck, Ph.D.
Shannon Hughes, Ph.D.
Dan Gallahan, Ph.D.
NCI Division of Cancer Biology
June 15, 2020
(Executive summary and Appendix 5 – Additional Data added August 2020)

Table of Contents

Executive Summary

In September 2016, the National Cancer Institute (NCI) launched the Cancer Systems Biology Consortium (CSBC) to support investigator-initiated research on basic cancer biology using systems biology approaches (https://csbconsortium.org/). The CSBC funding opportunity announcements define cancer systems biology as the integration of experimental biology and computational or mathematical modeling to build and test predictive models of cancer initiation, progression, treatment and metastasis. CSBC research is performed at U54 centers and through U01 projects (**Figure 1**). A U24 center functions as the Coordinating Center, promoting resource sharing, scientific collaboration, training, and outreach. Since its inception, the CSBC program has awarded 31 grants and, in FY2019, the budget was $37.5M (**Figure 2**). The CSBC is currently in its fourth year and, in fall 2020, NCI will determine the impact of the CSBC on cancer research and future programmatic direction. In support of this process, the following report provides information on progress towards the three main goals of the CSBC.

Goal 1: Advance understanding of mechanisms that underlie fundamental processes in cancer

The primary goal of the CSBC is to advance the mechanistic understanding of cancer using systems biology approaches. The impetus behind the CSBC is that systems biology approaches are uniquely able to address challenges in cancer research that other approaches cannot. The intent is that consortium research will be disseminated to the wider scientific community, and ultimately encourage the broad adoption of systems biology approaches to address challenges and discover new insights in cancer research.

- ***Scientific advances:*** Five major research themes (U54s) and questions (U01s) have emerged across the CSBC portfolio: (a) The role of tumor heterogeneity and evolution in cancer progression; (b) Biological mechanisms of therapeutic sensitivity and resistance; (c) Tumor-immune interactions in cancer progression and treatment; (d) Cell-cell interactions and complexities of the tumor microenvironment; and (e) Systems analysis of metastatic disease. A major thread that ties CSBC studies together is the consideration of cancer as a systems-level problem, taking into account the context in which molecules, cells, and tumors reside.
- ***Research dissemination and impact:*** The 343 CSBC articles published in 2017-2019 span a broad range of scientific disciplines, including cell biology and mathematical biology, and multiple cancer types, including breast, brain, lung, and pancreas (**Figure 3**). Bibliometric analysis indicates that almost all citations of 2017 CSBC publications are from researchers outside of the CSBC (with a self-citation rate of 1%) and that at least one third of these CSBC publications have a citation rate within the top 10% of publications from the same field and publication year (**Figure 4**).

Goal 2: Support the broad application of systems biology approaches in cancer research

A motivation for launching the CSBC initiative was to support the growth of a sustainable portfolio of cancer systems biology research. This long-term outcome will be accomplished through the combined efforts of multiple complementary past and future NIH and NCI programs. Indicators of sustainable funding for cancer systems biology include its strong presence in the federally funded research portfolio, the movement of research applications to an unsolicited grant pool, and the existence of an appropriate locus of review.

- ***Representation of cancer systems biology within the NIH grant application portfolio:*** The number of applications to NCI that include systems biology and mathematical modeling (the NCI SB portfolio) has consistently grown from 1999-2019 (**Figure 5**). Despite this, systems biology applications currently represent only a small proportion of the research applications submitted to NCI (2.2% in 2016-2020). Systems biology applications received by NCI focus on basic cancer biology, cancer treatment, and early detection, diagnosis, and prognosis.
- ***Contribution of the CSBC to the NIH grant application portfolio:*** CSBC publications have been cited by 2150 research applications submitted to NIH in 2017-2020, with 58% of these applications assigned to

NCI (**Figure 6**). In 2020, 25% of the NCI SB portfolio cited CSBC publications. These data suggest that a strong contribution of CSBC research to the NCI SB application pool and future federally funded research.

- ***Sustainability of support for cancer systems biology research:*** From 2004 to 2019, the fraction of unsolicited cancer systems biology applications has not increased; however, the success rate of these application is approaching that for all NCI applications, indicating partial progress towards sustainable funding for the field (**Figure 7**). CSBC funding continues to provide a major source of support for the field, with over 250 applications received since 2016. Challenges remain with identification of appropriate review loci for unsolicited applications (**Figure 8**). In 2014-2019, the minority (31%) of NCI SB applications were reviewed by a NIH standing study section. In study sections that received the most cancer systems biology applications over this window (>30 applications over five years), the success rate of all NCI applications was substantially higher than that of NCI cancer systems biology applications.

Goal 3: Support the growth of a strong and stable research community in cancer systems biology

An important goal of the CSBC is to support an emerging field until it becomes stable. While this is best achieved by performing impactful research, the CSBC also aims to build a strong workforce through support of research collaborations and the next generation of scientists. Additionally, the program aims to enable uptake of systems biology approaches by the cancer research community through the generation of shared data and resources.

- ***Strengthening the community through collaboration:*** A survey of CSBC participants indicated that 86% initiated new collaborations since joining the CSBC (**Figure 10**). Respondents indicated that conversations at annual investigator meetings, working groups, and the availability of supplemental funds supported the development of collaborations. NCI Administrative Supplement awards enabled 31 activities that supported collaborations within and outside of the CSBC, leading to publications and additional funding (**Figure 9**).

- ***Enabling update of cancer systems biology approaches through resource sharing:*** The CSBC has generated and shared more than 199 genomic data sets and 48 computational tools (**Figure 11**). Journal articles associated with the release of these resources have been cited by 1271 publications and 1492 NIH grant applications in 2017-2020. Over 50% of these articles (published in 2017 and 2018) have a citation rate within the top 10% of publications from the same field and publication year (**Additional Figure 3**). In late spring 2020, CSBC launched the Cancer Complexity Knowledge Portal to further support the broad accessibility of CSBC resources.

- ***Building an inclusive field through community expansion:*** The CSBC requires each U54 Research Center to include an Outreach Core and a pilot project funding program to support new systems biology projects and early stage investigators. In 2017-2019, 56 pilot projects were awarded, led by investigators not originally listed as CSBC key personnel (**Figure 12**). Sixteen projects led to additional funding, including NCI R01, U01 and R21 awards. Outreach Cores have facilitated a wide range of activities including symposia, training workshops, and community events.

- ***Recruiting of new cancer systems biologists through targeting training:*** From 2016-2020, CSBC funding has supported the participation of 305 trainees at all levels in CSBC research activities (**Figure 13**). The majority of former CSBC trainees (60%) continue to pursue systems biology research or are employed in a highly related field. Of the 25 moving to a tenure track position, 21 intend to pursue systems biology approaches. Additional activities to support trainees include the NCI-sponsored Junior Investigator Meeting and CSBC/PS-ON Summer Undergraduate Research Program, as well as a broad array of events targeting audiences from K-12 to post-graduate that are facilitated by CSBC U54 Outreach Cores.

Introduction

Evaluation Purpose and Charge to the External Panel

The onset and progression of cancer involves multiple molecular and cellular scale processes that adapt and evolve as a result of both internal and external events. These dynamic and interactive properties make cancer difficult to predict, prevent and treat. There has been significant progress in characterizing the genetics of cancer, as well as the downstream effects on the molecular and cellular pathways that are critical for the initiation and progression of cancer. However, many single parameter studies have also highlighted the need to understand cancers as integrated systems of genes, networks, and intercellular interactions.

In 2016, the National Cancer Institute (NCI) Division of Cancer Biology (DCB) launched the Cancer Systems Biology Consortium (CSBC) to support investigator-initiated research on basic cancer biology using systems biology approaches (https://csbconsortium.org/). The CSBC funding opportunity announcements define cancer systems biology as the integration of experimental biology and computational or mathematical modeling to generate predictive models that integrate multivariate perturbations with dynamic changes associated with cancer, such as initiation, progression, treatment and metastasis. The impetus behind the CSBC is that systems biology approaches are uniquely able to address challenges in cancer research that other approaches cannot. The CSBC program follows on other NIH and NCI efforts that first promoted the incorporation of systems biology approaches into biological and clinical research.

The data compiled in this report is intended to provide information on the NCI support of cancer systems biology research and progress towards CSBC program goals. The report is split into sections that address the three main goals of the CSBC:

- **Goal 1: Advance understanding of mechanisms that underlie fundamental processes in cancer**
- **Goal 2: Support the broad application of systems biology approaches in cancer research**
- **Goal 3: Support the growth of a strong and stable research community in cancer systems biology**

The CSBC is currently in its fourth year. In fall 2020, NCI will determine the impact of the CSBC on cancer research and future programmatic direction. In support of this process and in line with the goals of the program, NCI seeks to receive input from external experts regarding these general areas:

- What impact is CSBC having on the field of cancer biology?
- What is the role of CSBC in supporting a strong and stable research community?
- Should NCI continue to support a dedicated program in cancer systems biology?
- Are the current funding mechanisms (U24, U54, and U01) appropriate to achieve the program goals?
- Are there additional training and outreach activities that should be pursued to achieve program goals?

The role of the expert panel will be to provide high-level strategic recommendations for future program direction and implementation. This document aims to provide an overview of research highlights from the first four years of the program and quantitative data speaking to the specific questions above. In addition to reviewing the written information, you are invited to participate in the virtual 2020 CSBC Annual Investigator Meeting on September 16-17, 2020. The CSBC investigators are informed of your attendance and interaction with CSBC members is encouraged.

Additional data requests from panelists are welcomed and encouraged. Please note that this evaluation focuses on the CSBC program as a whole and how to best support the field of cancer systems biology going forward. The evaluation will not focus on the progress of individual grant awards, and access to grant-level data will be limited. Program-level summary statistics will be provided whenever possible. Your role and efforts are a critical aspect of NCI oversight and we highly appreciative of your time and insights.

Introduction to the CSBC

The CSBC aims to address challenges in cancer research through the combination of experimental biology and computational or mathematical modeling in a highly integrative and iterative manner. There are 3 components that make up the Consortium **(Figure 1)**, each described in the paragraphs below. In brief, CSBC research is performed at U54 centers and through U01 projects. A U24 center functions as the Coordinating Center, promoting resource sharing, scientific collaboration, training of a transdisciplinary workforce, and outreach to the community. Since its inception in September 2016, the CSBC program has awarded 31 grants **(Figure 2A)**. The FY2019 budget was $37.5M **(Figure 2B)**.

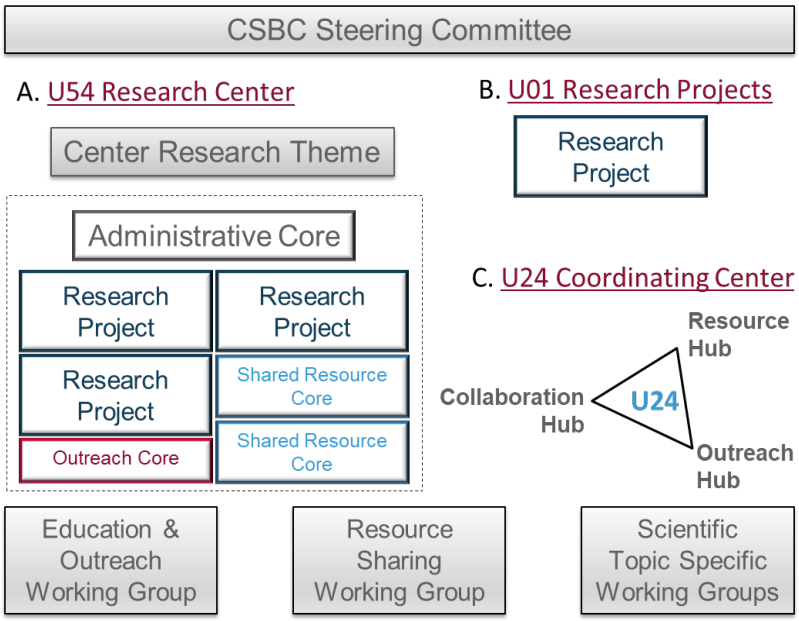


Figure 1: Structure of the NCI Cancer Systems Biology Consortium (CSBC). A. Each U54 Research Center consists of 2-3 research projects, up to 2 shared resource cores, an administrative core, and an outreach core. The U54 is organized around a central research theme that serves to integrate the efforts of the Center. B. Specific, circumscribed U01 Research Projects extend the cancer research topics pursued within the U54 Research Centers. C. The U24 Coordinating Center is organized around three 'hubs' that serve to catalyze collaboration and sharing of CSBC research products across the CSBC and the scientific community. Further information about the CSBC and funded projects can be found at www.csbconsortium.org.

The CSBC U54 Research Centers, solicited under *RFA-CA-15-014* and first awarded in 2016, form the major scientific, training, and outreach hubs of the CSBC initiative, providing expertise and resources to the Consortium while pursuing independent cancer systems biology research projects (**Figure 2A, Appendix 1**). The CSBC Research Centers consist of interdisciplinary teams, research and training and outreach programs, and infrastructure organized around a Center-specific investigator-identified research theme. The CSBC Research Centers develop and employ, individually and through collaborative Consortium activities, predictive computational, mathematical, or theoretical models that advance our current understanding of basic cancer biology and oncology.

First awarded in April 2017, the CSBC also includes U01 Research Projects solicited under *PAR-16-131* and, more recently, *PAR-19-287* (**Figure 2B, Appendix 1**). While the U54 Research Centers address high-level questions that require a broad range of expertise and multiple integrated projects, the U01 Research Projects provide opportunities for investigators to address specific, circumscribed questions through individual research grants with a scope and budget more similar to a traditional NIH R01.

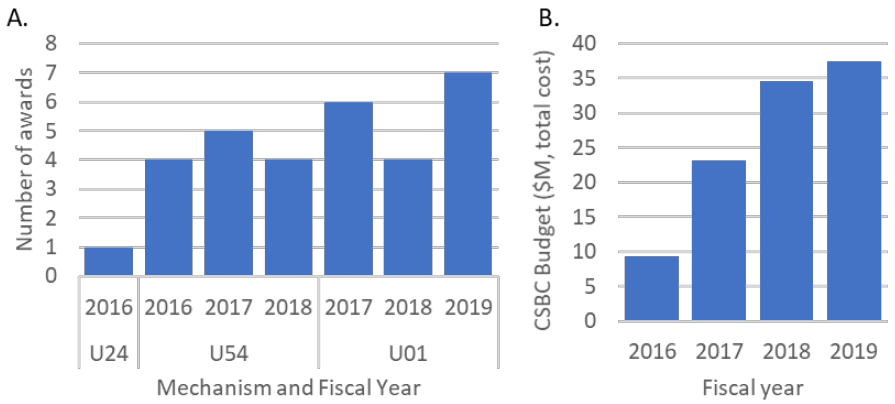


Figure 2: Overview of CSBC Awards and Budget. (A) As of May 1, 2020, the CSBC consists of 31 awards (13 U54 Research Centers, 17 U01 Research Projects, and 1 U24 Coordinating Center). (B) Total CSBC budget per fiscal year as extracted from NIH RePORTER. Grant details can be found in Appendix 1.

In addition to the CSBC, the NCI Division of Cancer Biology supports the Physical Sciences in Oncology (PS-ON) Program. While the programs are distinct in research focus, they are complementary with respect to some computational approaches and experimental technologies. To optimize the synergy between the two programs, a U24 Coordinating Center was solicited under *RFA-CA-15-015* and awarded in September 2016 (**Figure 1C, Appendix 1**). The purpose of the U24 CSBC/PS-ON Coordinating Center is to promote resource sharing, scientific collaboration, training of a transdisciplinary workforce, and outreach to the community.

Investigators from the U54 Research Centers, U01 Research Projects and U24 CSBC/PS-ON Coordinating Center are expected to collaborate across the Consortium, participate in Consortium activities and annual meetings, and share data and analytical resources. The CSBC convenes a monthly Steering Committee (SC) meeting that is open to all participating investigators and trainees. The SC meeting features updates from CSBC scientific working groups, information from NCI regarding grants policies and new opportunities, and 1-2 scientific talks from CSBC investigators. CSBC investigators are also required to attend the CSBC Annual Investigator Meeting, an in-person event that is designed to facilitate information sharing and stimulate collaboration. Details of CSBC working groups, annual meetings, and collaborative activities are provided in the subsequent sections.

Goal 1: Advance the understanding of mechanisms that underlie fundamental processes in cancer

Most cancers result from the dysregulation of multiple molecular pathways and cell-level interactions that govern normal tissue function. Cancer systems biology attempts to develop predictive computational models that integrate disparate types of experimental or clinical data, often in a setting representative of disease dynamics, to derive new, testable hypotheses. High-throughput technologies, such as genomic and epigenomic sequencing, transcriptomics, metabolomics, proteomics, and high-content imaging, have generated enormous amounts of descriptive data and, thus, systems analyses, and predictive modeling are necessary to integrate across these disparate datasets to derive an actionable understanding of cancer. It is in this spirit that the NCI launched the CSBC.

The primary research goal of the CSBC initiative is to address challenges and discover new insights in cancer research that are difficult to address using other approaches. To work towards this goal, the CSBC supports investigator-initiated research on basic questions in cancer biology through the integration of experimental biology and computational modeling. The intent is that the research, resources, and tools generated by these projects will be disseminated to the wider scientific community, be applied by others beyond CSBC, and encourage broad adoption of systems biology approaches in cancer research. Use of these approaches, by CSBC and the broader community, ultimately will establish mechanistic insights and testable hypotheses improving our understanding of cancer initiation, progression and treatment. These scientific advances will be accompanied by

enhanced attention to the research area including broadened sources of support for the field and strong representation of cancer systems biology in the federally funded research portfolio.
While this section of the report includes quantitative data that are intended to inform the panel on CSBC research, we acknowledge the limitations of these metrics for assessing scientific advances. NCI relies on panelist expertise to aid in assessing the impact of CSBC research. Panelist assessment of consortium research, supported by participation at the CSBC annual investigators meeting, and of the state of the field are critical to NCI's ability to support cancer systems biology in the future.

