# IBEST

## 2017 Annual Report

University *of* Idaho

The Institute for Bioinformatics and Evolutionary Studies

Supported by the National Institutes of Health under Award Number P30 GM103324


Table of Contents

I. Executive Summary.....1

II. Mission & Vision.....2

III. Goals & Key Performance Indicators.....3

VI. Broadening Interdisciplinary Research in Evolutionary Science.....4

V. Computational Resources Core.....7

VI. Genomics Resources Core.....17

VII. Optical Imaging Core.....33

VIII. Education and Training.....39

IX. Transition to Stable Business Plan Post COBRE.....41

Appendix I. IBEST Publications.....43

Appendix II. IBEST Proposals.....53

Appendix III. IBEST Awards.....54

Appendix IV. BCB Seminar Series.....55

Appendix V. Business for Scientists.....57

Appendix VI. Inland NW Genome Research Symposium.....58

# I. Executive Summary

This report documents the accomplishments of the Institute for Bioinformatics and Evolutionary Studies (IBEST) as a Level III entity at the University of Idaho for the calendar year 2017. IBEST was established as a Level III Research Institute at the University of Idaho in August of 2011, for a period of 5 years. During 2017, we applied for and received renewal of our charter for a period of three years.

Vision: IBEST aspires to be a globally-recognized center of excellence in interdisciplinary studies of evolutionary bioinformatics.

Mission: IBEST's mission is to empower University of Idaho researchers to understand and apply the evolutionary process. We provide an intellectually stimulating and collaborative environment as well as access to state-of-the-art computational and genomic core facilities with exceptional scientific staff.

There are four goals of IBEST: 1) broaden interdisciplinary research in evolutionary processes, 2) maintain and enhance research infrastructure (IBEST's core facilities in Genomics and Computational Resources), 3) promote interdisciplinary education and research training at all career stages (undergraduates, graduate students, postdoctoral fellows, and faculty), and 4) transition to a stable business plan post COBRE. Assessment metrics related to these goals are summarized here and detailed in this report.

1) IBEST faculty report having published 71 papers in 2016 and 72 so far in 2017.
   Faculty submitted 18 proposals through IBEST, requesting $7.59 M in funding.
   Faculty have secured 34 awards (including BEACON) totaling $3.13 M over the reporting period (with notification of additional $1.59 M to be awarded imminently).
   Forty-four faculty from 16 departments and 7 colleges have benefited from interactions with IBEST.

2) The CRC supported 70 users from 20 departments in 6 colleges with UI and 17 external users.
   The GRC supported 154 users from 54 institutions. Within UI, there have been 50 users from across campus.

3) Faculty report engaging 59 undergraduates, 66 graduate students and 24 postdoctoral fellows in IBEST related research (See Appendix I for scholarly output).

4) IBEST is continuing its transition to a post-COBRE funding environment and has initiated efforts to pursue an NSF Science and Technology Center Award.
   Awards have been funded by 4 federal agencies, 3 universities (as sub-awards), one state agency, and 3 non-government organizations.

# II. IBEST Mission and Vision

## Vision Statement

IBEST aspires to be a globally-recognized center of excellence in interdisciplinary studies of evolutionary science and bioinformatics and a leader in applying evolution to large problems facing society.

## Strategic Statement

In five years, we will have extended our excellence in evolutionary science to a broadening array of disciplines at the University of Idaho, including agricultural genomics, soil microbiomes, and wildlife/conservation genomics, and pathogen/host interactions.

## An Enhanced Focus

We will pursue new and emerging partnerships and apply genomic and bioinformatics approaches to accelerate the understanding of the ecological and evolutionary drivers of human health and welfare.

## Alignment with UI Strategic Plan

The goals and activities described in this document directly address Goal 1 of the UI strategic plan. Specifically, we will address Goal 1 Objective A, B, and C, as follows:

> We will continue to build collaborative partnerships that increase productivity of interdisciplinary research of IBEST faculty. This is precisely why we started IBEST.
>
> We will provide world-class research infrastructure (including equipment and human resources) in genomics and high-performance computing to attract and educate excellent graduate student and postdoctoral researchers. This is precisely why we founded the graduate program in Bioinformatics and Computational Biology and we will excel at enabling student-published research.
>
> We will grow UI's reputation by increasing high-impact, interdisciplinary publications, provide support for faculty to pursue large interdisciplinary grants. UI is a nationally recognized (by Elsevier) leader in Evolution, Ecology and Systematics and IBEST activities are central to this excellence.

# III. IBEST Goals and Key Performance Indicators

IBEST was founded on the premise that the best way to maximize the joy of discovery is through collaborative research. In the next five years, we will broaden the impact of evolutionary science across the University of Idaho.

**Table 1**. New goals and key assessment metrics established in the new charter.

| Goal | Assessment Metrics |
|---|---|
| 1) Broadening Interdisciplinary Research in Evolutionary Science. | Number of Publications, Including Interdisciplinary Publications.<br>Number of Extramural Grant Proposals Submitted and Awarded and Diversity of Funding Sources.<br>Number of Departments/Colleges Whose Faculty Have Benefited from IBEST. |
| 2) Maintaining and Enhancing Research Infrastructure. | Number of Users of Each Core Facility.<br>Number of Research Infrastructure/Instrumentation Grant Proposals Submitted/Awarded. |
| 3) Promoting Education and Research Training at Several Career Stages. | Number of Undergraduates, Graduate Students, and Postdocs Engaged in IBEST Research.<br>Scholarly Output of Those Students. |
| 4) Transition to Stable Business Plan Post-COBRE. | Number of IBEST Staff Supported by Permanent Funds.<br>Number of Programmatic Grant Proposal Submitted.<br>Diversity of Funding Agencies. |

# IV. Broadening Interdisciplinary Research in Evolutionary Science

The process of evolution, descent with modification leading to diversification, is the unifying principle of life sciences. Furthermore, understanding evolution is central to improving human well-being because evolutionary processes drive critical health challenges such as emerging infectious diseases, antimicrobial resistance, and even the origin and treatment of diseases such as cancer, mental illness, and obesity. Evolution also underlies agricultural challenges such as the emergence of pesticide resistance, the effects of invasive species, and improving the effectiveness of domestication. Furthermore, understanding evolution helps protect our natural heritage by informing conservation policy and providing insight into adaptation to climate change and how and why organisms respond to changing environments. Thus, IBEST coalesces scholars from the life and physical sciences, agriculture, natural resources, engineering, and social sciences, to leverage advances in fields as diverse as biology, physics, geology, mathematics, statistics, computer science, electrical engineering, husbandry, cultivation, forestry, conservation, geography, languages, and history.

In FY 2017, IBEST was responsible for $2.614 M in research expenditures, continuing our long history as a major component of the UI research endeavor, with a total > $25 M since 2011.

