



**BTI Computational
Biology Center**

Decoding the complexity of life

**Strategic Plan
2021-2026**

Table of Contents

Executive Summary	2
Introduction	3
Committees	4
Mission, Vision, and Strategic Goals	5
Strategy and Implementation	6
Timeline	12

Executive Summary

The BTI Computational Biology Center (BCBC) supports and contributes to BTI's strengths in bioinformatics, plant transformation, and metabolomics to tackle new frontiers. The Center promotes synergistic interactions amongst groups at BTI and supports integration with the greater community through research collaborations and educational programming. This strategic plan identifies five areas to address in the next five years to support our mission.

- 1) **Collaboration:** Structure our collaborations and build our network.
- 2) **Research:** Develop and apply computational tools to advance our understanding of life.
- 3) **Education:** Enhance bioinformatics consulting, education, and outreach activities based at BTI.
- 4) **Infrastructure:** Meet current and future computational infrastructure needs
- 5) **Funding:** Attract new and diverse avenues of funding to stimulate research and education efforts.

Introduction

In the spring of 2016, BTI began planning the BCBC – a process set in motion by Lukas Mueller’s concept of a bioinformatics-focused center at BTI. The goal was to build on the Institute’s strength in genomics and associated bioinformatics and raise our profile. A planning group was convened, including Zhangjun Fei, Fay-Wei Li, Lukas Mueller, Eric Richards, Susan Strickler, and Amy Yanosh, which began with a brainstorming session including members of the Cornell community (Ed Buckler, Michael Gore, Chris Myers and Susan McCouch). Later, additional BTI members were recruited for specialized input (IT, Outreach, and Fundraising). The planning group has become a Steering Committee dubbed ‘The Pirates’, which provides feedback on how BCBC can best serve the computational needs of the local community as well as engage with scientists, students, and society outside our immediate locale.

The BCBC launched as a center focused on computational approaches to biological diversity, with particular attention to both genomic and biochemical complexity in non-model plants. The planning group also felt strongly that BTI’s current efforts in bioinformatics training needed to be continued and additional avenues of training implemented. In the planning process, it was discussed how current and future training programs might mesh with related efforts on campus. There is considerable campus demand for bioinformatics training, and we understood that BTI does not have the capacity to fully satisfy this need. Therefore, the BCBC is principally focused on the BTI community and its circles of collaborators.

The 2020-2021 COVID-19 pandemic has resulted in the need for greater bioinformatics support for researchers, as well as new educational opportunities. Many students and researchers have been forced to work remotely and computational biology projects and training can usually be adapted to support this mode of work. The need for virtual communication during the pandemic has also encouraged people to explore new ways of connecting and has likely made them more receptive to learning and working through these digital mechanisms.

Strategic Planning Members

Sophia Darling
Gina Holl
David Stern
Susan Strickler

The Pirates BCBC Steering Committee

Paul Debbie - Director of Research
Zhangjun Fei (2017-2021)
Fay-Wei Li (2017-2020)
Lukas Mueller (2017-2020)
Andrew Nelson (2020-2023)
Joyce Van Eck (2019-2022)

Committees

Strategic Planning Members

Members of the BCBC Strategic Planning group have provided guidance on the direction of BCBC and assisted with the revisions of the strategic plan. Activities include advice on budgeting, instruction on the nuances of strategic planning, and overall constructive feedback.

The Pirates

The steering committee provides feedback and guidance on the direction of the center. Members of the faculty provide feedback and ideas regarding BCBC workshops, training, and meetings and often contribute to these events. Members serve a three-year term except for the Director of Research who is a permanent member. Representatives from the Postgraduate Society (PGS), IT, Education and Outreach, Communication and Development also join meetings yearly or as their expertise is needed. We additionally plan to seek outside advisership to help the center become more outward-looking.

BTI Computational Biology Center (BCBC)

Decoding the complexity of life

Mission, Vision, and Strategic Goals

Mission

The BCBC enables computational approaches by providing educational resources and developing new methods in a collaborative framework to advance understanding of the complexity of life.

Vision

Computational biology democratized through accessible people and resources

Strategic Goals (2021-2026):

- 1) **Collaboration:** Structure our collaborations and build our network
- 2) **Research:** Develop and apply computational tools to advance our understanding of life
- 3) **Education:** Enhance bioinformatics consulting, education, and outreach activities based at BTI
- 4) **Infrastructure:** Meet current and future computational infrastructure needs
- 5) **Funding:** Attract new and diverse sources of funding to stimulate research and education initiatives

Strategy and Implementation

In the following, we detail how we will achieve each of our goals over the next five years.