CSBC scientific advances

The major goal of the CSBC initiative is to advance the mechanistic understanding of cancer using systems biology approaches that build and test predictive models of disease initiation, progression, and response to treatment. Specific research themes (U54s) and questions (U01s) are investigator-initiated and the resulting grant portfolio can be broadly grouped into five categories: (a) The role of tumor heterogeneity and evolution in cancer progression; (b) Biological mechanisms of therapeutic sensitivity and resistance; (c) Tumor-immune interactions in cancer progression and treatment; (d) Cell-cell interactions and complexities of the tumor microenvironment; and (e) Systems analysis of metastatic disease. The accompanying "CSBC Research and Highlights (2016-2020)" provides a summary of CSBC contributions to the five broad areas above. Only a subset of the over 430 consortium publications (as of May 15, 2020) are highlighted in the review and browsing of publications and tools associated with CSBC research can also be completed through the Cancer Complexity Knowledge Portal. If desired, specific details for each CSBC grant can be found through the links provided in *Appendix 1***.**

Dissemination of CSBC research

The analysis presented here is an attempt to provide insight into the scientific research funded by CSBC and pertains to research productivity **(Figure 3)** and the impact of publications on subsequent scientific research **(Figure 4)**. Information critical to data interpretation are included in the summary below. Additional information about these analyses can be found in *Appendix 2: Methods* at the end of this report.

This section presents summary-level bibliometric analysis of publications funded by CSBC in 2017 through 2019. Publications from 2016 were excluded because CSBC research projects were just initiated and publications likely represent work done before the funding period began. An endpoint of 2019 was used to ensure a fixed data set for all analysis. Note the above review of scientific highlights does include the most recent publications from 2020. For the time period 2017-2019, CSBC grantees published 343 articles that acknowledged CSBC funding **(Figure 3A)**. Of these, 82% were scientific articles and 9% were reviews, perspectives and comments **(Figure 3B)**. In addition to oncology, the Web of Science Subject Categories most highly represented in CSBC publications include cell and molecular biology, multidisciplinary science, and mathematical and computational biology **(Figure 3C)**. Publications confirm that multiple cancer types are being studied within the CSBC including breast, brain, lung, and pancreas (**Figure 3D**). Journals with the highest CSBC publication count during this time period include *Nature Communications, PNAS, Cell, Cell Systems, and Cancer Discovery* **(Figure 3E)**.

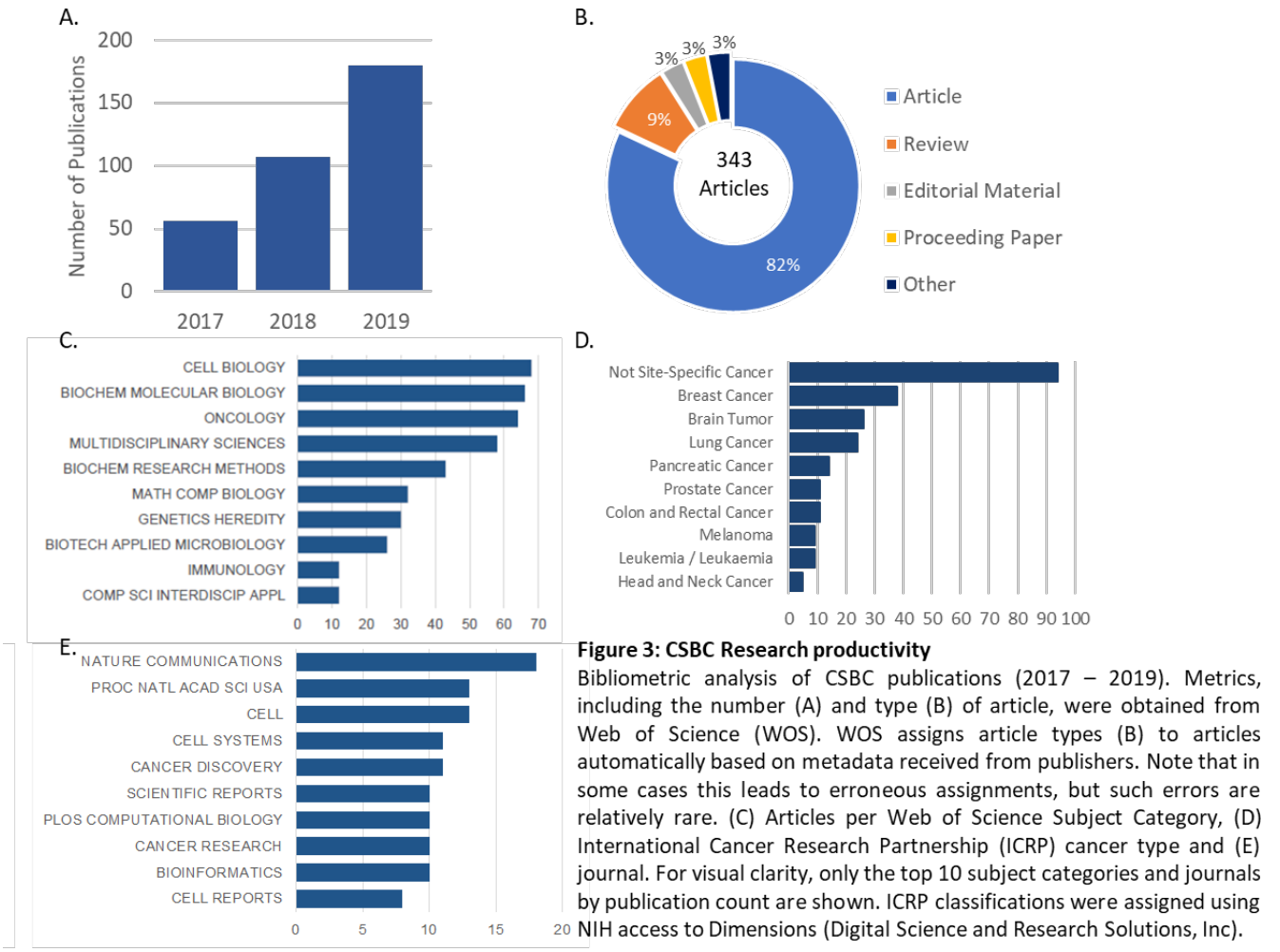


Figure 3: CSBC Research productivity
Bibliometric analysis of CSBC publications (2017 – 2019). Metrics, including the number (A) and type (B) of article, were obtained from Web of Science (WOS). WOS assigns article types (B) to articles automatically based on metadata received from publishers. Note that in some cases this leads to erroneous assignments, but such errors are relatively rare. (C) Articles per Web of Science Subject Category, (D) International Cancer Research Partnership (ICRP) cancer type and (E) journal. For visual clarity, only the top 10 subject categories and journals by publication count are shown. ICRP classifications were assigned using NIH access to Dimensions (Digital Science and Research Solutions, Inc).

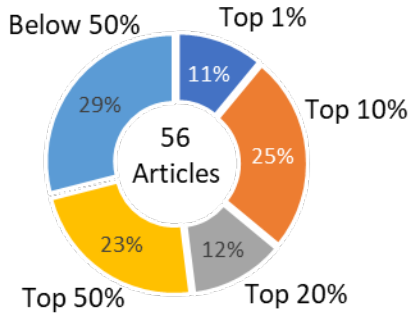
Analysis of scientific impact was completed via Web of Science and follows the principles established in the bibliometric literature to control for differences in citation volumes among scientific disciplines and publication age [1-3]. Per standard practice, CSBC bibliometric analysis is restricted to 2017 to ensure that publications are least 2 years old with enough citations for valid analyses, thus only a fraction of CSBC awards contribute to this analysis **(Figure 4A)**. The articles in 2017 (n=56 articles) have been cited 1435 times, averaging 26 citations per publication. The self-citation rate for CSBC publications from 2017 was less than 1%, indicating that almost all citations of CSBC publications were from the outside scientific community (most disciplines have self-citation rates between 2 and 10%)* [4]. Considering all articles published in the same Web of Science Subject Category and during the same year, 20 (36%) of 2017 CSBC publications have citation counts that rank in the top 10% **(Figure 4A,B)**. A table listing five CSBC publications with the highest citation percentile rank can be found in **Figure 4C**. These metrics indicate that CSBC publications are being cited by non-CSBC investigators and that at least one third of the CSBC publications are within the top 10% of cited publications in 2017.

**The self-citation rate is the percentage of citations received by these articles that were made by all authors of the papers in the analysis.*

A.

Bibliometric Indicator	Value
Number of articles	56
Number of citations (Times cited)	1435
Mean citation	26
Self-citation rate	< 1%
% of articles in the top 10% for citations	36%

B.



C.

PMID	Citation	Times Cited	Percentile in Subject Area
28514453	Philip M, et al. Chromatin states define tumour-specific T cell dysfunction and reprogramming. Nature 2017.	128	0.16%
28111070	Spitzer MH, et al. Systemic Immunity Is Required for Effective Cancer Immunotherapy. Cell 2017.	180	0.17%
28319113	Shen JP, et al. Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. Nature Methods 2017.	95	0.31%
28411207	Zhou M, et al. Transdifferentiation as a Mechanism of Treatment Resistance in a Mouse Model of Castration-Resistant Prostate Cancer. Cancer Discovery 2017.	76	0.44%
28380359	Tsujikawa T, et al. Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis. Cell Reports 2017.	100	0.59%

Figure 4: Impact analysis of CSBC publications in 2017.
Bibliometric analysis of publications funded by CSBC in 2017. Web of Science (WOS) was used to complete analysis. (A) The table lists common bibliometric indicators for CSBC publications in 2017. The self-citation rate is the percentage of citations received by these articles that were made by all authors of the papers in this analysis. The "% of articles in the top 10% for citations" indicator indicates that 36% of the 56 publications have citation counts that rank in the top 10% of all articles published in the same discipline and during the same year. (B) Distribution of articles among five percentile rank classes. Each article in this dataset is assigned to a percentile rank class based on its citation count in comparison with the citation counts of all articles published in the same year and Essential Science Indicators (ESI) subject category. (C) Top five CSBC publications by citation percentile rank.

Goal 2: Support the broad application of systems biology approaches in cancer research

A motivation for launching the CSBC initiative was to support the growth of a sustainable portfolio of cancer systems biology research. The success of the CSBC initiative should be accompanied by a strong representation of cancer systems biology in the federally funded research portfolio and broadened funding sources for the field. Because supporting a broad application of systems biology approaches in cancer research represents a long-term outcome of the CSBC initiative, the analysis here is provided to inform about progress towards this goal.

Representation of cancer systems biology within the NIH grant application portfolio

This section attempts to address the broad application of systems biology and computational modeling in the NCI grant application portfolio as well as the areas of cancer research most amenable to systems analysis **(Figure 5)**.

The analysis below presents summary-level data of research applications submitted to the NCI in 1999 through 2020 to provide context regarding growth of the cancer systems biology field. Here, "research" is defined based on the NIH Office of Extramural Research classification of grant mechanisms. For more, see *Appendix 2: Methods*. Over this time period, NCI supported multiple multi- and interdisciplinary programs that included large-scale data collection efforts and at least a partial focus on quantitative and computational biology. Some of those programs include the Integrative Cancer Biology Program (ICBP), the Quantitative Imaging Network (QIN), the Cancer Target Discovery and Development (CTD^2) program, the Clinical Proteomic Tumor Analysis Consortium (CPTAC), The Cancer Genome Atlas (TCGA), the Human Tumor Atlas Network (HTAN) and the Physical Sciences-Oncology Network (PS-ON).

To determine the growth of the cancer systems biology field from 1999-2020, an NCI systems biology grant application portfolio (NCI SB) was defined in a broad sense to encompass research grant applications submitted to NCI using the terms "system biology", "system analysis", and "mathematical modeling" (see *Appendix 2: Methods* for details regarding systematic search term choice). The number of submitted applications within NCI SB portfolio has consistently grown from 1999-2019 (5-year windows shown in **Figure5A**; note that 2020 is not included in this Figure because not all 2020 awards have been made to date), indicating an increased use of systems biology and computational modeling in cancer research. During the period of CSBC support (2016-2020), the NCI SB portfolio included 1617 research applications representing only 2.2% of the 72,839 research applications submitted to NCI in that time period. Note that the CSBC portfolio is a subset of the NCI SB portfolio. Therefore, while the number of applications within the NCI SB has increased with time it represents a small proportion of overall research applications submitted to NCI.

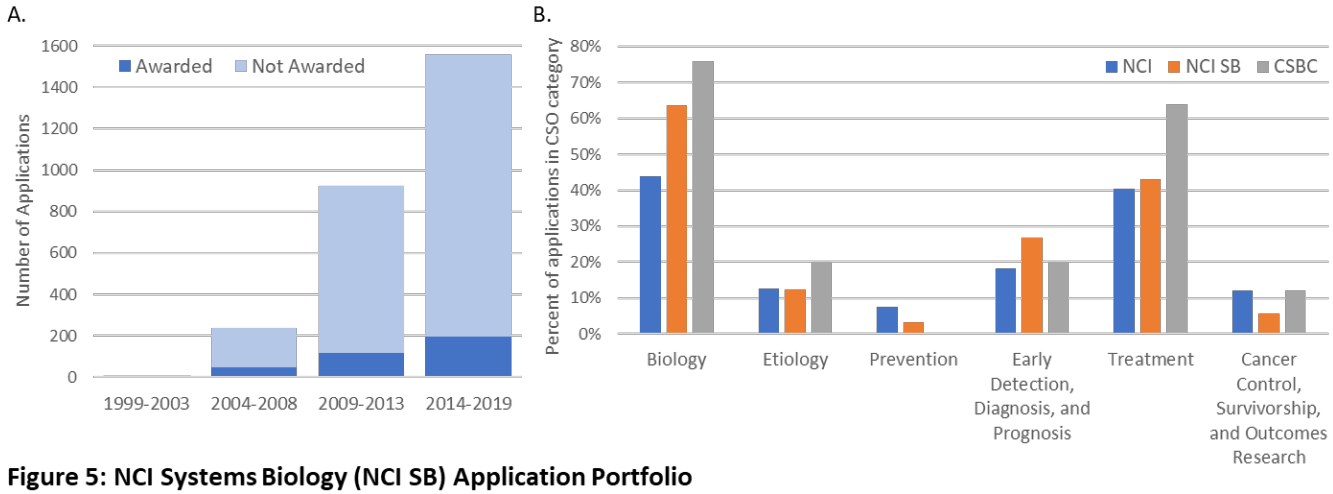


Figure 5: NCI Systems Biology (NCI SB) Application Portfolio
The NCI systems biology application portfolio (NCI SB) encompasses the research grant applications submitted to NCI using the terms "system biology", "system analysis", and "mathematical modeling" (see *Appendix 2: Methods* for details regarding search term choices and portfolio construction). (A) Number of NCI SB grant applications from 1999-2019. (B) Percentage of applications within each International Cancer Research Partnership (ICRP) Common Scientific Outline (CSO) category. The ICRP CSO provides a standard categorization of research projects into six areas of scientific interest in cancer research and was developed to compare research portfolios across agencies. NCI, NCI SB, and CSBC grants classified by ICRP CSO area. Note that not all applications receive a ICRP CSO categorization and multiple categories can be assigned.

Next, applications submitted to the NCI, NCI SB, and CSBC portfolios were classified using the Common Scientific Outline (CSO) developed by the International Cancer Research Partnership (ICRP) to explore whether specific areas of cancer research are considered by investigators to be particularly amenable to systems biology approaches. This coding system provides a standard categorization of research projects into six areas of scientific interest in cancer research and was developed to compare research portfolios across agencies **(Figure 5B)**. The percentage

of applications within the NCI SB portfolio coded within a specific category (note that an application can be coded with multiple categories) illustrates that systems biology and mathematical modeling approaches are overrepresented in the basic cancer biology and early detection, diagnosis, and prognosis categories versus the overall NCI application pool. The CSBC application portfolio is a particularly enriched in both basic biology and cancer treatment, as is expected and desired for grants managed by the NCI Division of Cancer Biology (**Figure 5B**). In sum, applications to NCI that include systems biology and mathematical modeling have increased over a 20-year period and tend to focus on basic cancer biology, cancer treatment, and early detection, diagnosis, and prognosis.

Contribution of the CSBC to the NIH grant application portfolio

To provide insight into the dissemination of CSBC research through the NIH, NCI and NCI SB application portfolios, the citation of CSBC publications within the three application pools **(Figure 6)** was examined. CSBC publications have been cited by 2150 research applications submitted to NIH in 2017 through 2020 (of which 248 have been awarded), with 58% of these applications assigned to NCI **(Figure 6A)**. The remaining applications were assigned to a broad range of institutes, suggesting relevance of CSBC findings and approaches to other disease states and biology. Within the NCI SB application portfolio, 155 (13%) research applications submitted between 2017-2020 cite CSBC publications **(Figure 6B)**. In 2020, 70 (25%) of SB portfolio research applications cite CSBC publications and only 9% of those applications were submitted by an investigator named as key personnel on any CSBC grant **(Figure 6C)**. These data suggest a strong contribution of CSBC research to the NCI SB application pool and additional contribution of CSBC findings to NIH and NCI applications more generally. Furthermore, CSBC applications are being cited within grant applications by investigators not associated with the CSBC.