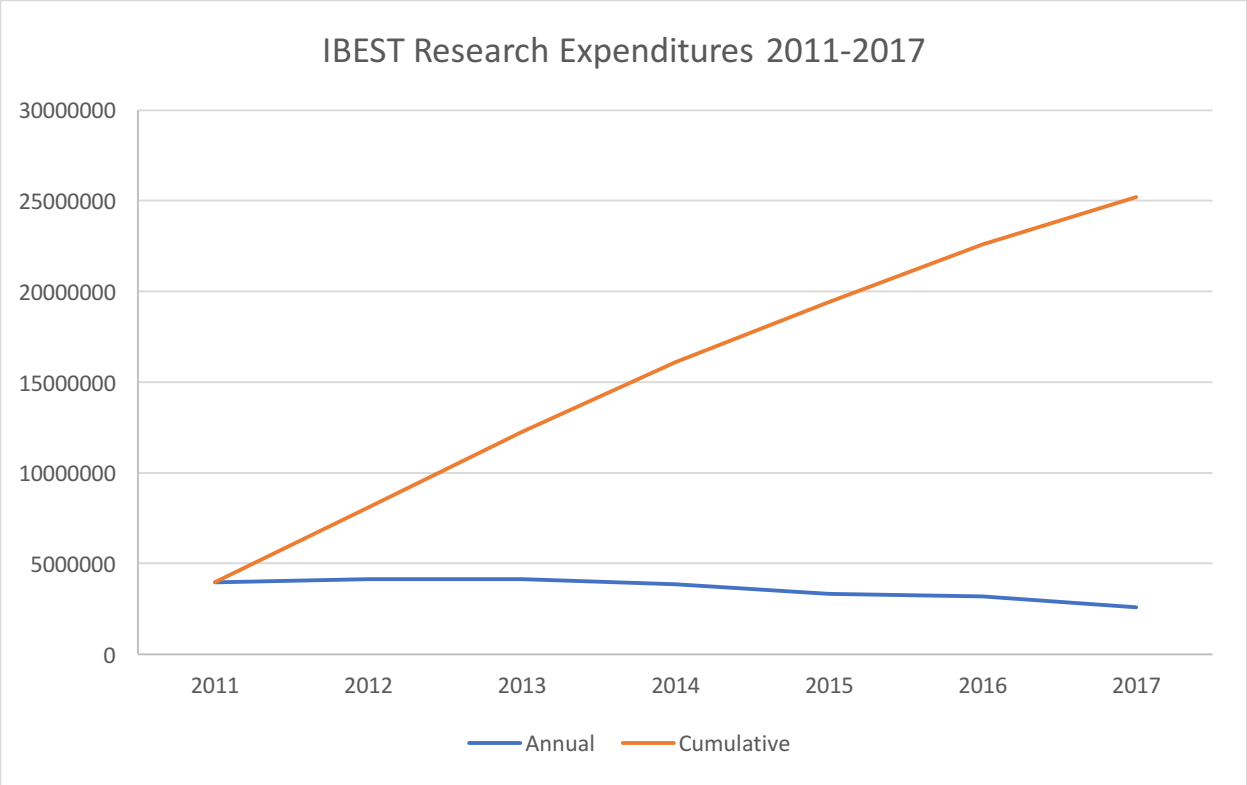


**Figure 1.** IBEST Research Expenditures since becoming a Level III institute.

Publications are the primary means of dissemination of science and are therefore one of our Assessment Metrics. IBEST faculty have been enormously productive; the list of IBEST publications from 2016 (71 publications) and 2017 (72 publications to date) are presented in Appendix 1. Indeed, over the past six years (2012-2017), IBEST faculty have published 426 peer-reviewed journal articles and these have appeared in a large diversity of high-impact journals. Indeed, the area of "Ecology, Evolution, Behavior and Systematics" has been uniquely identified as nationally prominent in a study of UI research competencies conducted by Elsevier as part of the Strategic Plan. Furthermore, Elsevier recognized the central role that IBEST has played in accelerating this impressive publication record. This recognition by an independent third party reiterates past evaluations from our External Advisory Committee (EAC) that have emphasized our national prominence in these areas.

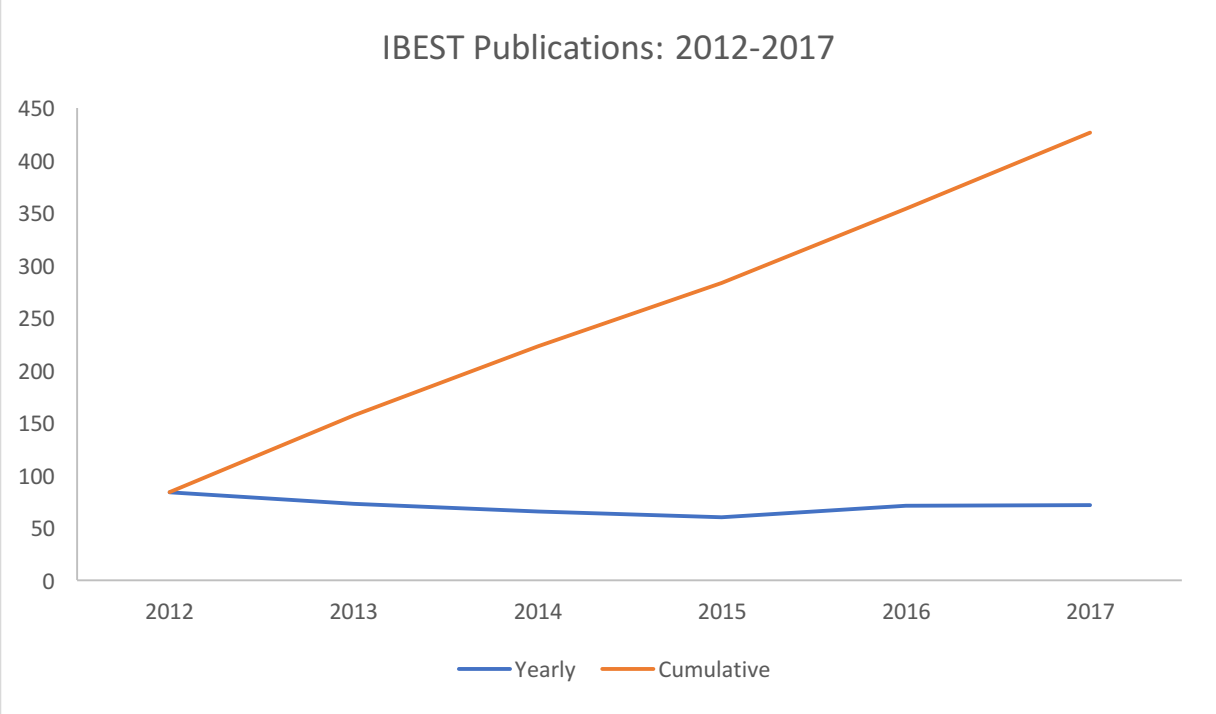


**Figure 2.** IBEST publications since becoming a Level III institute.

Research grant proposals that address evolutionary science may be submitted through IBEST by UI faculty from any unit. In such cases, the IBEST Business Manager assists principal investigators to prepare and submit their grant applications to the UI Office of Sponsored Programs (OSP) who in turn review and submit the application to the granting agency. The level of support provided to investigators varies depending on their level of experience and the agency requirements. At a minimum, the Business Manager works with the PI to prepare the budget and budget justification in accordance with UI policies and those of the granting agency. Once a grant is awarded the IBEST administrative staff help the PI recruit personnel, handle all purchasing and travel expenditures, and help the PI manage their budget. In FY2017, 18 grant applications requesting $7.6 Million in

extramural funds were submitted (Appendix II) and a total of $2.75 Million were awarded (Appendix III). In addition, in FY2017 $378,332 in new funds (direct cost) were awarded from BEACON grants. Furthermore, IBEST faculty have received notification from federal funding agencies that an additional $1.594 million is slated to be awarded based on proposal submitted in the last six months.

Forty-four faculty from 16 departments and 7 colleges have benefited from interactions with IBEST. These are summarized in Table 2.

**Table 2.** Colleges and Departments Impacted by IBEST.

| College | Department |
|---|---|
| COS | Biological Sci |
| COS | Math |
| COS | Statistics |
| COS | Physics |
| COS | Chemistry |
| Art & Architecture | Virtual Design |
| COBE | Accounting |
| COBE | Business |
| CNR | Fish & Wildlife |
| CALS | ID AG Exp Station |
| CALS | Animal/Vet Sci |
| CALS | Plant Soil & Entomology |
| CLASS | Politics & Philosophy |
| CLASS | English |
| ENGINEERING | Computer Sci |
| ENGINEERING | Electrical & Comp Engineering |

# V. IBEST Computational Resources Core

## Mission and Vision

The *mission* of the CRC is to provide state of the art computing and data management services to our customers. Our *vision* is to remain technologically current in hardware, software and services while partnering with customers to help them perform and disseminate their research, in a fiscally sustainable way. Our guiding principles are to maximize the reliability, availability, and effectiveness of our services while minimizing administrative costs.