Goal 1: COLLABORATION: Structure our collaborations and build our network. Through the bioinformatics consulting program, workshops, and other BCBC-led activities, the Center has established a number of relationships in the local community and beyond. Research at the BCBC is driven largely by partnerships built on these relationships which help to guide strategic choices of problems and experimental systems. We will continue to foster these relationships as well as forge new ones.

Goal 1 Strategy 1: Establish tiers of engagement (core, Tier 1, Tier 2, Tier 3) and define roles. Core members include the BCBC Director and the consultants. BTI funding will be allocated primarily to core members for consultants' salaries and communal infrastructure, but we will also seek funding from additional sources (see Goal 5). Tier 1 includes BCBC BTI faculty members and Cornell groups with which we have established a formal collaboration. Tier 1 members will help teach workshops and have an advisory role in the BCBC by helping select seminar speakers and workshop topics. Members from their labs will be selected for guest consulting spots for special themed office hours to occur once a month. Tier 2 includes the remaining BTI faculty, and Cornell and external groups with which we have formal collaborations. Tier 1 and Tier 2 members will have free access to consulting hours, BTI training courses, and server time (BTI has priority). Funding for BCBC contributions to projects led by Tier 1 and Tier 2 members are expected to come mainly from collaborative grant proposals to allow for more effective planning and allocation of BCBC resources although for-fee projects will be considered as resources allow. Tier 3 will include for-fee external collaborations, such as projects with industry or cost recovery from workshop attendees. These collaborators will have no advisory role in the BCBC and will only be taken on as time and resources allow. Tier 3 collaborations may move to Tier 2 in cases where successful funding for a longer-term project is procured.

Tactics

- Define current members and tiers
- Conduct surveys of members and the steering committee to guide selection of topics covered in training efforts (workshops, mini courses, etc.)
- Add new affiliates to office hour email and course advertisements if outside BTI
- Vote on seminar speakers in series
- Install member pipelines on our servers

Metrics

- Have one guest consultant at a themed office every two months
- Achieving balance of resources with demand (quantified through budgeting)

Goal 1 Strategy 2: Engage with Cornell. The vision for the Center positions BTI to benefit from many related activities on campus while building a unique research enterprise that will contribute to the life science community on the Cornell campus. BCBC-affiliated scientists currently encompass faculty at Cornell mainly in SIPS as well as those in the Digital Ag

initiative. We are also connecting with faculty in Computer Science, Statistics, and other areas. Campus activities will include periodic internal research seminars and discussion sessions, as well as hosting external speakers in conjunction with the seminar series in the Computational Biology department at Cornell.

Tactics

- Recruit new members from Cornell
- Incorporate new members into education programs
- Organize themed office hours
- Join Computational Biology Seminar Series to co-host seminar speakers
- Organize a BCBC journal club
- Establish network of community nanopore users

Metrics

- Make at least one new connection with a Cornell project leader per year leading to assistance with training efforts, journal club, or proposal submission
- Participation of at least one Cornell attendee at themed office hour days
- Attendance of at least 20 Cornell attendees per BCBC sponsored seminar
- Attendance of at least 5 Cornell attendees at journal club

Goal 2: RESEARCH. Develop and apply computational tools to advance our understanding of life. BCBC research will be guided by collaborations and consulting projects to spur research that leverages plant diversity to focus on their unique biology.

Goal 2 Strategy 1: Develop software, tools, pipelines, databases, algorithms, and best practices for bioinformatics problems. The Center will focus on creating and improving computational tools that are specialized for plants and associated organisms. Bioinformatics in plants often relies on software and pipelines written for animal genomes, and these do not take into account prevalent plant genome characteristics such as polyploidy and high repeat content. Consequently, these algorithms perform many tasks poorly in plant applications, such as genome assembly, genome mapping, and repeat and gene annotation.

Tactics

- Continue developing software in an open-source format and provide it through GitHub
- Continue utilizing cloud, virtualization, and containers in software development
- Define best practices for data analysis pipelines and encourage their use across all BCBC groups

Metrics

- Develop one tool every 2 years*
- Attract 5 unique visitors to BCBC GitHub page per month*
- Publish 3 peer-reviewed collaborative manuscripts per year*

Goal 2 Strategy 2: Apply tools to solve bioinformatics problems in the framework of the complexity of life. Evolution and adaptation is the ideal framework to support development of our tools as it encompasses many urgent questions relating to the environment, health, and agriculture, all problems that may be addressed with plant focused tools.

Tactics

- Continue development and maintenance of user-friendly databases to store data from research programs
- Support database and tool use with training and outreach activities
- Integrate data from different omics technologies

Metrics

- Attract 5 unique visitors to BCBC GitHub page per month*
- Publish 3 peer-reviewed collaborative manuscripts per year*

Goal 3: EDUCATION: Enhance bioinformatics consulting, education, and outreach activities based at BTI. BCBC will be a hub of activities that extends its reach through in person and virtual bioinformatics education.