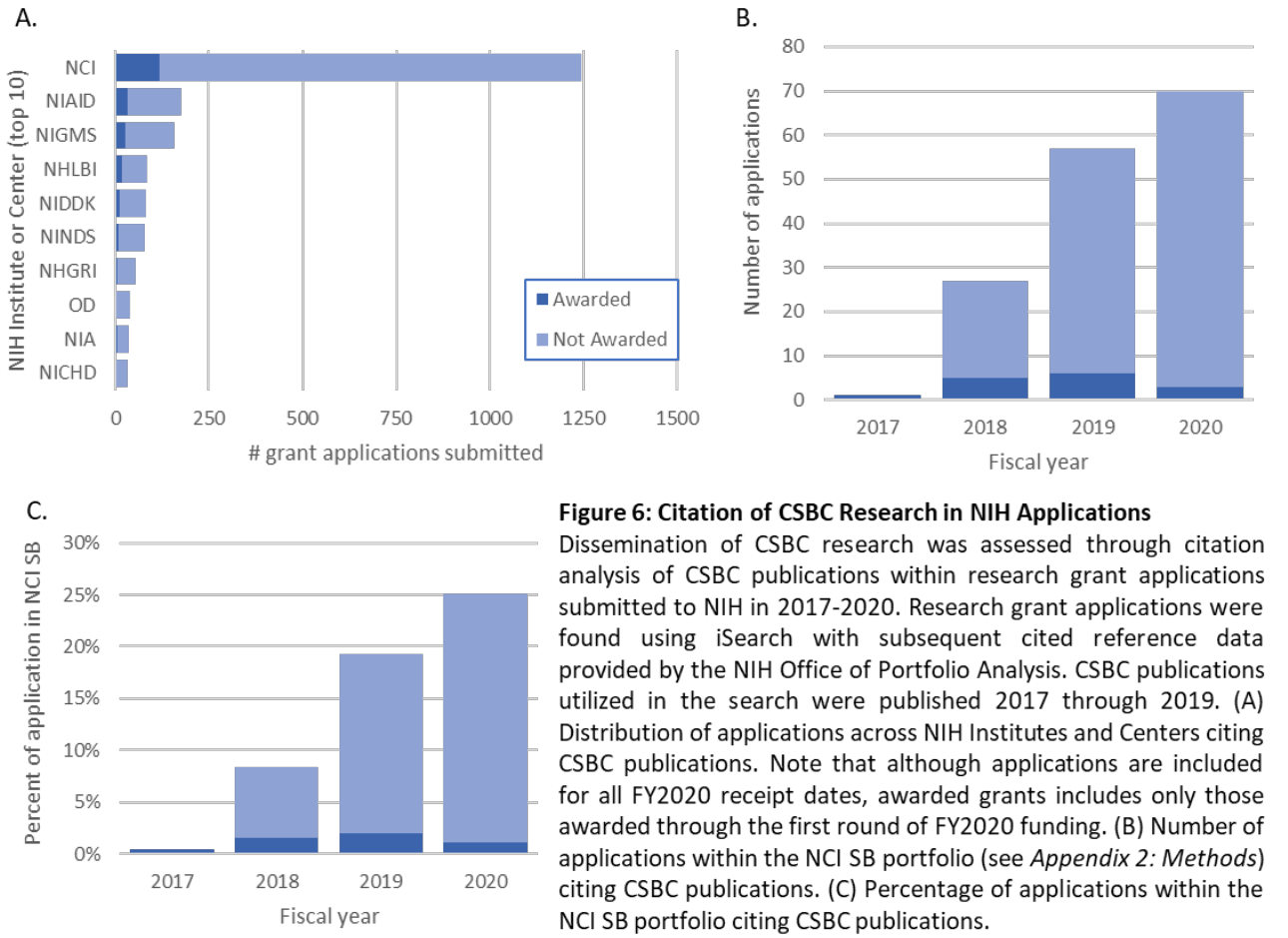
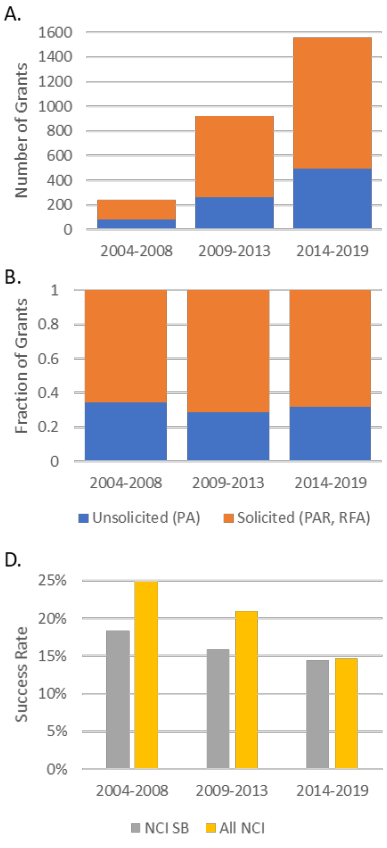


Figure 6: Citation of CSBC Research in NIH Applications
Dissemination of CSBC research was assessed through citation analysis of CSBC publications within research grant applications submitted to NIH in 2017-2020. Research grant applications were found using iSearch with subsequent cited reference data provided by the NIH Office of Portfolio Analysis. CSBC publications utilized in the search were published 2017 through 2019. (A) Distribution of applications across NIH Institutes and Centers citing CSBC publications. Note that although applications are included for all FY2020 receipt dates, awarded grants includes only those awarded through the first round of FY2020 funding. (B) Number of applications within the NCI SB portfolio (see *Appendix 2: Methods*) citing CSBC publications. (C) Percentage of applications within the NCI SB portfolio citing CSBC publications.

Sustainability of support for cancer systems biology research

In general, NIH and NCI special programs, such as the CSBC, are utilized to support an emerging field until sustained funding can be maintained through other means. An indication of sustainable funding for cancer systems biology approaches would be the movement of research applications to the unsolicited grant pool. Key indicators of progress towards this goal include where applications are submitted, application success rates, and application review outcomes. This section attempts to address the broadening of funding sources for cancer systems biology research (**Figure 7** and **Figure 8**).

The NCI SB application pool was analyzed to determine how applications were solicited over time: applications were submitted either to the unsolicited or NIH "parent" Funding Opportunity Announcements (FOAs) or to solicited or NCI-specific FOAs (**Figure 7A**). As demonstrated previously (**Figure 5A**), the NCI SB application pool increased between 2004-2019 (note that for this analysis the period between 1999-2003 is not included due to the low number of total applications, n = 12). The increase in applications submitted was to both unsolicited and solicited FOAs suggesting overall growth in the field. However, the proportion of systems biology-focused applications submitted to the NIH parent FOAs remained nearly constant between 2004-2019 (**Figure 7B**). Unsolicited systems-biology focused applications are primarily submitted to the NIH parent R01 and K99/R00 FOAs (**Figure 7C**) and 11% of the solicited applications between 2004-2019 were submitted in response to NCI-specific "parent" announcements for the P01 and R21 funding mechanisms which may also be considered true investigator-initiated research. An examination of FOAs that have received 20 or more system biology-focused applications from 2004-2019 shows that CSBC funding is considered a major source of support for the field, as evidenced by the high number of applications received: over 250 applications have been received in response to the CSBC FOAs since 2016. Finally, the success rate for applications within the NCI SB portfolio is approaching that for all NCI applications, suggesting that systems-biology approaches are becoming recognized as being critical within the cancer research field and is an encouraging indication that stable funding may be possible.



C.

FOA Number	FOA Title	Mechanism	Rank
PA13-302	Research Project Grant (NIH Parent FOA)	R01	1
PAR16-131	Emerging Questions in Cancer Systems Biology	U01	2
PA16-160	NIH Research Project Grant (NIH Parent FOA)	R01	3
PA19-056	NIH Research Project Grant (NIH Parent FOA)	R01	4
PA18-484	NIH Research Project Grant (NIH Parent FOA)	R01	5
PAR13-146	NCI Exploratory/Developmental Research Grant Program (NCI FOA)	R21	6
CA15-014	Research Centers for Cancer Systems Biology Consortium	U54	7
PAR14-169	Physical Sciences-Oncology Network (PS-ON): Physical Sciences-Oncology Centers (PS-OC)	U54	8
PAR19-287	Research Projects in Cancer Systems Biology	U01	9
PAR13-184	Collaborative Research in Integrative Cancer Biology	U01	10
PAR15-021	Physical Sciences-Oncology Network (PS-ON): Physical Sciences-Oncology Projects (PS-OP)	U01	11
PAR18-290	National Cancer Institute Program Project Applications (NCI FOA)	P01	11
PA11-260	NIH Pathway to Independence Award (NIH Parent FOA)	K99/R00	12
PAR19-101	Physical Sciences-Oncology Network (PS-ON): Physical Sciences-Oncology Projects (PS-OP)	U01	13

Figure 7: Solicitation of Systems Biology Research Grant Applications

Receipt of NCI SB applications was assessed to unsolicited (NIH parent or unspecified) funding opportunity announcements (FOAs) versus solicited (Program Announcement with Special Review [PAR] or Request for Applications [RFA]) FOAs. (A) Number of research grant applications submitted through unsolicited (Program Announcement [PA]) or solicited (PAR, RFA) FOAs between 2004-2019. Note that the period 1999-2003 is not included due to the low number of total applications (n = 12). (B) Fraction of grants submitted to solicited versus unsolicited FOAs. (C) Table of FOAs with greater than 20 systems biology-focused applications submitted (2004-2020). (D) Success rate of unsolicited research grant applications to NCI (2004-2019). Note that the success rate reported here only includes research grant applications as defined by NIH Office of Extramural Research classification of grant mechanisms.

Key to achieving stable funding of cancer systems biology research is an appropriate locus of review. CSBC applications submitted in response to RFA-CA15-014 (U54) and PAR-16-131/PAR-19-287 (U01) were and continue to be reviewed by Special Emphasis Panels convened new with each receipt date by the NCI Division of Extramural Activities. The overall success rate for the CSBC program, as of May 1, 2020, was approximately 23% for U54s (2016-2018) and 9% for U01s (2017-2019). To understand more generally where cancer systems biology research applications submitted to NCI are reviewed, we determined where all applications within the NCI SB portfolio were reviewed between 2014-2019. Of the 1560 applications in the NCI SB portfolio during that time frame, 1084 (69%) of the applications were reviewed by NIH or NCI Special Emphasis Panels (review panels generally convened with new members for each review) and the remaining 476 applications were reviewed in standing study sections convened by the NIH Center for Scientific Review (CSR). On average only 20-30 cancer systems biology focused applications were reviewed across all CSR standing study sections per round of receipt, making them a very small fraction of any review (**Figure 8A, 8B**).

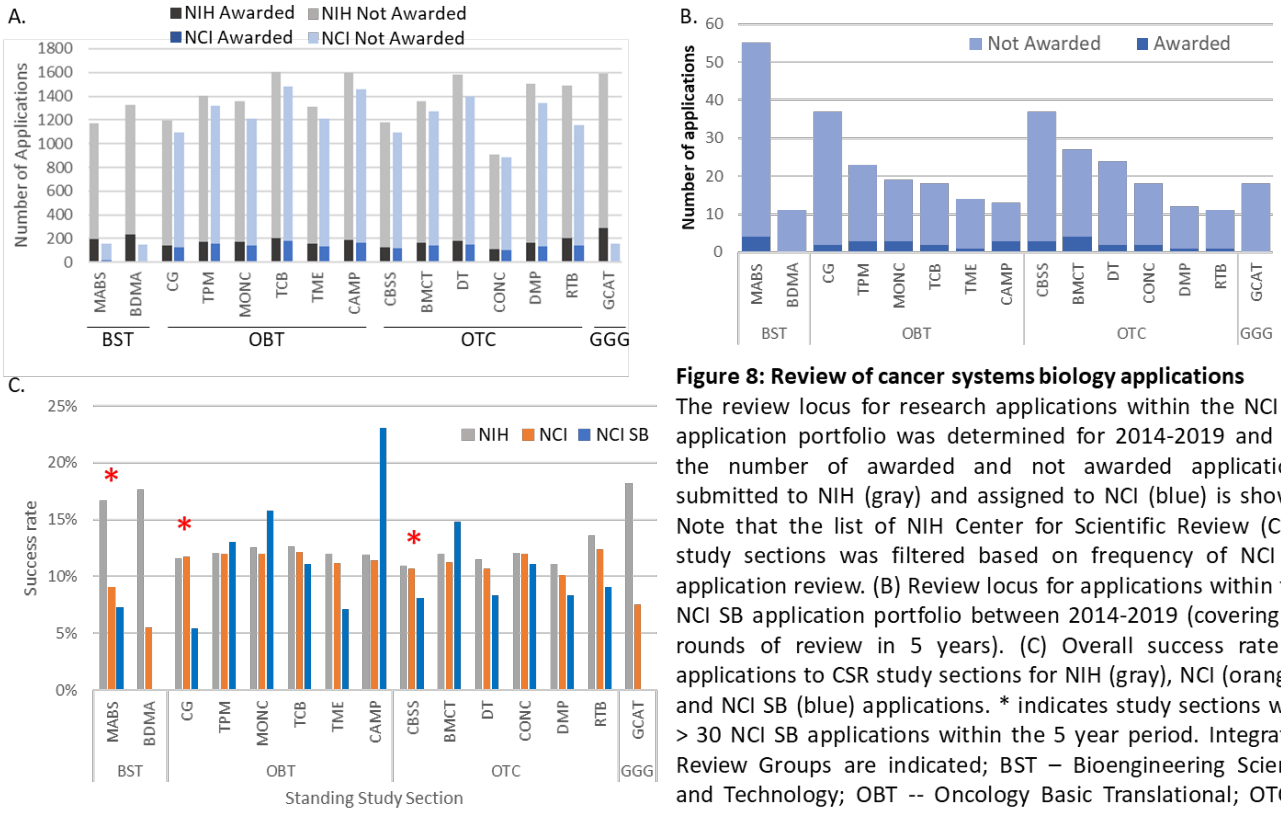


Figure 8: Review of cancer systems biology applications
The review locus for research applications within the NCI SB application portfolio was determined for 2014-2019 and (A) the number of awarded and not awarded applications submitted to NIH (gray) and assigned to NCI (blue) is shown. Note that the list of NIH Center for Scientific Review (CSR) study sections was filtered based on frequency of NCI SB application review. (B) Review locus for applications within the NCI SB application portfolio between 2014-2019 (covering 15 rounds of review in 5 years). (C) Overall success rate of applications to CSR study sections for NIH (gray), NCI (orange), and NCI SB (blue) applications. * indicates study sections with > 30 NCI SB applications within the 5 year period. Integrated Review Groups are indicated; BST – Bioengineering Science and Technology; OBT -- Oncology Basic Translational; OTC – Oncology Translational Clinical; GGG – Genes, Genomes, and Genetics. Links to study sections in *Appendix 2: Methods.*

Because of the round-to-round variability inherent in Special Emphasis Panel review and to better understand how applications submitted to unsolicited FOAs are distributed across CSR standing study sections, this analysis focuses on review of cancer systems biology applications within standing study sections. Considering only the fifteen CSR study sections that reviewed more than ten NCI SB applications between 2014-2019, NCI SB application outcomes were reviewed to determine how they fared in comparison to all applications submitted to NIH and to all those that were assigned to NCI (**Figure 8C**). In study sections that received the most cancer systems biology applications (>30 applications from 2014-2019, marked by asterisk in **Figure 8C**) the success rate of all NIH and NCI applications was substantially higher than that of cancer systems biology applications. This was independent of the focus of the study section [i.e. if the focus was on computational and mathematical approaches (MABS; Modeling and Analysis of Biological Systems) or if the focus was specifically on cancer (CG; Cancer Genetics or

CBSS; Cancer Biomarkers Study Section)]. Note that it is difficult to determine if these trends hold true across all 15 study sections due to the low number of applications over the 5-year period, but of the study sections that have reviewed greater than 20 applications (2014-2019), TPM (Tumor Progression and Metastasis) and BMCT (Basic Mechanisms of Cancer Therapeutics; now MCT1 and MCT2) appear to be more supportive of systems biology approaches.

Cancer-focused applications, in general, fared less well in the MABS study section and this was generally true across study sections with a focus on computational approaches (MABS, BDMA, GCAT). It is possible this is due to the breadth of research topics covered within those study sections, as cancer-focused applications are a minority of applications (**Figure 8A**). However, a confounder to this analysis is the different methods of paying awards employed across NIH Institutes and Centers (for example, disparate "pay line" cutoffs set by IC priorities). Difference in pay lines across the NIH likely explains why MABS, BDMA, and GCAT have overall higher success rates for all NIH applications than seen in the study sections falling within the OBT (Oncology 1-Basic Translational) and OTC (Oncology 2-Translational Clinical) Integrated Review Groups, in which the majority of applications are assigned to NCI. Nonetheless, NCI SB applications still fare worse in MABS (7% success rate), BDMA (no awards), and GCAT (no awards) than NCI applications in the OBT and OCT administered study sections. Note that the 9% success rate of CSBC U01 applications is similar to the overall success rate of NCI applications submitted to MABS although caution must be used when comparing these numbers because the number of CSBC U01s awarded is dependent on the availability of NCI funds with each round and there is no set pay line for CSBC U01 applications.

In sum, this section demonstrates that there is an increasing number of cancer systems biology applications being submitted to NIH, but that the fraction of unsolicited applications has not increased (**Figure 7**). This suggests progress towards stability of federal funding for the field but that there work yet to be done as there exist further opportunities to encourage applications to the NIH and NCI parent funding announcements. However, challenges remain with identification of appropriate review loci for unsolicited cancer systems biology applications (**Figure 8**).

Goal 3: Support the growth of a strong and stable research community in cancer systems biology

It is now widely appreciated that all biological scientists must receive some level of training in computational analysis and mathematical modeling. However, for systems biology studies to be fruitful, those analyses must be accompanied by a deep appreciation for the biology at hand, therefore equally as important is the training of computer scientists, engineers, and physicists in cancer biology. An important goal of the CSBC is to provide opportunities for bench and computational scientists to become fluent in the language of the other. It is through this goal that the CSBC aims to support the growth of a strong, stable workforce and research community in cancer systems biology. Activities towards this goal operate at the investigator, trainee, K-12, and community levels. This section will provide an overview of these activities related to (1) strengthening the cancer systems biology community through collaboration and resource sharing, (2) building an inclusive field through community expansion, and (3) recruitment of new cancer systems biologists through targeting training.

Strengthening the community through collaboration and resource sharing

Collaboration and resource sharing are key to building a thriving research field and these activities have been encouraged in the CSBC through a variety of mechanisms. First, to promote intra-center collaboration each U54 Research Center was required to identify a central research theme that serves to integrate the Center. Second, to promote inter-consortium collaborations the U24 Coordinating Center was awarded to manage trans-consortium working groups, promote CSBC research projects, and to promote collaboration with the closely related Physical Sciences-Oncology Network program. Third, every CSBC investigator has the opportunity to compete for one-year collaborative Administrative Supplement awards to jump-start new collaborations that arise from consortium activities. Finally, each U54 Research Center manages a required pilot project fund within the Administrative Core

to encourage introduction of new researchers to the Center. Each of these activities and their outcomes to date will be discussed below.

CSBC investigators participate in both administrative and scientific working groups that are coordinated by the U24 Coordinating Center at Sage Bionetworks. Participation in the two administrative working groups, the Education and Outreach Working Group (EOWG) and the Resource & Data Sharing Working Group, are mandatory for U54 Research Centers. EOWG activities will be described below. The Resource & Data Sharing Working Group is organized around 1-2 monthly scientific seminars with dual goals of introducing new tools and data and promoting collaboration. A total of seven scientific working groups have been convened based upon common interests. These working groups are led by CSBC or PS-ON investigators and have a range of goals, from very specific organization of the Tumor Deconvolution DREAM Challenge to targeted project development in the area of protein-protein interactions to broad discussion on the challenges in tumor metabolism research. Most scientific working groups meet on a monthly or quarterly basis.

In recognition that the unique strengths and new insights of the CSBC investigators could be combined across the consortium to increase the overall impact of their awards, NCI has utilized the one-year Administrative Supplement mechanism to support new collaborations between CSBC members, CSBC members and DCB-related programs (such as PS-ON), and CSBC members and investigators outside of the program. Administrative Supplement awards (2016-2019) have supported 41 distinct activities across a range of collaboration structures (**Figures 9A, 9B)**. Note that collaboration is not required to apply for a CSBC Administrative Supplement but is strongly encouraged and is prioritized in review.