## Infrastructure

The CRC contains an advanced mix of high performance computing clusters, powerful servers and reliable data storage components and is staffed by personnel with the knowledge and technical skills required to compress years of analysis into days. Our data center is a 1400 square foot facility in Room 124 in McClure Hall on the University of Idaho campus that has been specifically designed and renovated for our core. This room has a dedicated Uninterruptable Power Supply (UPS) with three-phase power and four-forced air handlers attached to redundant university chilled water systems. Optical fiber and copper interconnects provide high-speed data transfer for server and storage intercommunication and communication to the University backbone that is connected to the high-speed Internet 2 network. The features of our primary systems are described below.

*High Performance Computing*

CRC has one main compute cluster for research and genomic data analyses. We have continued to add computational capacity in 2017, and the main cluster now provides 1496 processor cores (up from 1160 processor cores last year) and over 6 terabytes of system memory. Along with the new capacity, we have transitioned fully to new cluster scheduling software that is better able to allocate jobs to our now more heterogeneous cluster. We added two new GPU nodes to make GPU accelerated computing accessible to all our users, enabling research involving machine learning. Cluster nodes are connected with 40Gb/s QDR Infiniband connections, providing fast, low latency data transmission for increased performance of HPC bioinformatics applications. We also added another ‘standalone’ server, for a total of twelve servers (496 total cores and over 4 terabytes total system memory) for applications that require large amounts of memory on a single system but do not take advantage of the parallel cluster resources. Three of our most powerful servers in this group contain 256 times the system memory of a standard desktop (1TB or 1024GB) and are used primarily for sequence assembly of next-generation sequencing data.

*Data Storage*

The CRC has maintained two tiers of primary storage, and this year introduced a third distributed file system. The first tier is comprised of fast but more expensive disk arrays, which we have expanded from 52TB to 130TB. The second tier is comprised of 16 servers with over 600TB (214TB formatted) of slower hard drives and provides primary

storage for large data sets. The new third tier of storage has 15 servers and over 400TBs of hard drives, with another 10 servers and 720 TB hard drives purchased. This third tier will provide greater performance, and has better self-balancing and healing properties. Additionally, we have approximately 300TB disk available for data archiving and backup storage within the McClure data center. In 2016 we added a robust offsite backup system by working with UI ITS to co-locate two IBEST data servers with 360TB combined capacity in the UI Library datacenter, and have added another server with 270TB disk this year. In addition the core provides in-house developed solutions to maintain data integrity and restoration.

*Support Systems*

The CRC maintains its own support infrastructure because this scale of core operations falls well outside that of the University of Idaho Information Technology and Enterprise Computing services. Our support infrastructure includes several servers for data storage and authentication of user accounts, domain name resolution, Internet address assignment, and secure connections to our private networks. The core also provides web and database services for online documentation and data sharing.

*Education and Training*

To support educational programs and inter-institutional collaborations we maintain three teleconferencing enabled conference rooms and a state of the art technology classroom. The classroom is frequently used by instructors from the College of Science and the College of Natural Resources. The classroom also has teleconferencing system, which allows us to offer workshops and classes from and to collaborating institutions such as Michigan State University, University of Texas at Austin, University of Washington, and North Carolina Ag and Tech.

*Power*

Providing the energy demands of the CRC systems is a challenging task. The energy needs to be clean and uninterrupted for proper operation of the systems and supporting infrastructure. This challenge is met by our 3-phase 80KV power supply battery backup system. This system was purchased in 2012 and we replaced the batteries in the fourth quarter of 2015.

New Infrastructure

To increase the capacity, throughput, and reliability within the CRC for our users, we have over the past year:

- Added 6 computational nodes - 320 processor cores and 10,752 GPU cores to our main cluster, and transitioned fully to the new scheduling software introduced last year.
- Replaced 8 failed nodes with more modern refurbished nodes.
- Added one new standalone server, with two Nvidia P100 coprocessor cards.
- Continued to improve our in-house developed billing system.
- Doubled the capacity of our Lustre (1$^{st}$ tier) file system.

- Increased the capacity of the Gluster (2nd tier) file system by 45%.
- Introduced a new Ceph distributed file system, which will provide increased performance and fault tolerance, reducing down time.
- Implemented a multi-factor authentication system for systems administrators, and an intrusion detection system for increased security.
- Installed a new network attached storage (NAS) server for a researcher in Physics.

Planned infrastructure

- We intend to continue adding additional data storage capacity to accommodate increasing numbers of users and their increasing data requirements.
- Add new cluster nodes with additional co-processors and/or GPUs for increased computational efficiency.
- We plan to retire our current virtual machine infrastructure, and replace it with servers running Docker – a containerization technology.

Under Consideration

We are considering various other changes to our infrastructure, including the following:

- New support for users of the Optical Imaging Core. These users generate exceptionally large datasets and need a reliable and cost-effective means of long-term storage.
- We are also considering expanding our account offerings to undergraduate students in a variety of departments.

## Innovation

Continuing Innovation in Technology and Services

The primary function of the CRC is to facilitate the innovation of our customers. We have deployed existing technology in innovative ways, offer services that are not available from most other computational core facilities, and developed unique in-house solutions to address user needs.

Examples of our innovative use of existing technology include:

- We use configuration management systems (the modules environment) to provide customized software services, including versioning. Most cores provide only one version of software, which makes it difficult to replicate prior work or to test new user-developed software. This mechanism is uniform across over 100 systems, so the learning curve for users is very shallow. This mechanism also makes it possible for us to install and test new software without disrupting system availability.
- Some of our hardware, such as the very large memory servers, are not commonly available. These enable users to pursue specialized applications such as alignments of very large genomic datasets, intense agent-based

simulations, and visualization rendering.
- Our existing data backup system was developed in house.
- Internal software development – We employ several technologies and write a significant amount of code to maintain our complex infrastructure with a small staff.

Examples of our innovative services include:
- The tight integration of the CRC and GRC in terms of personnel, hardware, software, and administration is highly innovative relative to most other computational core facilities.
- We provide a high level of support for customized software installation, configuration, script development, and *ad hoc* user services.
- We offer a local, secure file-sharing system as an alternative to DropBox and similar cloud storage services.
- We offer our own web-based account management, poster printing, and online documentation systems. These systems were developed in-house, and offer streamlined interfaces to our services and documentation and are easier for CRC staff to maintain. Thorough documentation of our services allows novice users a consistent reference, and reduces CRC staff user support load.

## Sustainability

To sustain the level of service required by investigators we must continually update hardware and software to remain an attractive option for researchers. There are two dimensions to sustainability in the CRC: maintaining our current services and updating services to remain on the cutting edge.

### Maintaining Current Status

In June 2014, we implemented a fee for service model with a single user fee for access to all systems. A single standard user subscription currently costs $1500 per year, down from $2000. We have been able to control costs through the extensive use of automation, commodity hardware, university support, and a stable user base. We currently have 70 paid users (42 standard accounts, and 28 satellite accounts) across 20 UI departments, and 6 colleges.

The bulk of the CRC equipment was purchased using COBRE funds, and so pursuant to federal guidelines, user fees fund personnel costs associated with administering the CRC, not hardware. In 2011 however, the CRC used University funds to lease several servers with the intent of billing users for the computational time used. Last year we made the final payment to the University for this leased equipment, and have incorporated the leased servers into various CRC systems. We are now in a position to support our staff on user fees (40 standard accounts, and 30 satellite accounts) provided that: 1) the significant salary support from the Dept. of Biological Sciences for the CRC Director position continues, and 2) we continue to use undergraduate students to fill the role of systems administrators.