Goal 3 Strategy 1: Make tools and resources easily accessible and promote their use. Tools developed through the BCBC research program must be presented in a user-friendly format with proper documentation.

Tactics

- Develop workshops to demonstrate usage of tools and pipelines
- Make presentations from conferences and workshops publicly available
- Continue to update the bioinformatics course page

Metrics

- Attract 5 unique visitors to BCBC GitHub page per month*
- Continue to have top attendance at webinars hosted by CyVerse
- Scoring an average of 90/100 or higher on workshop participant feedback forms*

Goal 3 Strategy 2: Update bioinformatics courses and workshops. By polling topics of interest at BTI Postgraduate Society (PGS) and faculty meetings, we identify areas of interest for further education. Additionally, courses based around usage of tools generated from BCBC research projects will be organized for the broader research community. These courses will be available for a fee to participants outside BTI and Cornell as resources permit.

Tactics

- Supplement current course with additional courses on topics selected by polling PGS and affiliates
- Implement “bring your own data” courses on mainstream topics, such as RNA-seq
- Train users on installed pipelines on a rolling deadline at no cost
- Create a virtual nonmodel plant genomics workshop.
- Create an online bioinformatics course for a fee
- Interface with other resources such as CyVerse, Data Carpentry, and Galaxy

Metrics

- Workshop maximum capacity achieved (usually 20 participants)
- Scoring an average of 90/100 or higher on workshop participant feedback forms*
- Contribution by at least one student in a bioinformatics publication per year

Goal 3 Strategy 3: Structure the consulting program. Internal and affiliate consulting requests are organized into tiers, as follows 1) Q&A sessions, 2) routine tasks for set fees, 3) new projects organized by work packages and fees based on time to complete. Q&A sessions are covered at weekly office hour sessions and acknowledgement of the BCBC will be requested in publications. For new projects, the Director of the BCBC arranges the structure of the project with the client. Arrangements consider issues such as publications, authorship, server time, consultant time, possible future grant proposals, and fees. If the request can be covered by an existing BCBC course, the client is referred to the relevant course for training. External projects requested by nonacademic labs or labs outside of the affiliate structure are only considered as part of a grant proposal or if time allows and incur a higher fee.

Tactics

- Continue weekly consultant meeting
- Maintain bioinformatics pipelines on BTI servers for do-it-yourself efforts
- Continue to evaluate for-fee pricing structure and increase rate as needed

Metrics

- Average at least 2 attendees per office hour each year
- Average at least 10 BTI/Cornell labs at office hour per year
- Achieve at least 60% cost recovery by 2026

Goal 3 Strategy 4: Expand outreach and public engagement efforts. The current REU program may be expanded to attract and support more applicants interested in bioinformatics. The center also supports community-focused incentives to attract the next generation of researchers, such as Girls Who Code and Girl's Day Out (YWCA Cortland). Events are coordinated with BTI Education and Outreach.

Tactics

- Work with BTI Education and Outreach to continue to attract bioinformatics REU students and have a wider reach
- Continue summer intern bioinformatics course
- Continue participation in Girl's Day Out

Metrics

- Reaching participant capacity at events
- Mentor two REU students per year
- At least one BTI news release of and E&O activity in which BCBC participated

Goal 4: INFRASTRUCTURE. Meet current and future computational infrastructure needs.

To facilitate the proposed activities, the Center will require a higher level of computational infrastructure and support compared to what is currently available. While some computational problems can be moved to the cloud, in some cases bandwidth limitations preclude bioinformatics applications from being run in a time- and cost-efficient manner in that environment. The large amounts of data that are produced today, however, need to be vetted and archived and backed up for future reference and analysis. Raw data from projects will be submitted to SRA when applicable and backup in house will not be necessary. Data from genome projects will be submitted to the relevant database. Intermediate data analysis files may be stored in the cloud and on an in-house storage system until submission to an aforementioned database, upon which they will be removed from the local system. The Center

will need to predict space needed for future applications, and also provide a backup capacity of commensurate dimensions. In addition to storage, the Center will need to provide analysis servers. To complement the current server infrastructure, we will plan for upcoming analysis and space needs. We will continue working with BTI IT for assistance with system upgrades and maintenance.

Tactics

- Predict yearly computational needs based on past metrics
- Implement in-house and cloud backup storage
- Encourage researchers to communicate their future needs through the online form
- Implement a job queuing system on analysis servers to ensure servers are used efficiently
- Procure additional server time at the Computational Biology Service Unit (CBSU) for use when all BTI servers are booked
- Utilize additional existing infrastructure including CyVerse, JetStream, and XSEDE

Metrics

- Amount of storage needed each year versus what is provided
- Amount of analysis power needed each year versus what is available

Goal 5: FUNDING. Attract new and diverse avenues of funding to build research, infrastructure, and education efforts. The Center will provide an opportunity for attracting new funds to BTI and will work with the Development and Grants offices to raise support for the Center from federal agencies, foundations and private donors.