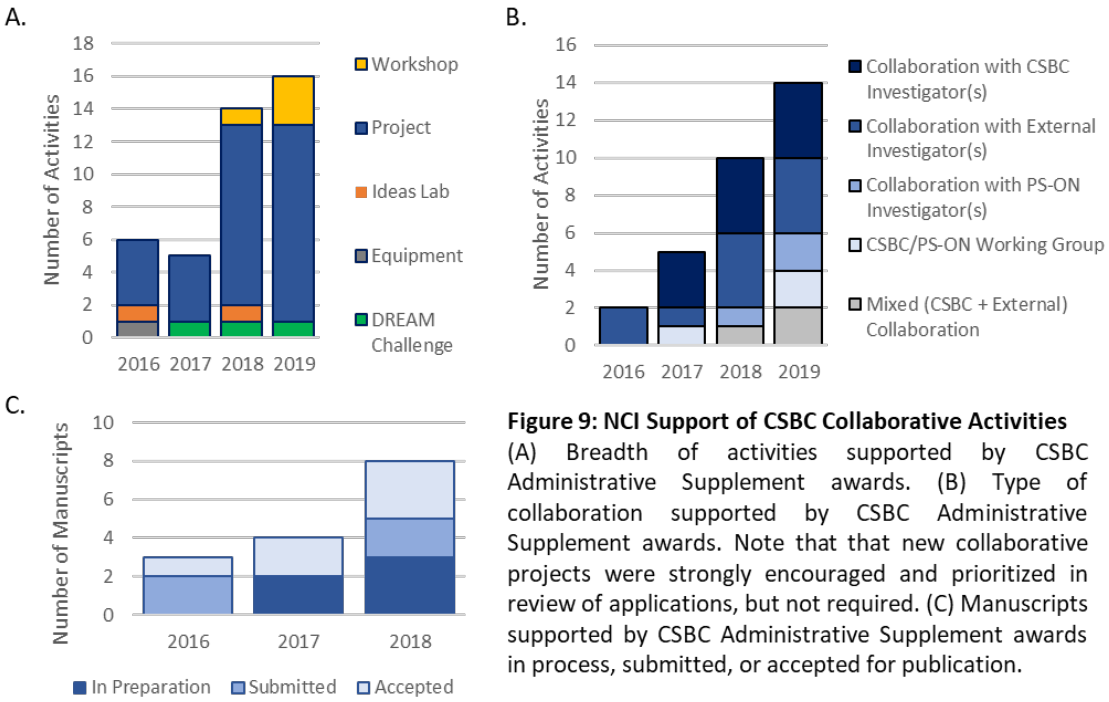


Figure 9: NCI Support of CSBC Collaborative Activities
(A) Breadth of activities supported by CSBC Administrative Supplement awards. (B) Type of collaboration supported by CSBC Administrative Supplement awards. Note that that new collaborative projects were strongly encouraged and prioritized in review of applications, but not required. (C) Manuscripts supported by CSBC Administrative Supplement awards in process, submitted, or accepted for publication.

Research projects represent 76% (n = 31) of Administrative Supplement-supported activities with an average total cost budget per award of $147,000. As of May 1, 2020, Administrative Supplements (2016-2019) have resulted in 6 peer-reviewed publications with several others submitted or in-preparation (**Figure 9C**). In addition to research projects, Administrative Supplement awards supported other collaborative activities such as DREAM Challenges to encourage method development for deconvolution of bulk RNA-seq data into individual cell types and automatic labeling of experimental and clinical metadata to facilitate computational analysis. Almost 500

computational scientists have registered and participated in the CSBC-sponsored DREAM Challenges. Ideas labs, intensive 3-5 day science "bootcamps" designed to rapidly build collaborations, were held by Sage Bionetworks (2016) and Memorial Sloan Kettering (2018) focused on interdisciplinary approaches to cancer metastasis and systems approaches to the cancer microbiome, respectively. Outcomes from the Ideas labs include publications [5, 6] and new grant funding (NCI R01 [PI Lee], Susan G. Koman [PI Bravo-Cordero]). Administrative Supplement funding supported collaboration-building workshops led by CSBC members such as the 2019 West Coast Cancer Systems Biology Symposium, an event organized by the OHSU U54 Research Center that ended with small pilot funding for trainee-led cross-CSBC projects. Finally, CSBC scientific working groups, such as the Protein-Protein Interaction working group and the Image Analysis working group have utilized Administrative Supplement funding to catalyze joint research projects and hackathon events. In sum, new collaborations and activities to promote nascent collaboration have resulted in multiple publications, additional funding, and increased investigator interactions.

A survey administered by NCI program staff to better understand how collaborative projects arise within the CSBC indicated that 86% of respondents initiated new collaborations since joining the consortium (n = 56, response rate of approximately 33% across all U54 and U01 MPIs and project managers). Survey results identified conversations initiated at face-to-face annual investigator meetings, the availability of supplemental funds, CSBC working groups, and CSBC Steering Committee meetings as important initiators of new collaborations (**Figure 10A**). Investigators were also asked which activities, if any, were less effective for promoting collaboration and resource sharing and although respondents found that no activities were truly ineffective, they found CSBC scientific working groups to be less effective than others (**Figure 10B**).

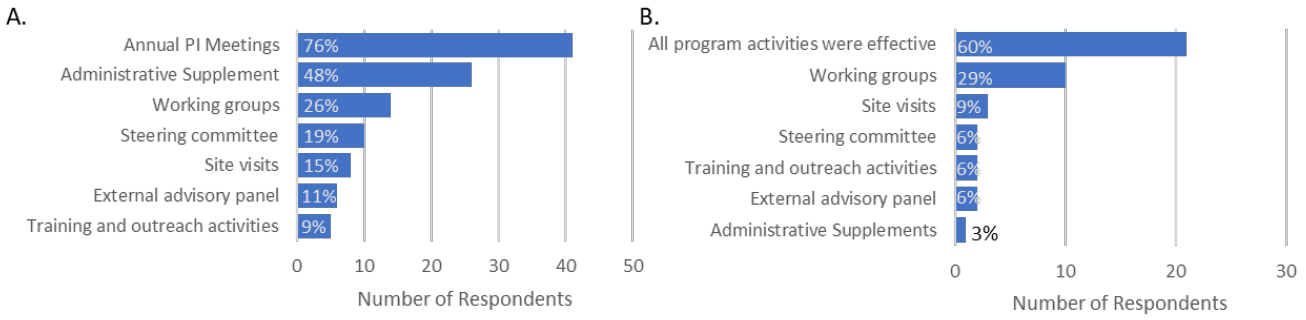


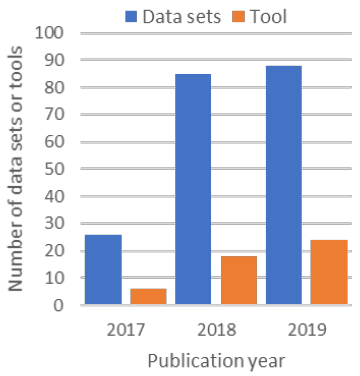
Figure 10: NCI Investigator Feedback on CSBC Activities
All CSBC principle investigator-level members were invited to participate in a survey focused on NCI-hosted CSBC program activities (n = 56 respondents). (A) Survey answers to "Choose one or two activities that were most effective in fostering collaboration". Percentage of question respondents shown in bar. (B) Survey answers to "Select the activities, if any, that were INEFFECTIVE in facilitating the programmatic goals of (1) establishing collaborations between CSBC investigators, and (2) sharing expertise and resources across the consortium." Note that Investigators were allowed to choose multiple answers and were not required to answer all questions. Although investigators from all award mechanisms (U24, U01, U54) were invited to answer, only U54 investigators participate in site visits and their own U54-specific External Advisory Panel.

Sharing of resources, such as new computational tools and data sets, can also lead to new collaboration and increased application of cancer systems biology through increased tool and resource availability. The majority of CSBC awards are required to share genomic data under the NIH Genomic Data Sharing Policy and 199 genomic data sets were reported in publications and shared with the community via NIH resources between 2017-2019 (**Figure 11A**). Note that this underestimates the true extent of data sharing because there are limited NIH-supported repositories to share non-genomic data with the community. Some Centers have set up individual web infrastructure to share data. Examples include the Harvard U54 (CyCIF), DFCI U01 (Interactome Database), Columbia U54 (PrePPI Database), UCSF/UCSD U54 (NDEx), and Stanford (LTMI). CSBC awardees have also shared 48 new computational tools via Github or other similar services (**Figure 11A**).

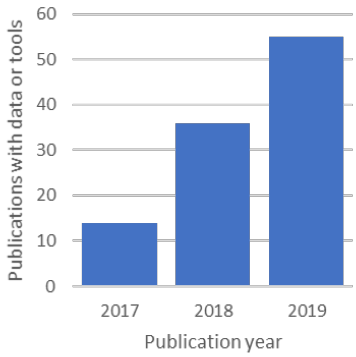
In total, 105 publications have reported new data sets or tools from 2017-2019 (**Figure 11B**). Bibliometric analysis of the 105 publications provides insight into which data sets and tools are valued by the scientific community. An analysis completed in May 2020 indicates that the top 10 publications have been cited 741 times in publications and 811 times in grant applications to the NIH (**Figure 11C**). Note that because all publications were included in the analysis that the top 10 may be more heavily weighted to publications prior to 2019. Overall this analysis suggests that data sets and tools generated by the CSBC are well-used in the scientific community.

At present CSBC data sets and tools are most easily discoverable through publication. However, the ultimate goal of the Cancer Complexity Knowledge Portal is to provide a convenient inventory of CSBC and PS-ON data sets, tools, publications, and other resources that is searchable by program, cancer research theme, and keyword. The portal launched in the late Spring of 2020 and the U24 Coordinating Center aims to improve the functionality of the portal through feedback from the CSBC community and beyond.

A.



B.



C.

PMID	Author	Journal	Year	Link to Tool and/or Data Set	Number of Citations (pubs)	Number of Citations (Grant Apps)
29961579	Azizi E et al.	Cell	2018	https://github.com/dpeerlab/seqc Data: PRJNA472387	129	104
28514453	Philip M et al.	Nature	2017	GSE89309	128	148
28380359	Tsujikawa T et al.	Cell Reports	2017	https://github.com/multiplexIHC/cppipe	100	144
29961576	van Dijk D et al.	Cell	2018	https://github.com/DpeerLab/magic/ Data: GSE114397	86	135
28411207	Zou M et al.	Cancer Discovery	2017	GSE92721	76	61
28246332	Carmona-Fontaine R. et al.	PNAS	2017	GSE93702	74	28
30078711	Goltsev Y et al.	Cell	2018	https://github.com/nolanlab/CODEX	51	89
29505029	Ma J et al.	Nat Methods	2018	http://d-cell.ucsd.edu/	38	19
29153838	Herring CA et al.	Cell Systems	2018	https://github.com/KenLauLab/pCreode Data: GSE102698	31	61
29610481	Zhang W et al.	Nat Genetics	2018	GSE10129	28	22

Figure 11: CSBC Shared Data and Tools
(A) CSBC data sets and tools were collated from CSBC publications (2017-2019) by searching for GEO, SRA, or BioProject accession numbers (data) or github or similar web repositories (tools). (B) Number of publications containing CSBC data sets or computational tools. (C) Top ten publications with respect to the number of citations (publications).

Building an inclusive field through community expansion

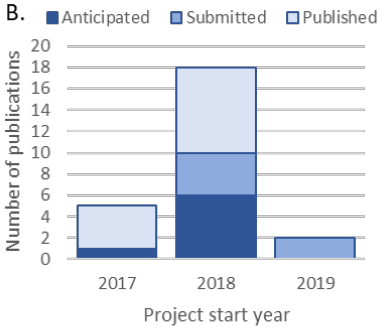
The most effective way to expand the cancer systems biology community is by demonstrating that systems biology approaches lead to important advances in cancer research and thus recruiting new participants through their desire to implement similar approaches to answer their own research questions. The accompanying CSBC scientific highlight review speaks to how CSBC investigators are doing just that. In this section we will discuss additional community building opportunities facilitated by NCI program staff and by CSBC investigators.

In addition to direct support of the CSBC, during the period 2016-2020 the NCI also co-organized six CSBC-associated scientific workshops that brought together CSBC investigators with domain experts to discuss the state of the science and to promote collaboration (2016 Synthetic Biology Approaches to Cancer Systems, 2017 Systems Approaches to Cancer Metastasis [7], 2017 Mathematical Oncology Meeting, 2018 DKFZ-NCI Workshop on Cancer Systems Biology, 2019 Mathematical Oncology Meeting, and the upcoming 2020 CSBC-CTD^2 Multidisciplinary Approaches to Understand Cancer Treatment Resistance). The 2017 and 2019 Mathematical Oncology Meetings continued a set of semi-annual conferences that launched in 2013 and are specifically tailored to the needs of the math modeling community. The efforts have produced the "Mathematical Oncology Handbook", a bioRxiv collection of short manuscripts about the various mathematical oncology and systems biology approaches. The NCI has also co-organized the 2016, 2018, and upcoming 2020 Systems Approaches to Cancer Biology (SACB) conference held at Woods Hole, MA. The SACB meeting is organized by the Association of Cancer Systems Biologists, a non-profit organization that arose from early stage investigators supported by previous NCI systems biology efforts. Over 50% of the 110 attendees at the 2018 SACB were *not* participants in CSBC indicating strong interest in the field. Finally, with the support of NCI program staff, CSBC investigators proposed and led several cancer systems biology focused plenary, mini-symposium, 'Meet the Expert', and educational sessions at the 2017-2019 AACR Annual Meetings to highlight how systems biology approaches are being employed within the cancer research community.

A.

CSBC pilot projects overview	
Indicator	**Value**
# pilot project funded	56
# projects receiving additional funding	18
# publications	12
# public data sets	5
# public tools	2

B.



C.

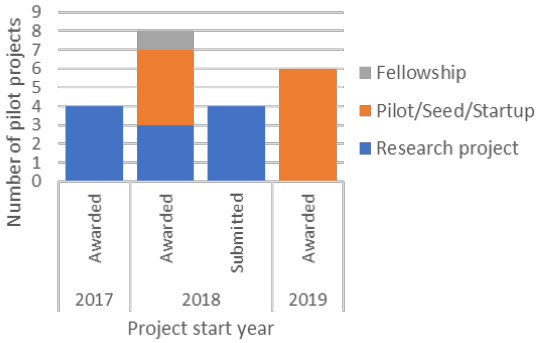


Figure 12: CSBC Pilot Projects
Each CSBC U54 Research Center administers a $50,000 direct cost pilot project fund. (A) Pilot project overview (2016-2019), data obtained from Research Performance Project Report and self-report. (B) Publications as result of CSBC pilot project support. (C) Research funding obtained as result of CSBC pilot project support.

CSBC investigators in U54 Research Centers are required to expand the scope of their Center to support new systems biology projects and early stage systems biologists through a $50,000 direct cost (per year) pilot project fund that is coordinated through the U54 Administrative Core. The pilot project solicitation and award process is individualized to the needs of each U54 Research Center, but most have indicated that pilot project funds are used to support the generation of preliminary data in support of federal funding with an emphasis on early stage investigators. Fifty-six pilot projects were awarded within CSBC U54 Research Centers from 2017-2019 (**Figure 12A**). In all cases, pilot projects were led by investigators who were not originally listed as key personnel on the CSBC award. Pilot projects resulted in multiple publications (**Figure 12B**) and successful grant funding, including 2 NCI R21s, 2 NCI U01s, and 3 NCI R01s research projects (**Figure 12C**). When surveyed, CSBC investigators commented that the pilot project program allowed them to quickly bring in matching or complementary expertise, initiate new collaborations, and support early stage projects and junior investigators in work that contributes to larger proposals. These data suggest that investigator-controlled distribution of modest, flexible pilot project funds is an effective method to expand the systems biology field beyond the current roster of CSBC investigators.

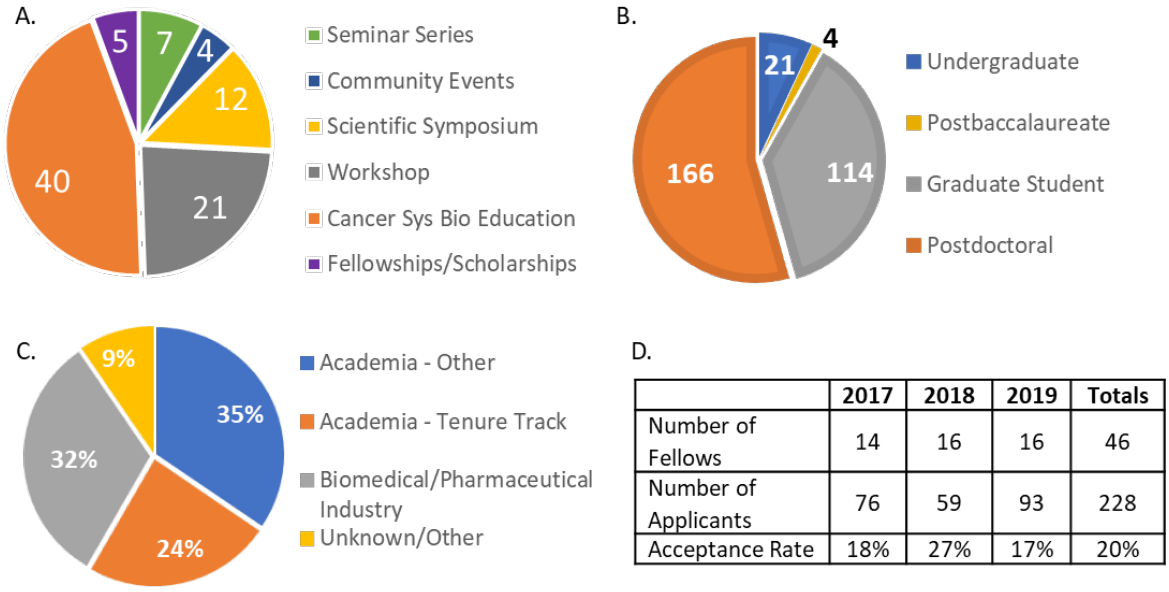
In addition to promoting new cancer systems biology research, each CSBC U54 Research Center includes a required Outreach Core with a minimum budget of $100,000 direct cost per year. The Outreach Core coordinates scientific training and outreach activities that are representative of the overall research theme of the Center and promote cancer systems biology at all career stages. Approximately 50% of CSBC Outreach Core activities are focused on training and will be discussed in the next section (**Figure 13**). The remaining Outreach Core activities are wide ranging and include scientific symposium, training workshops on specific computational or experimental methods, community events meant to introduce systems biology and computational biology approaches to the lay public, and cancer systems biology-specific seminar series. As an example, see the Stanford U54 CCSB Seminar Series which in 2019 featured eight CSBC investigators outside of the Stanford U54 as speakers. Appendix B includes a list of public scientific symposia and training workshops hosted by CSBC U54 Research Centers from 2017-2020. Symposia and workshops are regularly communicated to CSBC investigators via the monthly CSBC Steering Committee meeting. Multiple U54 Research Centers (OHSU, ASU, Vanderbilt) have offered support to junior investigators from other CSBC sites to join workshops and training opportunities. Finally, highlights of community events organized by CSBC U54 Research Centers include the Yale Cancer Answers radio show that is partially funded by CSBC and features Yale systems biology investigators, and the Arizona State University "Endless Forms Most Beautiful" cactus garden that provides a new way to think about genetic mutations in multicellular organisms. In sum, the required Outreach Core activities provide an additional opportunity to promote collaboration and provide a platform for dissemination of Center research, tools, and concepts to both scientific and lay audiences.