We introduced a new model this fall to complement our existing fee for service model. The new 'Condominium' model will allow researchers to purchase dedicated computational hardware, which the CRC will manage. This hardware will be incorporated into our main cluster, and unused compute cycles will be available for all CRC users. In this way, researchers are freed from the burden of system administration, and CRC users benefit from additional computational capacity. Along with the condominium model, we introduced a tier of free accounts for undergraduate and graduate students. These accounts have access to a limited set of the CRC's computational resources, but with all the software available. The undergraduate accounts are not time limited, and are intended to help the University achieve its strategic goal of increased research productivity on the way to becoming an R1 institution. The graduate accounts are limited to 90 days, and are intended as an introduction to high performance computing, and to see what the CRC can offer without monetary commitment.

We have continued to increase our campus wide impact and overall number of paid accounts and now have users in 20 different UI Academic Departments (Fig. 3). We continue to support external users from Reed College, and Washington State University, and a private company out of Wisconsin (Fig. 4). We intend to continue to court external users, while keeping our core user-base on the University of Idaho campus.

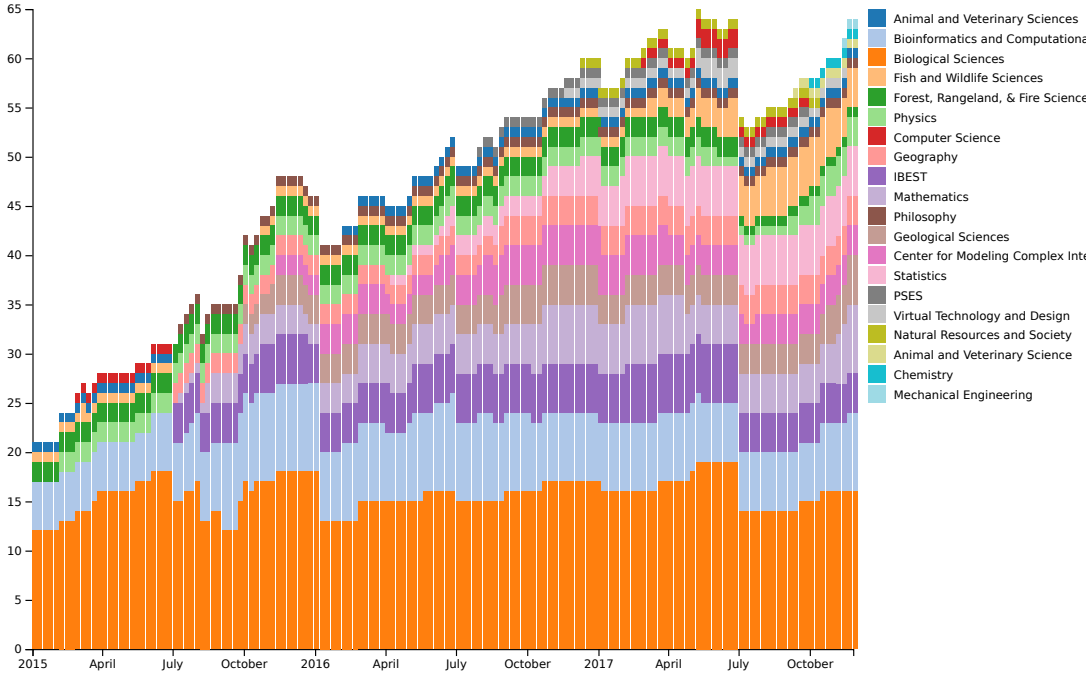


**Figure 3.** The number of University of Idaho users with active accounts for each week since 2015, colored by department.

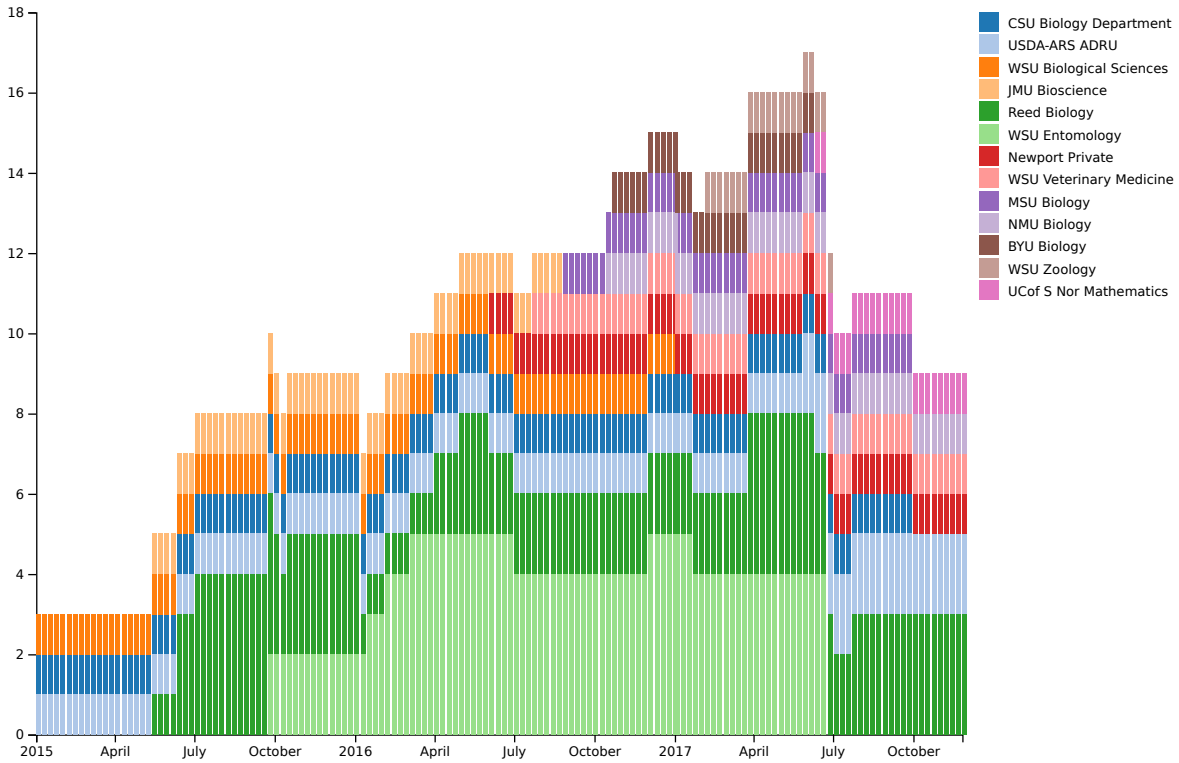


**Figure 4.** The number of external users with active accounts for each week since 2015, colored by institution and department.

Usage pattern trends and prospectus

In the fast-paced and intensely competitive research environment now common to higher education, our users tend to pick the shortest path to quick results rather than spend the time required to learn complex application programming interfaces. Thus, being able to simply log onto a powerful server and immediately run several threads of a bioinformatics application has historically proven more attractive for our users. Over the past year we have seen increased demand on the main cluster. To help CRC users overcome the intimidating knowledge barrier presented by job-scheduling software, we offer regular workshops where researchers can get one-on-one help converting their scripts and application calls to cluster enabled scripts.

This year we have also seen increased demand for GPU enabled computational infrastructure to be used for machine learning and neural network simulation. To help meet this demand, we purchased two new cluster nodes with GPUs, and added another as part of a Condo agreement with an existing user.

Keeping Current

Maintaining current hardware is a continuous challenge. Academic and corporate data centers assume a half-life of about two years for high-end equipment like ours. Thus, after approximately four years, the equipment is fully depreciated. Two of our most powerful systems were purchased 5 years ago (Nov 2012) and the bulk of our cluster

nodes are now 9 years old (purchased Nov 2008). As the older cluster nodes fail, it has not been cost-effective to repair them, but to simply replace them with refurbished nodes. At this time, 8/64 of the original cluster nodes have failed and been replaced with more modern hardware. The purchase of 32 new cluster nodes at the end of 2015 has helped keep our current user base satisfied, but new demands for GPU enabled servers have exceeded our current capacity.