Goal 5 Strategy 1: Build cost recovery strategies for consulting projects. Fees are based on an hourly rate. Collaborators are encouraged to include support through grant proposals. Commonly used pipelines are installed on analysis servers and users are encouraged to do their own analyses when possible. BCBC provides free training on these pipelines.

Tactics

- Encourage collaborators to support BCBC staff through grants

Metrics

- Achieve cost recovery of at least 60% by 2026*

Goal 5 Strategy 2: Apply for agency funding, solicit private donors, foundations, and other. Potential targets for funding include the NSF (AI Research Institutes & IOS/EDGE - Enabling Discovery through Genomics Tools), the Templeton Foundation, the Moore Foundation (Data-Driven Discovery), the Sloan Foundation (Data and Computational Research), the Bill and Melinda Gates Foundation (BMGF), Triad Foundation, and special state programs, such as the Southern Tier development programs (Upstate Revitalization Initiative, etc.). Infrastructure support may be obtained through the NSF (Major Research Instrumentation) program.

Tactics

- Apply for 5 grants per year
- Organize proposals with affiliates
- Experiment with crowdfunding at least one project
- Continue working with BTI Development to identify opportunities

Metrics

- Procure funding for 2 collaborative grant proposals per year
- Form connections with one new donor for BCBC

Goal 5 Strategy 3: Build BCBC brand. We aim to make BCBC a recognizable name in the bioinformatics research community. Many of the tactics above will contribute to this goal.

Tactics

- Improve our web presence by advertising events through various avenues such as CyVerse
- Use the BCBC logo on materials
- Hire and maintain staff that are aligned with BCBC branding and can advocate for our mission
- Create a bioinformatics blog

Metrics

- Attract 5 unique visitors to BCBC GitHub page per month*
- Continue to have top attendance at webinars hosted by CyVerse*
- Scoring an average of 90/100 or higher on workshop participant feedback forms*

* denotes a metric used in more than one strategy

Timeline

Tactic	Goal	Target Completion Date	Assignment
Create an online bioinformatics course for fee	Goal 2	1/21	Director, consultants
Solicit feedback from members on training topics	Goal 3	2/21	Director
Procure additional server time at CBSU for use when all BTI servers are booked	Goal 4	3/21	Director
Define best practices for data analysis pipelines and encourage their use across all BCBC groups	Goal 1	9/21	core
Plan BCBC Symposium	Goal 3	5/2022, 5/2024	core
Supplement current course with additional courses on select topics	Goal 2	9/22	Director, consultants
Support database and tool use with training and outreach activities	Goal 1	1/23	core
Incorporate guest consultants from Cornell/elsewhere with a special research themed day	Goal 2	2/23	Director

Tactic	Goal	Target Completion Date	Assignment
Develop workshops to demonstrate usage of tools and pipelines	Goal 2	8/23	Director, consultants
Implement bring your own data courses on mainstream topics, such as RNA-seq	Goal 2	1/24	Director, consultants
Organize a BCBC journal club	Goal 3	1/25	Director
Apply for 5 government grants per year	Goal 5	yearly	core, tier 1
Select speaker for Computational Biology seminar series at Cornell	Goal 3	yearly	Director
Organize proposals with affiliates	Goal 5	yearly	Director
Vote on seminar speakers in series	Goal 3	yearly	core, tier 1
Predict yearly computational needs based on past metrics	Goal 4	yearly	core
Continue summer intern bioinformatics course	Goal 2	summer, yearly	Director, consultants
Continue fall workshop with GOBii or other partners	Goal 2	fall, yearly	core, tier 1, tier 2

Tactic	Goal	Target Completion Date	Assignment
Continue developing software in an open-source format and provide through GitHub	Goal 1	ongoing	core
Continue development of user-friendly databases to store data from research program	Goal 1	ongoing	core
Train users on installed pipelines on a rolling deadline at no cost	Goal 2	ongoing	consultants
Continue weekly consultant meeting	Goal 2	ongoing	Director, consultants
Continue tracking consulting requests	Goal 2	ongoing	Director, consultants
Install member pipelines on our servers	Goal 3	ongoing	consultants
Continue working with development to reach donors	Goal 5	ongoing	Director, development
Usage of BCBC logo on materials	Goal 5	ongoing	Director, development
Maintain bioinformatics pipeline on BTI servers for do-it-yourself efforts	Goal 2	ongoing	consultants