Recruiting of new cancer systems biologists through targeting training

There are multiple, often strong, viewpoints regarding the most effective way to marry computational and experimental science. From a cancer systems biology perspective, what is clear is that the most effective practitioners have a strong, clear understanding of cancer biology or oncology and, at the least, a keen appreciation of computational and mathematical models, their underlying assumptions, and their appropriate application. A cancer systems biologist applies the tools of bioinformatics, statistics, computer science, physics, engineering, developmental, molecular, and cancer biology to discover how the parts of a system operate together to manifest in a pathological state. In this section we will describe the efforts taken within the CSBC U54 Research Centers and by the NCI CSBC Program Staff to train the next generation of cancer systems biologists.

Efforts led by CSBC Investigators

Most outreach activities pursued by CSBC U54 Research Centers focus on training efforts that span multiple education stages including K-12, undergraduate, and post-graduate (**Figure 13A**). Graduate and postdoctoral opportunities include fellowship programs, such as the Stanford Cancer Systems Biology Scholars program and UCSD/UCSF Cancer Systems Biology Training Program, which are both supported in part by the CSBC. A variety of undergraduate research opportunities are supported with U54 Research Center funds. For example, salaries for 28 undergraduate students were funded by the CSBC in 2019 across U54 sites. Training opportunities focused on high school students include the HackBio program created by the leader of the UT Health Science U54 Research Center Outreach Core. Another example is the "CaST Scholars Program", a CSBC-sponsored program originally designed by the Columbia University Outreach Core to provide fully supported research, community, and mentoring experiences for students from the City University of New York (CUNY) and was recently expanded to high school students from economically challenging backgrounds through partnership with the Palazzo Strozzi Foundation. Finally, multiple U54 Research Centers have reported activities aimed at elementary and middle school students. Examples include two hands-on modules from the Yale U54 Research Center aimed at introducing middle school students to coding through building and controlling electronic circuits and to biopolymers, such as actin, through 3D printed models.



	2017	2018	2019	Totals
Number of Fellows	14	16	16	46
Number of Applicants	76	59	93	228
Acceptance Rate	18%	27%	17%	20%

Figure 13: CSBC Outreach and Training Activities
Data on CSBC U54 Research Center outreach and training activities was collected from annual research progress reports and through direct survey of Center outreach staff. (A) Number and type of outreach activities reported across U54 Research Performance Progress Reports (RPPRs, 2017-2020). (B) Number and title of trainees participating in CSBC research, as reported in U54 RPPRs (2017-2020). (C) Current position of CSBC trainees who progressed in their career during the CSBC funding period (2017-2020).

Training and outreach activities are shared by the members of the consortium through the CSBC Education and Outreach Working Group (EOWG). Meeting monthly and chaired by CSBC investigators, the group provides a platform for education- and outreach-focused U54 staff and investigators to share information about the activities sponsored by the CSBC. The group also partakes in trans-consortium efforts and has spearheaded a specific effort to teach scientists about the value of communicating to the general public through mentored "public interest pieces" which have been created in collaboration with cancer patient advocates. An example of another science communication-focused activity presented to the group by one EOWG member and emulated in other Research Centers includes the "flipped science fair" where graduate students present posters to high-school students who serve as science fair judges. The EOWG also invites outside speakers to share their programs and experiences and a focus of the 2019-2020 meetings has been understanding how to increase the diversity of systems biology investigators.

Efforts led by CSBC NCI Program Staff

As of May 1, 2020 funding from the CSBC program supported the participation of 732 investigators, with 305 trainees at all levels, in CSBC research activities from 2016-2020 (**Figure 13B**). Most trainees remain within the program, but a total of 70 have left the CSBC by progressing in their careers with 41 having moved to new academic positions and 22 having taken positions within the biomedical and pharmaceutical industries (current position of 7 trainees was unknown at time of writing). A total of 19 postdocs and research/staff scientist-level investigators have started tenure track faculty positions after training in a CSBC lab, representing 24% of all trainees and research/staff scientist investigators that have left their CSBC positions between 2017-2020.

Starting in 2007, the NCI has organized a systems-biology focused "Junior Investigator Meeting" to support interdisciplinary graduate students and postdoctoral researchers during important career transitions. The current incarnation of the meeting is a crucial component of the CSBC's efforts to support the growth of the cancer systems biology field. The planning committee for the Junior Investigator (JI) Meeting is comprised exclusively of trainees, with input and support from the U24 Coordinating Center and NCI Program Staff (an example agenda

can be found in **Appendix 4**). The purpose of the meeting is to facilitate networking, encourage collaboration, and promote data and idea sharing specifically among early stage investigators who are just launching their independent careers. The meeting focuses on topics not always included in the early stage investigators' formal training such as challenges unique to interdisciplinary scientists, career options, and funding opportunities. In total, 91 CSBC-associated trainees have attended meetings in 2017, 2018, and 2019 (link contains Tweets from attendees on Twitter). Former systems biology-focused JI meeting participants from 2007-2014 are responsible for the organization of the Systems Approaches to Cancer Biology conference and the formation of the Association of Cancer Systems Biologists organization discussed under the section "*Building an inclusive field through community expansion"* above. Several current CSBC awardees, including PIs on U01s (Aaron Meyer, UCLA U01; Matthew Lazzara, UVA U01; Shannon Mumenthaler, USC U01) and U54s (Laura Heiser, OHSU U54; Jeffrey Chang, COH U54), were involved in JI meetings during their training suggesting that former participants have become successful members of the cancer systems biology field.

In addition to supporting post-graduate trainees within CSBC labs, the NCI also supports the introduction of undergraduate students to systems biology. Similar to the Junior Investigator meeting, the CSBC/PS-ON Summer Undergraduate Research Program is a continuation of a program started as part of the NCI's Integrative Cancer Biology Program (ICBP). Although focused specifically on systems biology from 2006-2016, in 2017 the program was broadened to include participation of the Physical Sciences-Oncology Network. From 2017-2019 the CSBC/PS-ON Summer Undergraduate Research Program has supported 46 rising Junior and Senior undergraduate students as they complete a 10-week research fellowship at CSBC or PS-ON research sites located across the country (**Figure 13D**). Students receive a $4000 stipend and housing at their research site as well as transportation to and from their fellowship location. Fellows also gather for a 2-day summer undergraduate research conference held at the NIH to present their research, network with the other fellows, and explore various topics related to future careers in scientific research. A highlight of the meeting is a session where students tour the NIH Clinical Center and interact with cancer patient advocates. Of the 46 students who have completed the program between 2017-2019, 21 are still undergraduates working towards their degrees and 25 have completed their undergraduate degrees. The status of 17 of the 25 graduates was determined for tracking purposes; 10 are graduate students in biomedical fields and 7 have science-focused careers (with two completing postbaccalaureate fellowships at the NIH, **Box 1**). Many of the students have co-authored papers with their research mentors at the CSBC/PS-ON Summer Undergraduate Research Program site.

In addition to the research projects completed at their research site, students are required to participate in bi-weekly videoconferences featuring a systems biology focused curriculum referred to as the "mini-DREAM challenge". The mini-DREAM challenge, hosted on the U24 Coordinating Center Synapse platform, introduces basic concepts in bioinformatics and mathematical modeling with an emphasis on combining insights gained from RNA-seq analysis with a mathematical model of cell migration, all within the broader context of understanding breast cancer metastasis. The course highlights the value of data sharing, methods comparison, and collaborative science. The mini-DREAM course represents the major collaborative project of the CSBC EOWG. All course materials were created by members of the working group and the course has evolved from 2017-2020 to become a modular teaching tool. All materials are publicly available, and the working group aims to publish the mini-DREAM challenge framework so that it may be used by others. Because the students arrive at the challenge with a very broad range of backgrounds (from no programming experience to no biology experience), the organizers are interested in conveying the challenges with creating content for the purpose of creating interdisciplinary scientists. The mini-DREAM challenge is intended to be a lasting contribution from the CSBC EOWG to the systems biology field.

Box 1

In 2017, Dylan Hirsch completed his summer undergraduate research experience in the lab of Dr. Christina Leslie, a CSBC U54 grantee at Memorial Sloan Kettering Cancer Center. During his time in the lab, his practical coding and statistics skills radically improved. Understanding the mathematical theory that underlies coding and statistics is critical, but doing research taught him how to put that theory into use in a way problem sets did not. He also gained exciting experience working with new data types that allowed him to peer into cellular behavior from new angles. Most importantly, he learned much more about the "soft skills" of science that are required to tackle open-ended computational biology research. Whereas learning about computer science or biology in a classroom comes with well-curated problems, computational biology research does not. He loved working with Dr. Leslie and the other researcher in her lab, and that experience working in the lab was what made him confident that research was the right direction for his future career path.

Hirsch said, "I cannot overstate how helpful the CSBC summer program was, not just to my professional development, but to my scientific training. In my mind, the internship really marks the beginning of my development from a student into a scientist."

After his experience at MSKCC, Hirsch wanted to learn more about omics-based research, which ultimately lead him to pursue a postbaccalaureate fellowship at the NIH after graduating from Johns Hopkins in 2018. In John Tsang's group at NIAID, he works on analyzing omics and deep immune phenotyping data from patients with rare immune diseases to learn more about clinical and basic immunology. His NIAID fellowship will be ending soon, but he is the recipient of an NSF graduate research fellowship and will begin pursuing his PhD in Bioengineering in the fall of 2020 at MIT.

Closing remarks

This report is intended to support the development of strategic recommendations for future CSBC program implementation and, more broadly, NCI's decisions about how to best support cancer systems biology going forward. While assessing the short term impact of basic research can be difficult, the data presented here are an attempt to provide insight on progress towards CSBC program goals: to advance understanding of mechanisms that underlie fundamental processes in cancer, support the development and broad application of these approaches in cancer research, and support the growth of a strong and stable research community in this field. Because this report focuses on program-level goals, program-level summary statistics have been provided whenever possible. We encourage you to request additional data as needed and invite you to interact with investigators and trainees during the 2020 virtual CSBC Annual Investigator Meeting in September. We are very appreciative of panelist effort to provide feedback on the progress, impact, and future direction of NCI support of cancer systems biology research.

Acknowledgements

Many NIH offices and personnel provided advice or assisted in identifying resources for this evaluation report. Particular acknowledgement goes to the following:

- NCI Division of Cancer Biology: Erin Wetzel and Tamika Carney participated in data collection and analysis regarding CSBC data sets, tools, training, and outreach activities. Nas Zahir provided input and valuable discussion.
- NCI Center for Research Strategy provided consultation on evaluation design and available NIH resources to assist with evaluation. Specific thanks to James Corrigan, Eddie Billingslea, Melissa Antman and Christine Burgess.

- NIH Library: Bibliometrics analysis in Figure 3 and 4 were provided by Ya-Ling Lu, Informationist at the NIH Library.
- Digital Science *Dimensions for NIH*: This report was written using data obtained from Digital Science's *Dimensions for NIH* platform. (Data access dates are included in *Appendix 2: Methods*, below.) Access was granted to subscription-only data sources under a license agreement. Holly Wolcott and Derek Denning, NIH contractors, provided advice on bibliometrics analysis and portfolio definition approaches, as well as assisting with data access and identification of similar applications to awarded CSBC applications.
- Holly Wolcott and Derek Denning from NHLBI provided advice on bibliometrics analysis and portfolio definition approaches, as well as assisting with data access and identification of similar applications to awarded CSBC applications.
- NIH Office of Portfolio Analysis: OPA staff provided consolation and assistance with identifying application citation information and assistance with the *iSearch* platform as needed to access grant application and publication information. Specific thanks to B. Ian Hutchins and George Santangelo.

Appendix

APPENDIX 1: NCI Cancer Systems Biology Consortium (CSBC) Portfolio (June 1, 2020)

Award Year	Grant Title (CSBC website link)	NIH RePORT Link	Contact PI	MPIs or U54 Project/Core Leads
CSBC U54 Research Centers				
2016	Coordination Center for Open Collaborations in Systems Biology	U24 CA209923	Justin Guinney (Sage)	
2016	Center for Cancer Systems Therapeutics (CAST)	U54 CA209997	Andrea Califano (Columbia)	Barry Honig (MPI), Cory Abate-Shen, Charles Karan, Andrew Kung, Ken Olive, Raul Rabadan, Peter Sims, Nicholas Tatonetti, Denis Vitkup, Harris Wang
2016	The CSBC Research Center for cancer systems immunology at MSKCC	U54 CA209975	Christina Leslie (MSKCC)	Alexander Rudensky (MPI), Joan Massague, Dana Pe'er, Andrea Schietinger, Joao Xavier
2016	Modeling the role of lymph node metastasis in tumor-mediated immunosuppression	U54 CA209971	Sylvia Plevritis (Stanford)	Garry Nolan (MPI), Edgar Engleman, Andrew Gentles, Christina Kong, John Sunwoo, Rob Tibshirani
2016	Systems analysis of phenotypic switch in control of cancer invasion	U54 CA209992	Andre Levchenko (Yale)	Gunter Wagner, Murat Acar, Sidi Chen, Rong Fan, Farren Isaacs, Mark Lemmon, Michael Murrell, Jessie Reinhardt
2017	Quantitative and functional characterization of therapeutic resistance in cancer	U54 CA217377	Scott Manalis (MIT)	Doug Lauffenburger (MPI), William Hahn (DFCI), Christopher Love, Alex Shalek, David Weinstock (DFCI)
2017	Cancer Cell Map Initiative (CCMI)	U54 CA209891	Nevan Krogan (UCSF)	Trey Ideker (MPI, UCSD), Alan Ashworth, Laura van't Veer, David Agard, Laura Esserman, Jennifer Grandis, Natalia Jura, Prashant Mali (UCSD), Pablo Tamayo (UCSD), Silvio Gutkind (UCSD), Jill Mesirov (UCSD)
2017	Measuring, Modeling, and Controlling Heterogeneity (M2CH) Research Center	U54 CA209988	Joe Gray (OHSU)	Rosalie Sears (MPI), Emek Demir (MPI), Claire Tomlin (UC Berkeley, MPI), Laura Heiser, Jim Korkola, Young Hwan Chang, Xiaolin Nan, Andrew Adey, Ellen Langer, Michel Nederlof, Damir Sudar
2017	Systems analysis of epigenomic architecture in cancer progression	U54 CA217297	Tim Huang (UT San Antonio)	Qianben Wang (Duke, MPI), Victor Jin (MPI), Zhijie (Jason) Liu, Wei Li (UCI), Chun-Liang Chen, Zhong Chen (Duke), Jianhua Ruan
2017	Combating subclonal evolution of resistant cancer phenotypes	U54 CA209978	Andrea Bild (City of Hope)	Ravi Salgia , Jeffery Chang (UT), David Bowtell (Peter Mac Cancer Center), Frederick Adler (Utah)
2018	Arizona Cancer and Evolution Center (ACE)	U54 CA217376	Carlo Maley (ASU)	Darryl Shibata (MPI, USC), Trevor Graham (Barts Cancer Institute UK), Athena Aktipis, Amy Boddy (UCSB), Christina Curtis (Stanford), Hannah Kokko (Zurich), Tara Harrison (NC State), Li Liu, Joshua Schiffman (Utah), Yinyin Yuan (ICR London)
2018	Phenotype heterogeneity and dynamics in small cell lung cancer	U54 CA217450	Vito Quaranta (Vanderbilt)	Jonathan Irish, Alissa Weaver, Ken Lau, Jonathan Lehman, Christine Lovly, Julien Sage (Stanford), Shyr Yu