As the data storage needs of our users have grown even faster than their computational needs, we have focused our new equipment purchases on ever-larger data storage capacity (Figure 4). However, as we add additional users, we will now need to continue to update our computational infrastructure as well.

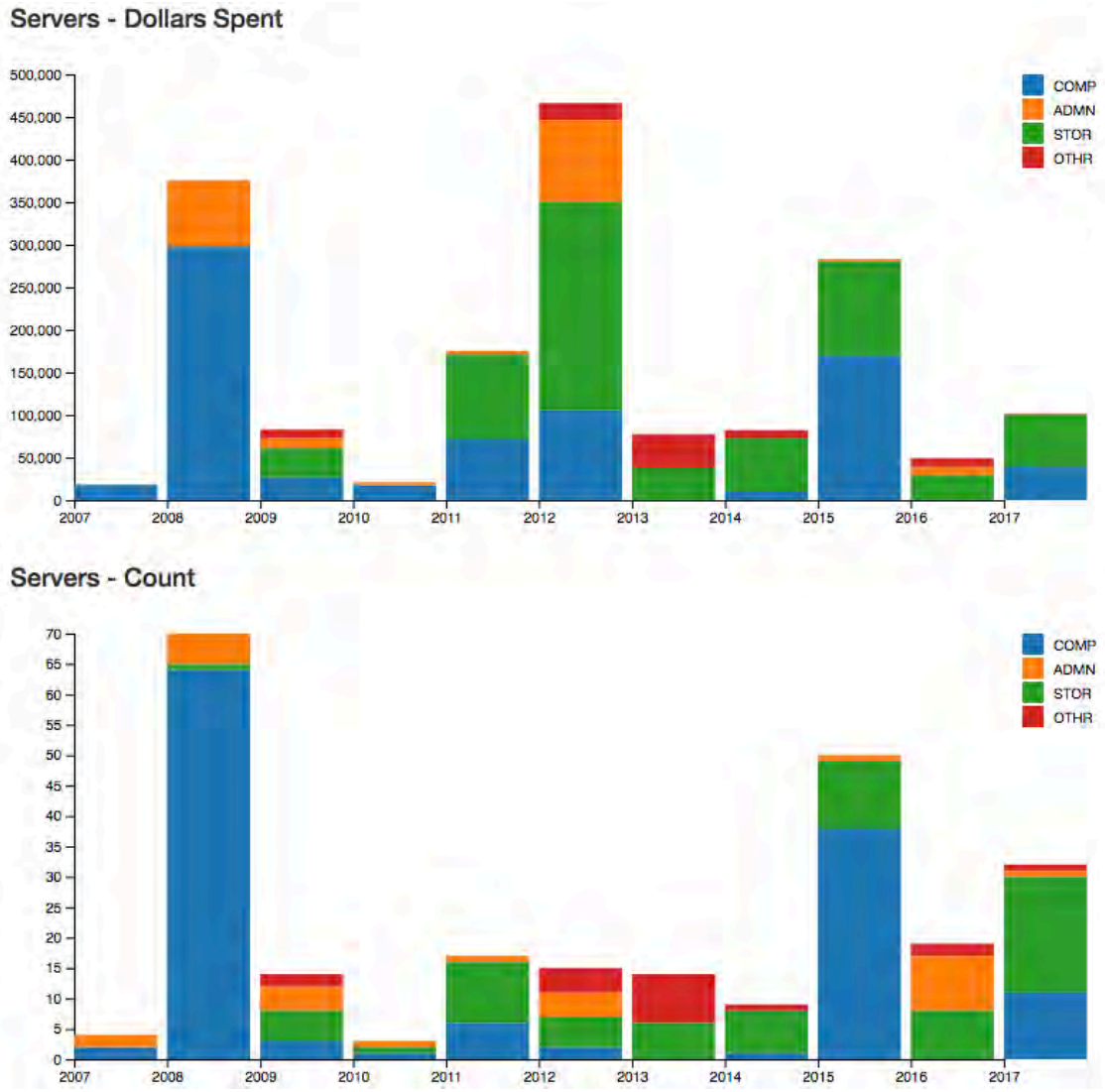


**Figure 5**. The number of servers and total amount spent each year, colored by their primary purpose.

Because our primary user data storage is a distributed file system composed of several individual servers that each store a part of the overall file system, a high-speed network is necessary to ensure adequate performance. As the amount of data stored and accessed by our users has increased, the standard networking technologies employed have struggled to deliver consistent performance. We therefore purchased higher speed network interfaces (Infiniband) to decrease latency and increase throughput by 4000% and installed that equipment in late 2015. Additionally, working with UI technology services, we updated our external facing switches from 1G to 10G so that our servers will be able to connect to the campus 10G network at full speed instead of the final link being only 1G. This increased bandwidth makes an offsite daily backup feasible, and we implemented such as system in 2016.

Plans

The sustainability of the CRC over the long term will require that we maintain self-generated revenue and retain institutional financial support. User fees alone cannot maintain centers such as the CRC given their high capitalization and maintenance costs. Therefore, institutional support will always be part of the core's revenue. Our goal is to support the bulk of salary expenses with self-generated revenue, and rely on future grant funding to continue to replace aging hardware.

In order to avoid costly maintenance contracts with hardware vendors, we rely heavily on commodity hardware with warranties included at the time of purchase. However, we keep hardware in service well beyond typical warranty periods (3-5 years), and so ongoing maintenance costs will likely increase.

Besides relying on commodity hardware, we also rely on open-source software almost exclusively. While this reduces overall operating costs, it places a greater burden on our staff in both time and education. With the historic compensation levels at the UI, we have found it impossible to attract systems administrators with previous experience in HPC and configuration management tools. Thus, we have in the past settled for systems administrators with little experience - but high aptitude - and trained them up over the period of a couple years. However, these now experienced systems administrators have then promptly moved on to private industry or higher education institutions that can compensate them at market rates. We have therefore decided to instead focus on our undergraduate staff. The amount of training necessary to get a talented undergraduate to the level of competency required for systems administration in our environment is comparable to the amount required for an individual with a bachelor's degree. The undergraduate students get the benefit of hands-on systems administration experience, and the CRC benefits from a simplified hiring process and reduced salary expenditures. This type of training requires a hardware environment tolerant of system administrator errors – which our production environment is not. We therefore built a unique 'classroom cluster' which has nearly all the complexities of our production cluster – but fits nicely in the IBEST classroom. This classroom cluster uses modern hardware similar to that used in our production environment, and provides an ideal environment for testing new technologies and software. Solutions developed for the classroom cluster can be

translated directly to our production environment. We intend to offer undergraduate classes for credit on HPC administration using the classroom cluster. These classes will serve as a talent pool for future CRC staff.

## Outreach

Do to staffing constraints, the CRC is less active with outreach than the other cores within IBEST. We provide workshops every Fall semester for CRC users and non-users. We hope to expand our educational reach with the undergraduate HPC administration class. The outreach activities described by the GRC are often facilitated by the tight integration between the Computation and Genomics Resources Cores.

## Opportunities

There are many on campus resources, both current and potential, that could increase the CRC user base or simplify CRC operations.

- The new COBRE Center for Modeling Complex Interactions includes a modeling collaboratorium and several potentially computationally intensive projects which has led to 20 user accounts, and computational nodes for the HPC cluster.
- We are now supporting other research units within the College of Science, including Geographical Sciences, and Geology Researchers. With the required software for these users now installed, we are in a better position to offer services to other faculty members. Additionally, with our free undergraduate accounts – we anticipate being able to impact even more departments on campus.
- Federal grants increasingly require data management plans to be submitted with any large grant. The CRC has the expertise to offer data management services for researchers, but would need to invest in additional hardware and formalize a service center category to offer this service.

We are also considering other opportunities to take advantage of existing on-campus resources. For example:

- We have expanded the use of undergraduate assistants for tasks such as inventory, classroom and communications support, hardware installation, and systems monitoring. In the past, this has been a reliable pipeline for developing and training future CRC staff.
- Work with the College of Business and Economics to help develop and implement a marketing strategy and a formal business plan for the CRC.
- Tap existing users to recruit new customers, for example at IBEST lunches or new faculty orientation.
- Include a university-funded CRC “gift certificate” as part of the startup package for new faculty.

We could also consider expanding our mission to support educational activities such as undergraduate research, courses, and workshops, or to support research from non-

evolutionary scientists such as physicists and computer scientists.

## Challenges

*Data Storage*: As the cost of DNA sequencing has fallen, the amount of data available to researchers from both on campus resources such as the GRC, and from public databases such as NCBI has increased dramatically. This readily available sequence data has found its way to our servers *en masse*, enabling CRC users to study previously intractable evolutionary processes. We are constantly adding storage capacity to accommodate current users. However, as our primary data capacity increases, backups of that data have become increasingly difficult to manage using open source backup solutions. We are currently employing large hard disk arrays with advanced file systems that allow for file compression, and reducing the number of backup copies currently being maintained.

*Networking*: The number of storage servers needed to accommodate user data has increased dramatically (doubled this year), but our internal network capacity is limited to the 10 switches that compose a 'virtual chassis' – and only two of these switches are 10g. To free up capacity on our main switches, we have brought several older switches out of retirement to use for non-critical networks. We will need to add additional switching capacity and run segregated internal networks to maintain our high performance storage systems.

**Figure 6**. *Hemphillia skadei*, a newly discovered Idaho endemic jumping slug described in 2017 (Lucid et al., in press) using IBEST-generated genetic data.

# VI. IBEST Genomics Resources Core

## Mission and Vision

The *mission* of the IBEST Genomics Resources Core (GRC) is to provide researchers at the University of Idaho access to cutting edge genomics technology and the bioinformatics tools needed to acquire, analyze, and visualize data. The vision of the GRC is to stay current in genomics technology and bioinformatics, remaining agile with respect to new techniques and approaches, and to build partnerships with research groups and other regional core facilities.

## Summary of Accomplishments

- Moved to a new lab space in the Integrated Research and Innovation Center (IRIC). This building was opened in January 2017, and is the university's premier interdisciplinary research center. The GRC was one of the first groups to move into the building, and was fully operational for the IRIC Ribbon Cutting event held on January 31st. The new location is more centrally located making the GRC more accessible to groups across campus.
- Tailoring molecular and bioinformatic techniques to accommodate the unique needs of each researcher.
- Custom two-step PCR process for generating amplicons and sequencing up to 2,976 amplicons per MiSeq run or HiSeq lane resulting in drastically reduced cost and increased coverage.
- To allow clients to assemble genomes without a reference the GRC is testing new long-read technologies including developing skills for generating and analyzing both Oxford Nanopore and PacBio long reads.
- The GRC's unique practices and capabilities have resulted in a large number of first-time clients from word-of-mouth advertising as well as an increasing number of repeat clients.
- Developed a software tool for Illumina read cleaning which introduces a number of new methods for cleaning Illumina reads, and facilitates rapid pipeline development with reduced I/O overhead by using Linux pipes (https://github.com/ibest/HTStream).

## Infrastructure and Personnel

The IBEST GRC is the only comprehensive facility on the University of Idaho campus that houses all the equipment and personnel necessary to aid researchers in every aspect of high-throughput genomics research. It provides the molecular expertise and equipment needed for most high-throughput sequencing studies, and develops partnerships with other service facilities when additional capacity or other specialized

equipment are warranted. The real benefit of the IBEST Genomics Resources Core facility, however, has been the integration of bioinformatics data analysis with data generation. The GRC offers consultation on experimental design, appropriate and best use of technologies, and bioinformatics support to perform analysis, quality assurance, interpretation, and visualization. Through a unique strategy known as “the triangle of collaboration,” an investigator, molecular scientist, and bioinformatician meet regularly as a team to discuss the goals and objectives for a project. This strategy helps improve the success rate of GRC projects, and reduces costs by generating informative data on the first attempt for a given experiment.

The GRC also maintains equipment that is accessible to faculty, staff and students of University of Idaho. This equipment, collectively called the “GRC User Core”, is primarily designated for high throughput sample preparation and quality assurance. Users are trained by GRC laboratory staff before scheduling time to use the equipment, and are responsible for any reagents needed to run their samples. When needed, GRC staff are available to help troubleshoot.

## Existing Infrastructure

The Genomics Resources Core Facility has the equipment necessary for applications of DNA sequencing technology, high throughput sample preparation, quality assurance, and bioinformatics analysis. The Core facility occupies two laboratory spaces in the IRIC building. IRIC 210 houses the controlled access component of the GRC which includes desk space for wet-lab staff and equipment required for preparing libraries and sequencing nucleic acid. The GRC Common User Core is located in IRIC room 142. The GRC has two offices, IRIC rooms 224 and 226 which house the bioinformatics scientists and are located a short distance from the wet-lab facilitating easy communication. The IRIC building is centrally located on the University of Idaho main campus in Moscow, Idaho. The Core facility infrastructure is described in more detail below.

## GRC DNA Sequencing Laboratory

DNA sequencing has become an indispensable tool for basic biological research, biomedical research, diagnostics, and molecular systematics. Current applications using DNA sequencing include whole genome shotgun sequencing, for de novo sequencing of previously unknown genomes; transcriptome sequencing; targeted re-sequencing; transposable element enrichment; single nucleotide polymorphism (SNP) discovery; metagenomics and amplicon sequencing for studies on microbial community composition; and many other applications. The Core facility also has equipment and robotics for high throughput sample preparation associated with activities upstream of DNA sequencing, such as library preparation. This equipment enables researchers to

streamline sample preparation, thereby reducing the costs of operating the core. Presently, the core has the following equipment in its DNA Sequencing Laboratory:

- DNA Sequencing

  Illumina MiSeq Sequencing Platform: Paired-end sequencing of up to 600bp per library-fragment and 15Gb of DNA sequence per run.

  Illumina HiSeq Sequencing Service: Paired-end sequencing for projects requiring higher-than-MiSeq read-density; libraries prepared in the GRC, but outsourced to collaborating facilities (this process is transparent to clients at the University of Idaho because the GRC handles billing and data transfer).

  Oxford Nanopore Sequencing Platform: Due to the experimental nature of the Oxford Nanopore platform, the GRC does not yet offer this type of sequencing as a standard service. However, the GRC has developed the expertise necessary to partner with collaborators interested in utilizing this technology on an experimental basis.

- Library Qualification and Quantification

  Life Technologies StepOnePlus: Quantification of sequenceable libraries via qPCR.

  Advanced Analytical Technologies Fragment Analyzer: Capillary array based high-throughput quality assessment of all DNA and RNA samples.

  Agilent 2100 Bioanalyzer: Sizing, quantification, and quality control of DNA, RNA, proteins and cells in low-throughput fashion.

- Library Preparation and Size-Selection

  Fluidigm Juno: Creates sequencing libraries of up to 2400 amplicons per 192 sample chip for targeted-resequencing. Highly automated for minimal hands-on time and high throughput.

  Fluidigm Access Array: Creates sequencing libraries of up to 480 amplicons per 48 sample chip for targeted-resequencing.

  Wafergen Apollo 324: Automates next generation sequence library preparation workflows for Illumina, Ion Torrent, and 454.