2018	Complexity, cooperation, and community in cancer	U54 CA217378	John Lowengrub (UC Irvine)	Arthur Lander (MPI), Marian Waterman (MPI), Anand Ganesan, Richard Van Etten, Christopher Hughes, Kai Kessenbrock, Natalia Komarova, Devon Lawson, Dominik Wodarz, Suzanne Sandmeyer
2018	Systems pharmacology of therapeutic and adverse responses to immune checkpoint and small molecule drugs	U54 CA225088	Peter Sorger (Harvard)	Marcia Haigis, Arlene Sharpe, Artem Sokolov, Keith Flaherty, Frank Hodi (DFCI), Conor Evans, Genevieve Boland, Nicole Leboeuf, Sandro Santagata, Ryan Sullivan (MGH)
CSBC U01 Research Projects				
2017	Systems approaches to understanding the relationships between genotype, signaling, and therapeutic efficacy	U01 CA215798	Doug Lauffenburger (MIT)	Kevin Haigis (MPI, DFCI), Ken Lau (Vanderbilt), Wilhelm Haas (MGH)
2017	Phenotype transitions in small cell lung cancer	U01 CA215845	Vito Quaranta (Vanderbilt)	Carlos Lopez (MPI), Ken Lau
2017	Precision lung cancer therapy design through multiplexed adapter measurement	U01 CA215709	Aaron Meyer (UCLA)	Eric Haura (MPI, Moffitt), Forest White (MIT)
2017	An integrated systems approach for incompletely penetrant onco-phenotypes	U01 CA215794	Kevin Janes (UVA)	Kristin Atkins, Jennifer Harvey, Christiane Fuchs (Helmholtz-Muenchen), Fabian Theis (Helmholtz-Muenchen)
2017	Model-based prediction of redox-modulated responses to cancer treatments	U01 CA215848	Melissa Kemp (Georgia Tech)	Cristina Furdui (MPI, Wake Forest)
2017	Quantifying multiscale competitive landscapes of clonal diversity in glioblastoma	U01 CA220378	Kristin Swanson (Mayo Arizona)	Leland Hu (MPI), Ross Mitchell (MPI), Nahn Tran (MPI)
2018	A plasticity and reprogramming paradigm for therapy resistance at the single cell level	U01 CA227550	Arjun Raj (UPenn)	Ravi Radhakrishnan (MPI) & Ashani Weeraratna (MPI, Johns Hopkins), Abhyudai Singh (Delaware), Junwei Shi
2018	Eco-evolutionary dynamics of non-small cell lung cancer to immunotherapy: Response and resistance	U01 CA232382	Alexander (Sandy) Anderson (Moffitt)	Robert Gatenby (MPI), Scott Antonia (MPI, Duke)
2018	Rewiring of regulatory networks in breast cancer by transcription factor isoforms	U01 CA232161	Marc Vidal (Harvard)	Martha Bulyk (MPI), Juan Fuxman Bass (MPI, Boston University), David Hill
2018	Multiscale systems biology modeling to exploit tumor-stromal metabolic crosstalk in colorectal cancer	U01 CA232137	Stacey Finley (USC)	Shannon Mumenthaler (MPI) & Paul Macklin (MPI, Indiana)
2019	Optimal control models of EMT for the design of	U01 CA243007	Matthew Lazzara (UVA)	Todd Bauer, Babatunde Ogunnaike (Delaware), Ben Stanger (UPenn)

	pancreas cancer combination therapy			
2019	Single-cell characterization of tumor and microenvironment co-evolution in peripheral T-cell lymphomas	U01 CA243073	Raul Rabadan (Columbia)	Teresa Palomero (MPI)
2019	Experimental-computational synthesis of altered immune signaling in breast cancer	U01 CA232216	Peter Lee (City of Hope)	Russell Rockne (MPI), Andrei Rodin (MPI)
2019	Identification of adaptive response mechanisms in breast cancer by information theory and proteomics	U01 CA238720	Forest White (MIT)	Nataly Kravchenko-Balasha (MPI, Hebrew University)
2019	Predictive modeling of the EGFR-MAPK pathway for triple negative breast cancer patients	U01 CA238475	Charles Perou (UNC)	Timothy Elston (MPI)
2019	Reverse sensitivity analysis for identifying predictive proteomics signatures of cancer	U01 CA227544	H. Steve Wiley (PNNL)	Wei Qian (MPI), Herbert Sauro (MPI, Univ. Washington)
2019	Systems analysis of aggressive prostate cancer pathology	U01 CA231978	Jim Costello (UC Denver)	Scott Cramer (MPI)
2020	Towards rational design of combination therapeutic targets	U01 CA243072	Kai Tan (CHOP)	Sarah Tasian (MPI)
2020	Multiscale computational models for predicting optimum immune checkpoint and targeted therapy schedules	U01 CA243075	Trachette Jackson (Michigan)	Alexander Pearson (UChicago), Randy Sweis (UChicago)
2020	Systems analysis of cell-cell communication networks and immune activity in the melanoma tumor microenvironment	U01 CA238728 pending	Kathryn Miller-Jensen (Yale)	Marcus Bosenberg (MPI)

APPENDIX 2: Methods

Publication data

Publications produced by CSBC grantees in 2015 through 2020 were identified in March 2020 through the *iSearch* platform provided by the National Institutes of Health Office of Portfolio Analysis (OPA) (n=376). Bibliometric analysis in this report considers two different time periods. The majority of the analyses (**Figure 3**) are based on the period 2017 to 2019 to include recent publications funded by CSBC. Publications from 2020 were excluded to ensure a fixed data set. The citation indicators presented in **Figure 4**, however, are based on 2017 only, to ensure that publications in this analysis are at least 2 years old and have had enough citations for valid analyses. Because of the availability of different bibliometric tools, slightly different subsets of these data were used, as summarized below.

- Figure 3A-C and E used publications available in Web of Science (n=343). Citation impact data were obtained from the Web of Science *InCites* tool.
- Figure 3D used CSBC publications available in the Digital Science and Research Solutions, Inc software "*Dimensions for NIH*" (accessed on Feb. 10 2020). This tool provides alternative metadata and categorizations, but it does not draw on the same publication data source as Web of Science or NIH OPA *iSearch*. Here, all publications identified through *Dimensions for NIH* with assigned PMIDs (2017 to 2019) were used (n=377).

Grant application data

NCI grant application data were identified on April 25, 2020 through NIH OPA *iSearch*. The NCI grant application data set analyzed includes all competing applications (Application Type 1, 2 or 9) administered by NCI from 1999 through 2020 (n=247,210). Analysis on application made use of the following application-level annotations available from *iSearch*: Award Status; Activity Code; Parent RFA Notice; RFA PA Number; Study Section. Individual grants were additionally annotated with derived fields including whether they were unsolicited (submitted to a NIH "Parent" Funding Opportunity Announcement (FOA) or similar) or solicited (submitted to a NCI-specific FOA).

Additional annotations for each grant application were obtained from other sources as summarized below:

- Research application **(Figures 5-8)**: Data from NIH Office of Extramural Research (OER) Categorization of Activity Codes (https://grants.nih.gov/grants/funding/ac_search_results.htm) were integrated via Activity Code. See more details below.
- NCI SB grant application **(Figures 5-8)**: Applications identified through Digital Science's *Dimensions for NIH* keyword search of complete applications were integrated based on application ID. Data were accessed on May 1 2020. See more details below.
- The International Cancer Research Partnership Common Scientific Outline (ICRP-CSO) **(Figure 5B)**: Application-level annotations for all awarded NCI applications were accessed from Digital Science's *Dimensions for NIH* on May 1 2020, and were integrated based on application ID.
- Citing CSBC publications **(Figure 6)**: Application IDs that cite CSBC publications from 2017 to 2019 were identified by the NIH OPA. Data were integrated based on Application ID and included applications to any NIH Institute or Center. NIH-wide citation data provided by OPA were analyzed separately (**Figure 6A)**.
- Current Regular Standing Study Sections or Continuing SEPs: Study Section-level information accessed from https://public.csr.nih.gov/StudySections/StandingStudySections and integrated based on Study Section name.

Additionally, analysis of application success rates drew on NIH-wide application data. Data were downloaded from *iSearch* for competing applications submitted to NIH in Fiscal Years 2014 through 2019 and assigned to one

of the following study sections: MABS, BDMA, CG, TPM, MONC, TCB, TME, CAMP, CBSS, BMCT, DT, CONC, DMP, RTB, GCAT. Data were further annotated with fields described above for NCI application data.

Identification of "research applications"

Applications are designated as "research" based on NIH Office of Extramural Research (OER) Categorization of Activity Codes (https://grants.nih.gov/grants/funding/ac_search_results.htm). OER categories considered as "research" include: Research Projects, Research Program Projects and Centers, Research-Related Programs, Institutional Training And Director Programs, and Cooperative Agreements with a few exceptions. Funding for Intramural, SBIR/STTR, or Clinical Center research were excluded, as were Research Project Awards for Education (ex. F32). K99/R00 projects were included. Specific activity codes that fall under included categories, but have been excluded from the "research" categorization, are as follows:

- Research Projects: R25 (Education Projects); R38 (Mentored Research Pathway to Residency); RL5 (Linked Education Project); RL9 (Linked Research Training Award); UE5 (Education Projects - Cooperative Agreements); R41, R42, R43, R44, SB1 (SBIR/STTR)
- Research Program Projects and Centers: P30 (Cancer Center)
- Research-Related Programs: SI2 (Intramural Clinical Scholar Research Award)
- Institutional Training and Director Programs: D43 (International Research Training Grants)
- Cooperative Agreements: U09 (Scientific Review and Evaluation); U13 (conference); U43, U44, UT1 (SBIR/STTR)

Note that activity codes that are not found in NCI competing applications from 1999 through 2020 were not categorized.

Definition of the NCI systems biology grant application portfolio (NCI SB)

The NCI systems biology application portfolio (NCI SB) encompasses the research grant applications submitted to NCI using the terms "system biology", "system analysis", and "mathematical modeling". Although broad, the search terms were selected based upon a systematic analysis as described below.

Three steps were used to select appropriate keywords to capture the NCI SB portfolio.

1. Common keywords (or "concepts") in awarded CSBC grants were identified using the natural language processing provided in Digital Science's *Dimensions for NIH* (accessed on May 1 2020):
    - Keywords were identified (n=3301) and those found in greater than 5 percent of awarded applications (n=811) were kept for further analysis.
    - Keywords also identified in at least one CSBC Funding Opportunity Announcement (RFA or PAR) were retained (n=23 out of 82 total concepts identified in FOAs).
2. Keywords shared with similar scientific areas were filtered:
    - This step is to exclude identified terms that are not specific to systems biology.
    - First, Digital Science's *Dimensions for NIH* was used to identify non-CSBC grant applications with similarity to individual awarded CSBC grant applications, using a machine learning approach ("*MoreLikeThis*" search in Solr, https://lucene.apache.org/solr/guide/6_6/morelikethis.html) that assigned a similarity score based on application titles, abstracts, and specific aims. These data were accessed on Feb. 27 2020.
    - Program staff reviewed a subset of these similar applications to determine a score threshold beyond which Program would generally not consider the application to be "cancer systems biology" as defined by the program (similarity score of less than 0.5). Keywords found to be present in 5% or more of these "semi-similar" applications were filtered from the search list.
3. Keywords utilized to define the "NCI SB" portfolio were selected:
    - After filtering, 16 terms remained (here listed by frequency found in CSBC awards): system biology, research project, dynamic, system biology approach, system analysis, tumor

heterogeneity, individual cell, tumor evolution, mathematical modeling, scale, quantifying, ecosystem, tumor phenotype, behavior, tumor initiation, tumor stroma.

- Three queries were employed to assess the degree of similarity of applications to the NCI Program definition for cancer systems biology and the breadth of portfolio defined.
- Each query was run on a set of competing research applications submitted to NCI in Fiscal Year 2019 (n=14,479 applications).
- Query 1 – "system biology OR system analysis OR mathematical modeling" - identified 345 applications and review suggested strong similarity of identified applications with definition of the field as indicated by CSBC FOAs.
- Query 2 – Query 1 + "tumor heterogeneity" OR "tumor evolution" OR "tumor phenotype" – identified 868 applications.
- Query 3 – Query 1 + Query 2 + "dynamic" OR "scale" – identified 3368 applications.
- Review suggested that the Query 2 and 3 expanded portfolios included grants less fitting of NCI Program definition of systems biology (i.e. narrow focus on biology with no computational approach or systems-level view).
- Thus, all NIH applications from Fiscal Years 1999 through 2020 were searched using Query 1, identified applications were downloaded on 3/28/2020, and used in subsequent analysis. NCI-specific applications were filtered and used as the "NCI SB" portfolio.
- Note: Substitution of "systems biology" or "systems analysis" for "system biology" and "system analysis" did not change results.

CSBC Data set and Tools

The CSBC Coordinating Center provided publication-level annotation of data sets produced by CSBC from 2017 through 2019. Data sets were identified through GEO and SRA unique IDs (GSE, SRP) associated with a PMID via PubMed. Program staff provided publication-level annotation of tools produced by CSBC by searching publication text for GitHub links or other indications of a publicly shared method or model. These publications were further filtered to identify only those that introduced a specific method. To acquire the number of citations in publications or in NIH applications, data were integrated with publication and application data described above.

NCI investigator feedback on CSBC activities

To assess whether CSBC programmatic activities were effective at (1) establishing collaborations between CSBC investigators and (2) sharing expertise and resources across the consortium, all CSBC principal investigator-level members were invited to participate in a survey in January 2020. Survey responses were welcome from any member of the consortium. Fifty-six individuals responded to the survey, including roughly 33% of CSBC PI and MPI-level investigators. For this survey, collaboration was defined to require, at a minimum, (1) sharing information and resources, (2) defined roles in the pursuit of a shared objective, (3) frequent communication, and (4) shared decision making.

Pilot project and administrative supplement award information

Pilot project and administrative supplement outcome data were collected from U54 research progress reports and data requests to U54 centers. Data indicating "submitted" or "anticipated" are based on optional self-report and so may be underestimates.

APPENDIX 3: CSBC U54 Research Center Symposia and Workshops

Center Name	Scientific Symposium	Training/Workshop	Dates
ASU		Cancer and Embryo Development Workshop	January 17-18, 2020
ASU		Methods for Somatic Phylogenetics Workshop	November 7-8, 2019
Columbia	Columbia Systems Biology Department Symposium		October 2016- present
Columbia		Columbia University School of Professional Studies Masters Course in Systems Biology	January - May 2017
Columbia	Cancer Genomics and Mathematical Data Analysis Symposium		February 7-8, 2018
Harvard		Introduction to singlecell RNA-seq	July 17-18, 2019
MIT	Quantitative Systems Pharmacology Day		April 1, 2019
MSKCC		10th Annual RECOMB/ISCB Conference on Regulatory & Systems Genomics with Dream Challenges	November 19-21, 2017
MSKCC		ICML 2017 Workshop on Computational Biology	July 9, 2017
MSKCC		11th Annual RECOMB/ISCB Conference on Regulatory & Systems Genomics with Dream Challenges	December 8-10, 2018
OHSU		Correlative Light and Electron Microscopy (CLEM), Multispectral SuperResolution Microscopy (MSSRM), and Cyclic-Immunofluorescence (CyclF).	May 7-11, 2018
OHSU	West Coast Cancer Systems Biology Symposium		May 14-16, 2019
Stanford	CCSB Annual Symposium		May 2017-2019
UCI		Advanced Fluorescence Imaging and Dynamics	October 22-26, 2018
UCI	MechBio Conference 2018 – the Mechanome in Action		July 26-27, 2018
UCI	The 2019 UCI Campus-wide Symposium on Basic Cancer Research		May 3, 2019
UCI	Joint UCI-National Taiwan University Symposium in the Life and Health Sciences		December 3, 2019
UCI	17th Annual UCI Immunology Fair		December 5, 2019
UCI		Single Cell Analysis Workshop with Jupyter Hub	May 23, 2019
UCI		Advanced Fluorescence Imaging and Dynamics	October 21-25,2019
UCSF	Cell Mapping Symposium		September 13-14, 2017
UCSF		Integrative Genomic Analysis with GenePattern	November 7, 2017 and December 5, 2017

UTHSCSA		U54 Education and Outreach Student Workshop for our Cancer Systems Biology Program	July 1, 2017
UTHSCSA		Mass Cytometry and DNA Encoded Chemistry in Drug Discovery Workshop	May 28, 2018
Vanderbilt		Bioinformatics: Big Data in Biomedical Research Workshop	August 6-10, 2018
Vanderbilt		Discovery Oriented Data Science Workshop	August 13-17, 2018
Vanderbilt		Summer Institute Workshop	August 5-16, 2019
Vanderbilt		Tumor Heterogeneity Workshop	August 21-22, 2019
Vanderbilt		Single Cell Biology Workshop	January 22-24, 2020
Yale		Microfluidics in System Biology	May 19, 2018
Yale	Yale Systems Biology Institute and Cancer Systems Biology @Yale Symposium		May 18, 2018

APPENDIX 4: 2019 CSBC/PS-ON Junior Investigator Meeting Agenda

Day 1 August 22, 2019		**AGENDA**
Start	**End**	
8:00	**8:30**	Registration
8:30	**8:45**	NCI Introductions
8:45	**9:30**	"Speed-dating"; Introductions/answer questions (4X10mins)
9:30	**10:30**	**Keynote: Kris Wood, PhD. from Duke University**
10:30	**10:45**	Break
10:45	**11:05**	**Session 1 Tumor Heterogeneity & Evolution** ***15-minute Talks:*** **Pasquale Laise, Columbia University:** Systematic network-based analysis identifies novel molecular subtypes conserved in multiple pancreatic cancer cohorts and at the single cell level
11:05	**11:25**	**Cassandra Burdzian, MSKCC**: Single Cell Dissection of Epigenetic and Tumor Ecosystem Dynamics During Pancreatic Cancer Progression
11:25	**11:45**	**Daniel Temko, DFCI**: Deciphering Epigenetic Heterogeneity in Triple Negative Breast Cancer (TNBC)
11:45	**12:05**	**Ming Guo, MIT**: Cell swelling, softening and invasion in a 3D breast cancer model
12:05	**12:25**	**Eszter Lakatos, Barts Cancer Institute**: The evolutionary processes shaping the neoantigen landscape during tumour growth
12:25	**12:45**	***4-minute Pop-talks:*** **Adam Palmer, Harvard:** A majority of clinically successful combination cancer therapies consist of independently effective drugs that overcome tumor heterogeneity but do not exhibit synergistic interaction **Kyle Singleton, Mayo Clinic:** Identifying distinct imaging ecologies in male and female glioblastomas **Giorgio Gaglia, Brigham and Women's Hospital:** HSF1 phase transition modulates stress adaptation and cell fate decisions **Shambhavi Singh, Univ. of Virginia:** Profiling heterogeneous cell-states in luminal breast

		tumors at diagnosis
		Lee Curtin, Mayo Clinic: Characterizing glioblastoma subpopulation dynamics through a combination of patient data, machine learning and mechanistic modeling
12:45	**2:45**	Lunch and Poster Session
2:45	**3:05**	**Session 2 Tumor Metabolism** ***15-minute Talks:*** **Yin Tang, Institute for Systems Biology:** Pleural Effusion-Based Single-Cell Metabolic Phenotyping of Lung Cancer Patients for Informative Diagnostics
3:05	**3:25**	**Lu Ling, Cornell University:** Multiphoton imaging of breast cancer cell metabolic dynamics
3:25	**3:45**	**Walid Abdelmoula, Brigham and Women's Hospital:** Tumor-Specific Metabolic Signatures identified in GBM PDX models using Artificial Neural Networks Analysis of Mass Spectrometry Imaging data
3:45	**4:05**	**Xiang Nan, Houston Methodist:** Imatinib and metformin combination regimen inhibits multiple convergent signaling pathways in Ewing sarcoma: a systematic drug repositioning study
4:05	**4:25**	**Emma Fong, University of Southern California:** Metabolic shifts of KRAS mutant organoids in response to stromal cell-mediated secreted factors
4:25	**4:45**	***4-minute Pop-talks:*** **Bishal Paudel, Vanderbilt:** Disrupting Redox Balance Enhances the Effects of Targeted Therapies in Melanomas **Andrew Raddatz, Georgia Tech and Emory:** Agent-based modeling of the bystander effect due to chemotherapeutic-based ROS generation in HNSCC **Stephen DeCamp, Harvard:** Unjamming and energy metabolism in the epithelial layer **Austin Lefebvre, UC, Irvine:** Mitochondrial Recruitment by Metastatic Breast Cancer is Mitigated by PI3K Inhibition **Joshua Lewis, Georgia Institute of Technology:** Personalized genome-scale metabolic models identify novel diagnostic and therapeutic strategies for radiation-resistant tumors
4:45	**6:30**	Break
6:30	**8:30**	Networking time- Tommy Joe's, 7940 Norfolk Ave, Bethesda
Day 2 August 23		
8:00	**8:20**	**Session 3 Tumor Therapeutics and Resistance** ***15-minute Talks:*** **Luca Gerosa, Harvard:** Single-cell ERK signaling dynamics drive adaptive drug resistance of BRAF V600E cancers
8:20	**8:40**	**Portia Thomas, Vanderbilt:** Evaluating Immune Cell Subtypes in Small Cell Lung Cancer
8:40	**9:00**	**Renee Brady, Moffitt:** Optimizing docetaxel scheduling to delay progression in metastatic prostate cancer patients receiving hormone therapy
9:00	**9:20**	**Nan Jin, UCSF:** The immunological mechanism responsible for NSAIDs effects in PIK3CA-altered head and neck squamous cell carcinoma

9:20	**9:40**	**Simone Pisano, Houston Methodist:** Dendritic cells-derived exosomes as a novel therapeutic approach for metastatic ovarian cancer
9:40	**10:00**	***4-minute Pop-talks:***
		Olga Nikolova, OHSU: Drug response signatures as part of the PRECEPTS Framework (Predictors of Cellular Phenotypes to guide Therapeutic Strategies) ***Jason Conage-Pough, MIT:*** Determining the critical vulnerabilities of Glioblastoma using Proteomics ***Mohsen Malehmir, MGH:*** Epigenomic correlates of immunotherapy response and resistance ***Bill Hong, Northwestern:*** Multifunctional Polymer-Caged Nanobins (PCNs) as a Smart Theranostic Platform: Cancer-Targeting Ability, pH-Responsive Drug Release, and Enhanced MR Contrast ***Louis Hinkle, Houston Methodist:*** Beta blockers improve cancer vaccine efficacy in a dendritic cell-driven manner
10:00	**10:15**	***Break***
		Session 4 Tumor Microenvironment and Metastasis ***15-minute Talks:***
10:15 - 10:35		***Almudena Espin Perez, Stanford:*** Computational approaches to identify molecular mechanisms that influence distant metastasis
10:35 - 10:55		***Jason Andrechak, U Penn:*** Engineered Marrow Macrophages Suppress Growth of Metastatic Models in Immunocompetent Mice
10:55 - 11:15		***Milos Nikolic, U of Maryland:*** Intracellular Mechanics of Cancer Progression Revealed by Brillouin Microscopy
11:15 - 11:35		***Panagiotis Mistriotis, Johns Hopkins University:*** Confined Cell Migration Induces Nuclear Volume Expansion and Blebbing by Triggering RhoA-Mediated Nuclear Influx
11:35 - 11:55		***Zaid Mohamed, MD Anderson:*** Spatial immune- and stromal-based characteristics of biophysical subtypes of pancreatic ductal adenocarcinoma (PDAC)
11:55	**12:15**	***4-minute Pop-talks***
		Ibrahim Chamseddine, Moffitt: Deconvolution of microenvironmental and drug effects on tumor growth using in silico organoids ***Martial Boutcheung Djidjou, Moffitt:*** Mass-spectrometry data analysis nominates new mechanisms of the bi-directional signaling between lung cancer cells and lung fibroblasts ***Huu Nguyen, MIT:*** Multicellular tumor-on-a-chip model for studies of early metastasis ***Eugenia Volkova, Johns Hopkins:*** Ex Vivo Analysis of Matrix Stress Relaxation Origins and Development in Sarcoma
1:30	**2:30**	***Session 5: Kristen Atkins, MD, University of Virginia: Interactive Session on Science Communication***

APPENDIX 5: Additional data provided to expert panel in August 2020 by NCI staff

Dear Cindy, John, Ed, Teresa, and Ken,

Thank you for your thoughtful review of the materials that we provided about the CSBC program. Your additional questions were appreciated and have already provided us some excellent ideas as we move forward. Please find the answers to your questions below. Some of the questions have been grouped together in order to provide more wholistic answers. Our response includes both new data (indicated by the prefix 'Additional') and references to the original evaluation report. We look forward to further discussion at the CSBC Annual Investigator Meeting.

Best regards,
Shannon and Hannah

Questions in black text; Answers in blue text.

1. Will panelists have an opportunity to speak to investigators before the Annual Meeting?

We have worked with STPI and the CSBC Steering Committee chairs to set up a conversation between the panel and CSBC investigators. The conversation will take place on August 31, 2020 and will be moderated by Drs. Laura Heiser (OHSU), Kevin Janes (UVA), and Julie Bletz (Sage Bionetworks). Discussion questions you drafted and edited will be provided to the CSBC investigators ahead of the meeting so that the time can be maximized. NCI Program Staff will not attend the meeting.

2. What is the projected lifetime for a U54 grant? What is on the Horizon for the CSBC program?

A common lifespan of large NIH programs is 10 years, making the possible lifetime of associated grants approximately the same. However, if the CSBC is renewed all U54s would need to recompete for funding so 10 years of funding is not guaranteed.

NCI's goal in supporting the CSBC has been and continues to be promoting systems biology approaches in cancer research. Ultimately, we want to facilitate a transition to a self-sustaining field where investigator-initiated research employing systems biology approaches is supportable through traditional NIH review. Widespread citation of CSBC research by the scientific community in publications and grant applications indicate that systems biology approaches are appreciated, and the associated biology is of high value (**Figures 3, 4, and 7** pages 7-8, 10). Furthermore, the number of NCI grant applications proposing to use a systems biology approach has grown over time (**Figure 5**, page 9). However, there remain challenges associated with review of cancer systems biology applications in standing CSR study sections (**Figure 8**, page 12) which suggest that the CSBC continues to serve an important role in growing the field.

As we look towards the future of cancer systems biology research, it is imperative that we understand if the current structure and focus of NCI support is optimal for building the field, doing impactful cancer research, and training the next generation of cancer systems biologists. There are many areas of basic cancer research that are ripe for a systems biology approach, some of which are listed as topics of interest within the current CSBC U01 funding announcement:

- Decoding dynamic tumor-stroma and/or tumor-immune system interactions, including understanding the mechanisms by which the dynamics of these interactions influence tumor development and/or progression;
- Building integrated models of chemical, molecular, structural, network, and localization information across space and time to understand tumor initiation, progression, and metastasis;
- Quantifying how individual cell states and behaviors (tumor and non-tumor) shape tumor ecosystems, including mathematical modeling of single-cell dynamics or development and testing of combined experimental and computational approaches that quantify how individual cell behaviors manifest at the tumor, organ, and whole-body levels;

- Determining the mechanisms by which external factors, such as the microbiome, affect tumor initiation, progression or treatment.

We encourage discussion with the CSBC investigators to identify areas where particular support might be needed.

3. What is the strength of U54s and U01s compared to traditional R01s in systems biology? How does the impact of CSBC publications compare to external systems biology publications? How do members outside of the CSBC program 'break-in' to the program (i.e. obtain a CSBC grant and/or assimilate into the cancer systems biology community)

The CSBC program utilizes Cooperative Agreement mechanisms (U01 and U54s) versus the traditional R01 and P01 mechanisms in order to facilitate consortium activities and collaboration. For example, the U01 and U54 mechanisms allow NCI to require that grantees set aside funds to attend consortium meetings and require participation within a Steering Committee and within working groups (i.e. Education and Outreach Working Group). The U54 internal pilot project fund is also possible due to the specific funding mechanism. Additionally, the issuance of a Request for Applications (RFA) to support the U54 and the Program Announcement with Special Review (PAR) to support the U01 allowed for a NCI nexus of review focused specifically on cancer systems biology. Finally, the consortium structure provides participants with the opportunity to take advantage of collaborative opportunities and funds specific to the CSBC.

Unfortunately, we cannot currently provide a quantitative comparison of research output between the U01 and R01 mechanism because the CSBC U01 projects are too young for bibliometric impact analysis (publications should be at least two years old for metric stability, but the earliest CSBC U01s started in 2018). However, the CSBC PIs may be able to provide helpful information about the relative strengths of the U01 mechanism. Several U01 awardees will be present at the CSBC Investigator Roundtable, including Kevin Janes, Steven Wiley, Carlos Lopez, Aaron Myer, Matthew Lazzara, and Nhan Tran. We look forward to performing the quantitative comparison between the U01 and R01 mechanisms when we are able because a comparison group of R01s could also help ensure that the publication impact metrics are interpretable with regards output at similar funding levels. Please note that the impact metrics for the 2017 U54 publications presented in **Figure 4** of the report are field-normalized, to control for differences in citation practices across disciplines, which allows these metrics to be interpreted directly as publication impact relative to articles in the same discipline.

Participation in consortium activities and collaborative opportunities may also be of interest to researchers currently outside of the CSBC. Several types of CSBC activities are designed to foster interactions between CSBC and non-CSBC investigators, including workshops and meetings, administrative supplement opportunities and pilot funding programs. For example:

- The NCI has co-organized Systems Approaches to Cancer Biology (SACB) conference held biennially at Woods Hole, MA. Over 50% of the 110 attendees at the 2018 SACB were *not* participants in CSBC. In addition, during the period 2016-2020 the NCI also co-organized six CSBC-associated scientific workshops that brought together CSBC investigators with domain experts to discuss the state of the science and to promote collaboration. More details can be found on page 17 of the report.
- The NCI has utilized the one-year Administrative Supplement mechanism to support new collaborations across a range of structures, including collaborations between CSBC members and external investigators). In addition to research projects, Administrative Supplement awards have supported other collaborative activities such as DREAM Challenges (collectively involving almost 500 computational scientists), Ideas labs (intensive 3-5 day science "bootcamps" designed to rapidly build collaborations), collaboration-building workshops led by CSBC members such as the 2019 West Coast Cancer Systems Biology Symposium, and hackathon events. Further information can be found in **Figure 9** and pages 13-14.
- CSBC investigators in U54 Research Centers are required to expand the scope of their Center to support new systems biology projects and early stage systems biologists through a $50,000 direct cost (per year) pilot project fund that is coordinated through the U54 Administrative Core. When surveyed, CSBC investigators commented that the pilot project program allowed them to quickly bring in matching or complementary expertise and initiate new collaborations. A summary of pilot projects and outcomes can be found on page 17 of the report.

4. Do CSBC trainees continue in the systems biology field? Note that systems biology is an excellent mechanism to engage underrepresented minorities as systems biology research is not location dependent. CSBC outreach activities could be used to focus on engaging these groups.

Of the CSBC trainees (graduate students, postdocs, and research/staff scientists) who have moved on from their CSBC-sponsored positions (**Figure 13**, page 19), 60% continue to pursue systems biology research or are employed in a highly related field (i.e. computational biologist at a pharmaceutical or biotechnology company). Those who are no longer in the systems biology field, and whose current employment status was available, include former trainees in academic or industry positions pursuing more traditional basic cancer biology research or technology development, individuals who are practicing medicine, and individuals who have entered a variety of non-research related fields such as science communication, business consulting, and K-12 education. Trainees who have accepted a tenure track faculty position overwhelmingly stay in the systems biology field, with 21 of 25 new assistant professors indicating that their labs intend to pursue systems biology approaches. These new faculty members are most frequently hired into biomedical engineering and computer science departments but have also found positions within cancer centers (i.e. pathology) and more traditional biology disciplines (i.e. biochemistry and cell biology).

We wholeheartedly agree with the panel that systems biology is an excellent mechanism to introduce underrepresented minorities to cancer research. Examples across the CSBC U54 Outreach Cores that speak to the interest of the CSBC community in recruiting and training a diverse systems biology community include targeted undergraduate research opportunities at Arizona State University (**Additional Appendix 1**), a partnership between Columbia University and the Palazzo Strozzi Foundation to recruit students from underserved schools in New York City into the CaST Scholar Program (**Additional Appendix 2**), and CSBC and NIH R25-supported efforts in systems biology training at the University of California, Irvine (**Additional Appendix 3**).

The NCI CSBC/PS-ON Summer Undergraduate Research Program can also introduce systems biology and physical oncology research to students who may not have similar opportunities at their home Institution. A program description is provided on page 20 of the evaluation report and limited demographic information about applicants and participants is provided here. NCI did not collect race and ethnicity data during the application process, but gender information was collected as an optional field. Of the 298 undergraduate students that applied to the program between 2017 and 2020, 62% identified as female and 48% identified as male. Applications are distributed to CSBC and PS-ON research sites and participants are chosen by the project mentors. Although some sites/mentors have indicated that priority is given to underrepresented minority students or students who do not have the opportunity to perform similar research at their home Institutions, this has not been a requirement of the program thus far. When including the 2020 cohort of students who are just wrapping up the program, 62 students have participated between 2017-2020. Of those participants, 64% identified as female and 36% identified as male. Based on personal knowledge of the participants, and not self-reported data, 16% of the participants are underrepresented populations in STEM.

5. To assess community impact of the CSBC program, could you please provide more information about the data and tools generated in the program?

We also have been interested in better understanding the extent that CSBC data and tools have been disseminated to, and applied by, the research community. Data relevant to this topic can be found below. First, though, we wanted to briefly discuss how the CSBC hopes to impact the cancer research community at a high level, and how the support of publicly available data and resources feeds into these goals. In this context, the CSBC aims to contribute to the cancer research community in three ways:

(1) To advance the mechanistic understanding of cancer by employing systems biology approaches, initially within consortium research projects, and subsequently by informing research performed by the wider scientific community. We

anticipate that the biological insights generated by the CSBC will be the main contribution of the consortium and that those insights will be utilized by individual investigators outside of the CSBC to inform their ongoing research efforts.

(2) To support the growth of sustainable portfolio of cancer systems biology research. We anticipate that the community of cancer systems biology researchers will experience progress towards sustainable funding through a competitive success rate for unsolicited cancer systems biology applications, and the existence of an appropriate locus of review for these applications. Data pertaining to this goal can be found in **Figures 5-8** and pages 8-13.

(3) To build a strong workforce through support of research collaborations and the next generation of scientists. We anticipate that this will benefit the cancer systems biology community through the development of strong research teams with diverse expertise, and successful entry into the federally funded workforce for early stage investigators. For more information about progress towards this goal, see **Figures 9-13** and pages 13-21.

An important secondary outcome of CSBC research is the generation of new datasets and computational tools for use by the community. Developing and disseminating data and tools are important components of supporting an emerging field, in that they help to facilitate the uptake of cancer systems biology approaches by the broader cancer research community. In addition to the information on data and resource sharing, below, CSBC investigators will be able to provide insight into the various ways that CSBC impacts the cancer research community.

a. What/how much data and tools are generated?

At the time that the CSBC evaluation report was prepared, the CSBC had generated and shared more than 199 genomic data sets and 48 computational tools (**Figure 11A and B**, page 16). A list of these resources can be found in **Additional Table 1**. The most highly cited resource publications (**Figure 11C**) introduce tools to process and analyze single-cell sequencing (SEQC, MAGIC, p-Creode) and multiplex imaging data (12plex-IHC, CODEX), and to simulate cell structure and function (DCell 1.4). They also include bulk (GSE89309, GSE92721, GSE93702, GSE101209) and single-cell (PRJNA472387, GSE114397, GSE102698, GSE114397) transcriptomic and epigenetic data sets. In addition, the consortium has generated and shared a variety of non-genomics datasets, including CyCIF, the DFCI Interactome Database, the Columbia PrePPI Database, NDEx (UCSD), and the Stanford Lung-Tumor Microenvironment Interactome.

b. Who is accessing and using the data? How are these data and tools being used externally to CSBC? How transferrable are these tools outside of CSBC?

Bibliometric analysis suggests that CSBC-generated resources are heavily used (**Additional Figure 1**). We examined articles funded by the CSBC and associated with the release of shared data and tools, published in 2017 and 2018 (per standard bibliometrics practice). This set of publications received 23 citations per publication on average, and more than 50% of these articles have a citation counts within the top 10% of all peer-reviewed articles, when normalized for research area and publication year. We also examined basic usage statistics for the six GEO datasets associated with the ten most highly cited CSBC resource-associated publications (**Additional Figure 2**). These datasets (published in 2017 and 2018) have been downloaded 5,005 times.

To provide insight into who is using CSBC-generated resources, we examined articles that cite CSBC resource-associated publications (**Additional Figure 3**). In 2017-2020, these articles received 1430 citations (**Additional Figure 3A**). Of the citing publications, 62% are research articles, 34% are reviews, and the remainder are proceedings papers, letters, and editorial material (**Additional Figure 3B**). Oncology is the research area (or Web of Science Subject Category) that is most highly represented, and the top ten research areas also include cell biology, immunology, biotechnology, and mathematical and computational biology (**Additional Figure 3C**). We further examined which NIH institutes receive grant applications that cite CSBC resource intense publications (**Additional Figure 4**). These publications have been cited 1104 times across 22 NIH Institutes and Centers in 2017-2020. While 56% of these applications were assigned to NCI, a substantial fraction was assigned elsewhere, suggesting that CSBC-generated data and tools are being generalized for use beyond cancer. (Note that the **Additional Figure 4** includes 2020 data, though awards have not yet been made. For applications submitted in 2017-2019, the NCI success rate was 15%.)

c. How long does it take to access and use the data or tools?

Additional Figure 2C provides journal publication and dataset release dates for the six GEO datasets associated with highly cited CSBC resource publications. All datasets were publicly released by their publication date, with five of the six released one to five months prior to article publication. Total annual FTP downloads for these datasets, which were published in 2017 and 2018, peaked in 2018 (at 2282 downloads), suggesting that the data were rapidly picked up by the scientific community.

d. Are there mechanisms for users to provide feedback on the utility of the data/tools?

At this point, there is no such consortium-level mechanism. The Cancer Complexity Knowledge Portal provides an opportunity to facilitate this type of feedback in the future, and also to better capture many aspects of resource access and usage. Launched in late spring 2020, this portal aims to provide an inventory of CSBC and PS-ON data sets, tools, publications, and other resources that is searchable by program, cancer research theme, and keyword.

6. Please provide up to 5 examples of tools or data have been translated or is on the way to being translated into some kind of clinical application beyond a fundamental understanding of cancer.