  Sage Biosciences BluePippin: Automated and customizable PFGE-based size-selection of DNA fragments between 90bp and 50kb with no cross-contamination.

  Covaris M220: Highly reproducible DNA-shearing between 150bp and 5kb.

  Invitrogen E-Gel System: Size-selection and visualization of library and DNA respectively.

- Sample Quantification

  Molecular Devices Plate-Reader and Invitrogen Qubit 3.0: Fluorometric quantification of DNA and RNA (hundreds of samples or single samples depending on device) yielding more accurate and reliable concentrations than NanoDrop.

## GRC User Core: High Throughput Sample Preparation and Quality Assurance

By acquiring new instruments in the GRC User Core for high-throughput sample preparation and quality assurance, the GRC provides researchers with the ability to increase sample quality while simultaneously reducing sample-to-sample variability and the time required for procedures. Equipment in the GRC User Core used for high sample throughput and quality assurance include:

- DNA, RNA, and Library Qualification

  Qiagen QIAxcel: Providing "digital gels" for all DNA and RNA less than 3000 bp in high throughput fashion.

  Molecular Devices SpectraMax Paradigm: Multimode modular microplate reader currently capable of high-throughput quantification of DNA & RNA.

- Sample -prep DNA & RNA purification

  Thermo Scientific KingFisher Flex: Automated high speed purification of nucleic acids, proteins, and cells in a 96well format using agnostic reagents and kits.

  Qiagen QIAgility: Highly customizable liquid handler for qPCR assay setup and other tasks benefiting from accurate and reproducible pipetting.

  Boreal Genomics Aurora: Gel based isolation, purification, and concentration of DNA from highly contaminated sources using Boreal's proprietary SCODA electrophoresis.

  Diagenode Bioruptor Plus (UCD-300): High-volume sonication/shearing of DNA, chromatin, cells, and tissue.

  BioRad T100: Basic touch-screen thermal-cycler for labs lacking this capability.

  Qubit 2.0 Fluorometer: Fluorometric quantification of DNA and RNA (hundreds of samples or single samples depending on device) yielding more accurate and reliable concentrations than NanoDrop.

  GRC staff continuously monitor current technological methods and trends for potential new equipment that will contribute to the mission of the GRC, both in the DNA sequencing laboratory and the GRC User Core. Each piece of equipment is evaluated for its ability to increase potential service offerings, improve the

quality or reduce the price of existing services, increase automation and throughput, and/or augment the existing equipment in the GRC User Core. These features are considered from the perspective of the stated mission – to facilitate cutting edge research in "real time evolution."

## Recent Infrastructure Investments

During calendar year 2017, the GRC has not made any large infrastructure investments.

## Planned Infrastructure Investments

The GRC is currently evaluating the purchase of the following equipment for addition to the GRC Sequencing Laboratory:

LightBench MK II: The LightBench, manufactured by Coastal Genomics, is an automated gel-based fragment analysis system which allows for highly accurate DNA/RNA size selection. The GRC is evaluating whether this instrument could lower cost and time investment in sequencing library quality control through automation, while improving the consistency and quality of sequencing runs by more effectively removing small fragments (primers and primer dimers) and genomic DNA which can lower the quality of sequenced reads, or cause sequencing runs to fail in some cases.

## Personnel

The IBEST Genomics Resources Core facility operates as a "turnkey" facility in which project design, sample preparation, data generation, and data analysis are integrated within a single facility. Therefore, the GRC has two main components: the "wet" lab and the "dry" lab, with the GRC Director overseeing both laboratories. The "wet" laboratory is staffed by professionals with molecular biology expertise and is where data are generated from samples provided by investigators. The "dry" laboratory is staffed by bioinformatics data scientists and is where data generated in the "wet" lab (and in other facilities) are analyzed, summarized and interpreted. A significant amount of communication and coordination occurs between the "wet" and "dry" laboratories.

The GRC stays nimble by continuing to develop new partnerships with other service facilities and by purchasing equipment to automate molecular methods, allowing a small staff to perform the same quantity and quality of work as a core facility with a larger staff that lacks as many automated workflows.

Genomics Resources Core Director

The current GRC Core Director, Dr. Samuel Hunter, joined the core in January 2016. Dr. Hunter earned a Ph.D. in Bioinformatics and MS in Statistics from the University of Idaho after earning a B.S. degree with a double major in Biology and Computer Science from the College of Idaho. Dr. Hunter worked for the GRC as a Bioinformatics Data Scientist from 2011 to 2014 focusing on Microarray data analysis and high throughput sequencing

analysis, especially genome assembly, variant calling, and methods development. In 2014 Dr. Hunter left the GRC to work as a Computational Biologist at Dana-Farber Cancer Institute in Boston, Massachusetts. While there he focused on methods development for clinical cancer sequencing, creating a software tool for CNV detection which was integrated into the Dana-Farber/Brigham and Women's Hospital personalized cancer sequencing pipeline to aid in clinical diagnosis. His current duties with the GRC include management of day to day operations, existing projects and client relations, outreach, identification of new opportunities, technologies, and clients, retaining and recruiting staff, data analysis, and advising students.

Bioinformatics Data Scientist

This position is responsible for bioinformatics and analysis of genomics data, and is currently vacant, but will be filled following the successful completion of an ongoing search. It was previously occupied by Dr. Alida Gerritsen who served through June 2017.

Genomics Laboratory Manager

Mr. Daniel New is responsible for the day-to-day operation of the GRC "wet" laboratories which includes the DNA Sequencing Laboratory and GRC User Core. Dan earned a B.S. degrees in Microbiology and Molecular-Biology/Biochemistry from the University of Idaho in 2005 while concurrently working as an undergraduate researcher to learn basic molecular techniques from 2003-2005. Prior to joining the Core in 2010, Mr. New was a Research Associate at Washington State University in the College of Veterinary Medicine where he gained experience in RNA extraction, relative-qPCR, mammalian cell-culture and transfection, microarray printing/processing, Sanger sequencing/instrumentation, PFGE, MLVA, and Kirby-Bauer assays. Starting in early 2016, Mr. New has worked closely with Dr. Gerritsen to learn basic bioinformatics skills necessary for routine data delivery.

## Genomics Laboratory Scientist

In May 2016 the GRC hired Matthew Fagnan as a temporary employee. Mr. Fagnan completed a B.S. in Bioengineering at WSU in May 2016 and spent the summer working closely with Mr. New and Dr. Gerritsen to gain the necessary skills for working in a high-throughput sequencing laboratory. These skills include training and familiarity with Qubit, TBS-380, Fluidigm 48.48 Access Array, Fluidigm Juno, Fragment Analyzer, aPCR, Wafergen Apollo, Ampur Bead cleaning, Linux, as well as experience with customer service and research and development. Following this successful evaluation/training period, the GRC hired Mr. Fagnan as a permanent employee to fill the Genomics Laboratory Scientist position. As part of an ongoing GRC objective to address recruitment and retention difficulties, as well as to improve redundancy and stability, Mr. Fagnan will continue to work with Mr. New, Dr. Hunter and Dr. Gerritsen to develop more advanced wet lab and bioinformatic skills. Additionally, Mr. Fagnan has chosen to pursue an MS in Statistics (with a focus on biostatistics) at the UI while continuing to work in the GRC.

## Bioinformatics Research Assistant

David Streett was a graduate student in Bioinformatics and Computational Biology until he graduated at the end of the Summer 2016 term. During his half time appointment in the GRC, he was responsible for the development of bioinformatics software and analysis pipelines. He received his B.S. in Biochemistry in May of 2015. His Master's thesis project was a software package for preprocessing Illumina sequencing reads. A publication is currently in preparation and the software is available online at https://github.com/ibest/HTStream.