It is, of course, the ultimate goal of all basic cancer biology research to discover mechanisms and insight that impact patient care and disease outcome. While the focus of the CSBC is a fundamental understanding of cancer, some of the CSBC awards have produced results that have been translated into the clinic. The following five examples were taken from the CSBC Research and Highlights review provided with the original evaluation report.

- In many tumors genetic heterogeneity does not necessarily predict phenotypic heterogeneity. How a multitude of sometimes disparate somatic mutations across patients manifest in a rather limited number of disease processes or phenotypes remains an open question in cancer biology. The City of Hope U54 team addressed this question through an integrated analysis of longitudinal whole exome and single-cell RNA sequencing across multiple timepoints in breast cancer patients demonstrating that drug-resistant subclones converge on well-known tumor hallmark pathways that are amenable to further therapeutic targeting [8]. Similar studies with an expanded longitudinal patient cohort have suggested new drug combinations for triple-negative and ER+ breast cancers [9] and launched a recently approved clinical trial at the University of Utah that will test the combination of HDAC and CDK4/6 inhibitors.
- Treatment with traditional or targeted therapies can uncover previously unappreciated cell plasticity and the OHSU Measuring, Modeling, and Controlling Heterogeneity Center found that genetically similar TNBCs exhibit significant transcriptional plasticity upon exposure to therapies targeting PI3K or MEK [10]. The plasticity is dynamic, steerable, reversible, and predictable via a mathematical model that describes cell state-switching [10-12]. Current work in the OHSU U54 Center is focused on understanding the epigenetic mechanisms underlying drug-induced plasticity across breast cancer sub-types. The computational methods developed within the OHSU Center, and supported by the CSBC, are being employed in the Serial Measurements of Molecular and Architectural Responses to Therapy (SMMART) program at OHSU to directly guide treatment of TNBC patients.
- Understanding which cells within a heterogeneous tumor population facilitate resistance is a focus of multiple CSBC grants. Coupling single-cell transcriptomics with an exquisitely sensitive dynamic cell mass measurement [13] allowed the MIT U54 Research Center to link transcriptional differences directly to drug response on a single-cell level in liquid and solid tumors [13, 14]. The approach facilitated identification of single drug resistant cells, and the potential molecular mechanisms of resistance. The coupled transcriptome-phenotype measurement is currently being tested in the clinic to identify cells that underlie minimal residual disease in leukemia patients [15] and has moved towards more widespread clinical use via Travera.
- Currently, patients are matched with targeted therapies using sequencing panels to identify common mutations that confer sensitivity to drug. While this approach has resulted in great success for a small number of targets, it does not account for more complex regulatory structures such as feedback mechanisms and alternative signaling pathways

that mediate resistance. Another approach is to infer and target small clusters of genes that regulate multiple orthogonal networks [16]. The Center for Cancer Systems Therapeutics at Columbia University refers to these clusters as "tumor checkpoints" [17] and has employed a suite of publicly available computational tools for identifying and targeting them across tumor types [10, 18-21]. The computational tools developed at the Columbia U54 Center have been licensed to Darwin Health to facilitate adoption by the pharmaceutical industry.

- In addition to contributing to immune cell evasion, single-cell heterogeneity drives local invasion of primary tumors and is prognostic for disease outcome. Employing a cell state model [22] and a novel 3D *in vitro* culture system that allows for purification and analysis of specific LUAD cell subsets the Yale U54 Research Center found that "leader" cells, those that pioneer collective invasion into the ECM, are more metabolically active than "follower" cells [23], do not require an epithelial-to-mesenchymal transition (EMT) for efficient invasion [24], and can be identified in LUAD patients using a mutational signature derived from genes on chromosome 16 [25, 26]. In complementary studies in glioblastoma, a micropatterned device (RACE assay) was utilized to quantify the proportion of migratory to proliferative cells upon tumor resection. Patients whose tumors had high proportions of RACE assay-identified migratory cells exhibited a short time to disease recurrence [27].

7. How transferrable is the information gained from CSBC across cancer types?

The question of how transferrable the information gained within one cancer context is important to variety of cancer research questions, such as the development of targeted therapies, and is one that the field struggles to answer. In contrast to large data collection efforts, such as the TCGA and ICGC, projects within the CSBC portfolio are investigator-initiated and hypothesis driven with data collection completed in a targeted fashion (versus in a survey fashion). Therefore, the specific data (or information) gathered within a particular project may not be immediately transferrable to another cancer context. However, the fundamental concepts that are discovered may be transferrable, as evidenced by citation of CSBC work across disciplines and NIH Institutes (**Figure 6**). For example:

- The OHSU U54 Research Center is investigating cellular plasticity in response to targeted therapies in triple negative breast cancer (bullet #2 above). Published results point to an epigenetic mechanism that drives cells into specific cell states. In more recent work, the OHSU U54 Research Center used the principles learned from the breast cancer study to investigate drug response of pancreatic cancer cells finding a similar epigenetic mechanism.
- The Stanford U54 Research Center has identified education of the immune system via tumor cell spread to local lymph nodes as a general mechanism that facilitates increased distant metastasis in models of head and neck carcinoma, melanoma, and breast cancer. One general principle that has applied to all three contexts is the ability of early disseminating cells to evade natural killer cells. The MSKCC U54 Research Center found that a similar mechanism is important for early dissemination in lung cancer.
- Finally, projects that are mapping how protein-protein interactions are altered by somatic mutation or alternative splicing, such as those within the Cancer Cell Map Initiative U54 Center and the DFCI U01 Project, will produce data amenable to querying across tumor types.

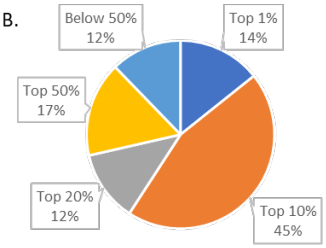
We encourage you to also bring this question to the CSBC investigators during your interactions.

Additional Figures 1-4:

A.

Bibliometric Indicator	Value
Number of articles published in 2017-2018	49
Number of citations (Times cited)	1140
Mean citation	23
% of articles in top 10% for citations	59%

B.

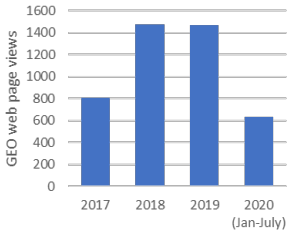


Additional figure 1: Impact analysis of CSBC publications associated with the release of shared data and tools. Bibliometric analysis of publications funded by CSBC, associated with the release shared data and resources, and published in 2017 and 2018. (Per standard bibliometrics practice, only articles at least two years old are used to ensure indicator stability.) Web of Science (WOS) was used to complete analysis. **(A)** The table lists common bibliometric indicators for relevant CSBC publications. The "% of articles in the top 10% for citations" indicator indicates that 58% of the 49 publications have citation counts that rank in the top 10% of all articles published in the same discipline and during the same year. **(B)** Distribution of articles among five percentile rank classes. Each article in this dataset is assigned to a percentile rank class based on its citation count in comparison with the citation counts of all articles published in the same year and Essential Science Indicators (ESI) subject category.

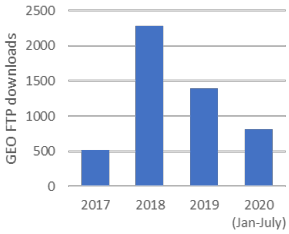
A.

Accession	Title	Journal publication	Dataset release	FTP downloads
GSE93702	Metabolic origins of spatial organization in the tumor microenvironment	Mar 2017	Feb 2017	595
GSE89309	Chromatin states of tumor-specific T cells	May 2017	May 2017	927
GSE92721	Transdifferentiation as a mechanism of treatment resistance in a mouse model of castration-resistant prostate cancer	Jul 2017	Apr 2017	646
GSE102698	Colonic single-cell RNA-seq	Jan 2018	Nov 2017	789
GSE101209	Gene expression changes associated with high density collagen microenviroment in cancer cells	Apr 2018	Nov 2017	920
GSE114397	Single cell RNA-seq and ATAC-seq of EMT induced by TGFbeta stimulation and Zeb1 overexpression	Jul 2018	May 2018	1128

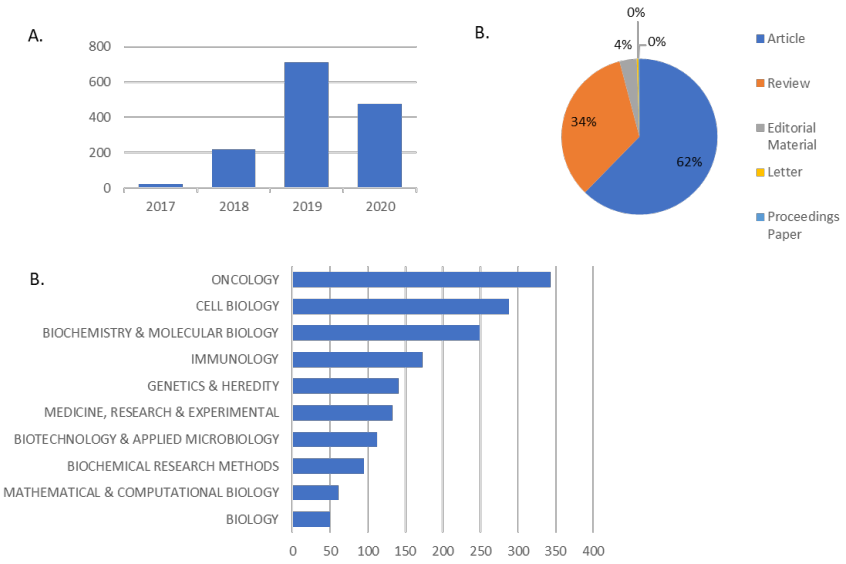
B.



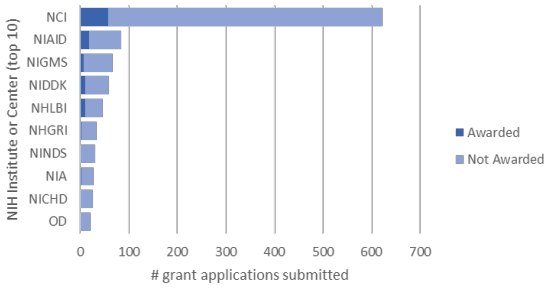
C.



Additional Figure 2: Access and usage of highly-cited GEO datasets
Access statistics for GEO datasets within the top 10 CSBC resource-associated publications, based on citations. **(A)** Basic publication and access statistics, including FTP downloads. **(B)** Total web page views per year for six datasets. **(C)** Total FTP downloads per year of six datasets.



Additional Figure 3: Articles that cite CSBC resource-associated publications
Analysis considered all publications (2017-2020) that cite CSBC articles associated with the release of data or tools (published in 2017-2019, for consistency with report). Analysis was performed in Web of Science. (A) Number of citing articles. (B) Type of articles that cite CSBC resource-associated publications. (D) Top 10 Web of Science Subject Categories for citing articles.



Additional Figure 4: NIH grant applications that cite CSBC resource-associated publications
Citation analysis within NIH research grant applications (submitted 2017-2020) was performed to provide insight on how CSBC-generated resources are being used by the research community. Citation analysis was performed on articles that acknowledge CSBC funding and are associated with the public release of data or tools, and that were published in 2017-2019. Distribution of citing applications across NIH Institutes and Centers (top 10 by citation count shown).

REPORT REFERENCES

1. Bornmann L, Leydesdorff L, Mutz R: **The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits**. *J Informetr* 2013, **7**(1):158-165.
2. Bornmann L, Marx W: **How good is research really? - Measuring the citation impact of publications with percentiles increases correct assessments and fair comparisons**. *Embo Rep* 2013, **14**(3):226-230.
3. Waltman L, Schreiber M: **On the calculation of percentile-based bibliometric indicators**. *J Am Soc Inf Sci Tec* 2013, **64**(2):372-379.
4. Snyder H, Bonzi S: **Patterns of self-citation across disciplines (1980-1989)**. *J Inform Sci* 1998, **24**(6):431-435.
5. Boone PG, Rochelle LK, Ginzel JD, Lubkov V, Roberts WL, Nicholls PJ, Bock C, Flowers ML, von Furstenberg RJ, Stripp BR *et al*: **A cancer rainbow mouse for visualizing the functional genomics of oncogenic clonal expansion**. *Nat Commun* 2019, **10**.
6. Xavier JB, Young VB, Skufca J, Ginty F, Testerman T, Pearson AT, Macklin P, Mitchell A, Shmulevich I, Xie L *et al*: **The Cancer Microbiome: Distinguishing Direct and Indirect Effects Requires a Systemic View**. *Trends Cancer* 2020, **6**(3):192-204.
7. Suhail Y, Cain MP, Vanaja K, Kurywchak PA, Levchenko A, Kalluri R, Kshitiz: **Systems Biology of Cancer Metastasis**. *Cell Syst* 2019, **9**(2):109-127.
8. Brady SW, McQuerry JA, Qiao Y, Piccolo SR, Shrestha G, Jenkins DF, Layer RM, Pedersen BS, Miller RH, Esch A *et al*: **Combating subclonal evolution of resistant cancer phenotypes**. *Nat Commun* 2017, **8**(1):1231.
9. Chi F, Liu J, Brady SW, Cosgrove PA, Nath A, McQuerry JA, Majumdar S, Moos PJ, Chang JT, Kahn M *et al*: **A 'one-two punch' therapy strategy to target chemoresistance in ER+ breast cancer**. *bioRxiv* 2020:2020.2003.2012.989251.
10. Risom T, Langer EM, Chapman MP, Rantala J, Fields AJ, Boniface C, Alvarez MJ, Kendsersky ND, Pelz CR, Johnson-Camacho K *et al*: **Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer**. *Nature communications* 2018, **9**(1):3815.
11. Chapman MP, Risom T, Aswani AJ, Langer EM, Sears RC, Tomlin CJ: **Modeling differentiation-state transitions linked to therapeutic escape in triple-negative breast cancer**. *PLoS Comput Biol* 2019, **15**(3):e1006840.
12. Langer EM, Allen-Petersen BL, King SM, Kendsersky ND, Turnidge MA, Kuziel GM, Riggers R, Samatham R, Amery TS, Jacques SL *et al*: **Modeling Tumor Phenotypes In Vitro with Three-Dimensional Bioprinting**. *Cell Rep* 2019, **26**(3):608-623.e606.
13. Cetin AE, Stevens MM, Calistri NL, Fulciniti M, Olcum S, Kimmerling RJ, Munshi NC, Manalis SR: **Determining therapeutic susceptibility in multiple myeloma by single-cell mass accumulation**. *Nature communications* 2017, **8**(1):1613.
14. Kimmerling RJ, Prakadan SM, Gupta AJ, Calistri NL, Stevens MM, Olcum S, Cermak N, Drake RS, Pelton K, De Smet F *et al*: **Linking single-cell measurements of mass, growth rate, and gene expression**. *Genome Biol* 2018, **19**(1):207.
15. Luskin MR, Murakami MA, Manalis SR, Weinstock DM: **Targeting minimal residual disease: a path to cure?** *Nat Rev Cancer* 2018, **18**(4):255-263.
16. Paull EO, Aytes A, Subramaniam P, Giorgi FM, Douglass EF, Chu B, Jones SJ, Zheng S, Verhaak R, Abate-Shen C *et al*: **A Modular Master Regulator Landscape Determines the Impact of Genetic Alterations on the Transcriptional Identity of Cancer Cells**. *bioRxiv* 2019:758268.
17. Califano A, Alvarez MJ: **The recurrent architecture of tumour initiation, progression and drug sensitivity**. *Nat Rev Cancer* 2017, **17**(2):116-130.

18. Alvarez MJ, Subramaniam PS, Tang LH, Grunn A, Aburi M, Rieckhof G, Komissarova EV, Hagan EA, Bodei L, Clemons PA *et al*: **A precision oncology approach to the pharmacological targeting of mechanistic dependencies in neuroendocrine tumors**. *Nat Genet* 2018, **50**(7):979-989.
19. Echeverria GV, Ge Z, Seth S, Zhang X, Jeter-Jones S, Zhou X, Cai S, Tu Y, McCoy A, Peoples M *et al*: **Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state**. *Sci Transl Med* 2019, **11**(488).
20. Sanchez-Martin M, Ambesi-Impiombato A, Qin Y, Herranz D, Bansal M, Girardi T, Paietta E, Tallman MS, Rowe JM, De Keersmaecker K *et al*: **Synergistic antileukemic therapies in NOTCH1-induced T-ALL**. *Proc Natl Acad Sci U S A* 2017, **114**(8):2006-2011.
21. Shen Y, Alvarez MJ, Bisikirska B, Lachmann A, Realubit R, Pampou S, Coku J, Karan C, Califano A: **Systematic, network-based characterization of therapeutic target inhibitors**. *PLoS Comput Biol* 2017, **13**(10):e1005599.
22. Haney S, Konen J, Marcus AI, Bazhenov M: **The complex ecosystem in non small cell lung cancer invasion**. *PLoS Comput Biol* 2018, **14**(5):e1006131.
23. Commander R, Wei C, Sharma A, Mouw JK, Burton LJ, Summerbell E, Mahboubi D, Peterson RJ, Konen J, Zhou W *et al*: **Subpopulation targeting of pyruvate dehydrogenase and GLUT1 decouples metabolic heterogeneity during collective cancer cell invasion**. *Nature communications* 2020, **11**(1):1533.
24. Richardson AM, Havel LS, Koyen AE, Konen JM, Shupe J, Wiles WGt, Martin WD, Grossniklaus HE, Sica G, Gilbert-Ross M *et al*: **Vimentin Is Required for Lung Adenocarcinoma Metastasis via Heterotypic Tumor Cell-Cancer-Associated Fibroblast Interactions during Collective Invasion**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2018, **24**(2):420-432.
25. Zoeller EL, Pedro B, Konen J, Dwivedi B, Rupji M, Sundararaman N, Wang L, Horton JR, Zhong C, Barwick BG *et al*: **Genetic heterogeneity within collective invasion packs drives leader and follower cell phenotypes**. *J Cell Sci* 2019, **132**(19).
26. Pedro B, Rupji M, Dwivedi B, Kowalski J, Konen JM, Owonikoko TK, Ramalingam SS, Vertino PM, Marcus AI: **Prognostic significance of an invasive leader cell-derived mutation cluster on chromosome 16q**. *Cancer* 2020.
27. Smith CL, Kilic O, Schiapparelli P, Guerrero-Cazares H, Kim DH, Sedora-Roman NI, Gupta S, O'Donnell T, Chaichana KL, Rodriguez FJ *et al*: **Migration Phenotype of Brain-Cancer Cells Predicts Patient Outcomes**. *Cell Rep* 2016, **15**(12):2616-2624.