## Interns

During the summer of 2017, the GRC experimented with hosting undergraduate interns. The interns Neale Ellyson and James Styer worked with GRC staff to learn molecular techniques and gain experience working in a sequencing core. James continued working with the GRC during the Fall 2017 semester to fulfill lab-rotation requirements for his degree.

## Bioinformatics Post-Doc

During the summer of 2017, the GRC worked with Dr. Haiqing Sheng in the College of Agricultural and Life Sciences to hire a temporary employee to complete a bacterial genome assembly and annotation project. The GRC identified a candidate (Dr. Mingrui Duan) for this position, and she was hired with funding from Dr. Sheng. The GRC provided desk space, access to computational resources, and training for Dr. Duan during this project. Following completion of the genome assembly project, Dr. Duan was hired as a postdoc through a collaboration between IBEST and Dr. Brenda Murdoch in the College of Agricultural and Life Sciences. Dr. Duan will work with GRC staff and Dr. Murdoch's team to develop an ovine genotyping platform with research and agricultural applications for marker assisted selection and parentage analysis. Dr. Duan has an MS in Biochemical Engineering and a Ph.D. in Molecular Biology.

## Bioinformatics Analysis Resources

The GRC does not maintain any specialized equipment for data management or bioinformatics analysis; instead, it maintains a strong partnership with the University of Idaho IBEST Computational Resources Core facility. This tight integration between the GRC and CRC has numerous advantages. First, the CRC provides the storage and computational power necessary for the analysis of the large-scale genomic data sets that are produced by the GRC. Second, the collaboration between the cores provides a great deal of agility with regard to the development of new bioinformatics techniques and analyses. This fosters innovation and creative activity that are the hallmark of IBEST, and differentiates the GRC from other more "traditional" genomics core facilities around the US and the world.

## Services and Innovation

The Genomics Resources Core offers "*genomics project management*" to customers by integrating services in all three phases of genomics research: project planning and consultation, genomic data generation, and bioinformatics data analysis. In contrast, most core facilities around the country focus mainly on data generation, leaving investigators to struggle with immense data sets with little help. Our integrated approach is very unusual, and a key component to our continued success. This has led to a large amount of off-campus clients (both U.S. and International) through "word-of-mouth advertising" which is balanced with our on-campus workload (see Fig. 1). To track and manage the growing GRC user base, the Core implemented iLab project management and billing system in late FY2014. Over the past two years, the Core has been able to use iLab to accurately track usage data from internal and external users and effectively bill for bioinformatics time.

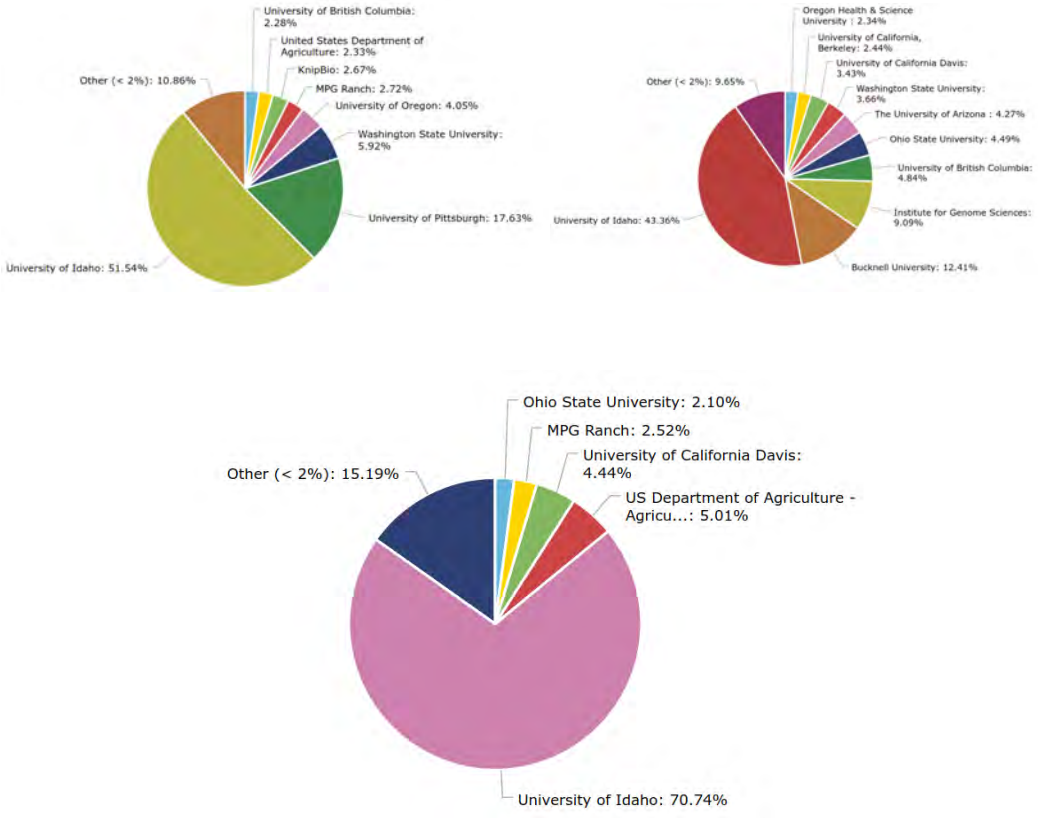


**Figure 7.** The top 10 institutions (by total revenue) in billed work from the GRC during FY2015 ($185,052), FY2016 ($212,420), and FY2017 ($200,324).

## Project Consultation

Core facility staff consult with investigators to discuss project aims, expectations, experimental design, appropriate and best use of technology, sample quantity and quality issues, and data analysis needs. During consultation, a project time line is formed, expected costs are discussed, deliverables are identified, and a user agreement is reviewed. Having these discussions early in a project provides an opportunity for Core personnel to offer their expertise, advice, and assistance to enhance the proposed project and sidestep potential problems. Initial consultation is a service that the GRC currently provides free of charge.

This service is especially important to researchers developing grant proposals, where a detailed quote and sophisticated understanding of the protocols and analysis are likely to increase chances for funding and ensure accurate budgeting. In 2017, the GRC has provided letters of support and/or consultation for several researchers seeking funding including Dr. Onesmo Balemba (UI), Dr. Diana Mitchell (UI), Dr. Ryan Driskell (WSU), Dr. Sanford Eigenbrode (UI), Dr. Eva Top (UI), Dr. Daolin Fu (UI), Dr. Jianli Chen (UI), Dr. Jeff Langman (UI), Dr. Shirley Luckhart (UI), and Dr. Paul Rowley (UI among others. Providing this service free of charge ensures that researchers come to the GRC to develop a detailed plan at an early stage of their project and can develop a cost structure for proposed experiments, including sequencing and bioinformatics. This approach helps keep overall costs low, expectations realistic, and potentially costly problems minimal in the latter stages of a project.

## Genomics Data Generation

The Genomics Resource Core facility operates and maintains equipment (described above) that allows high throughput sample preparation, quality assurance, and generation of high throughput DNA/RNA sequence data. While the Genomics Resources Core operates much of the equipment necessary to perform the work proposed by its clients, there are instances when projects require technologies that are not present in the facility. In these cases, the GRC facilitates access to the technology through cooperation and collaboration with other regional core facilities. For example, when investigators require the additional capacity provided by the Illumina HiSeq platform, the GRC staff prepares Illumina libraries that are sent to other institutions for sequencing (such as University of California Berkeley or the University of Oregon), and the data are then sent back to the GRC for processing and analysis. The GRC has also developed a relationship with the University of Washington PacBio sequencing core in expectation of the need for long-read sequencing in the future. The fact that the sequencing was done “off-site” is seamless and causes no additional work for the investigator. This expands the range of services the GRC can offer without incurring additional capital expense. A time series of expenditures by type is shown in Figure 8 and top ten services by total cost in Figure 9.

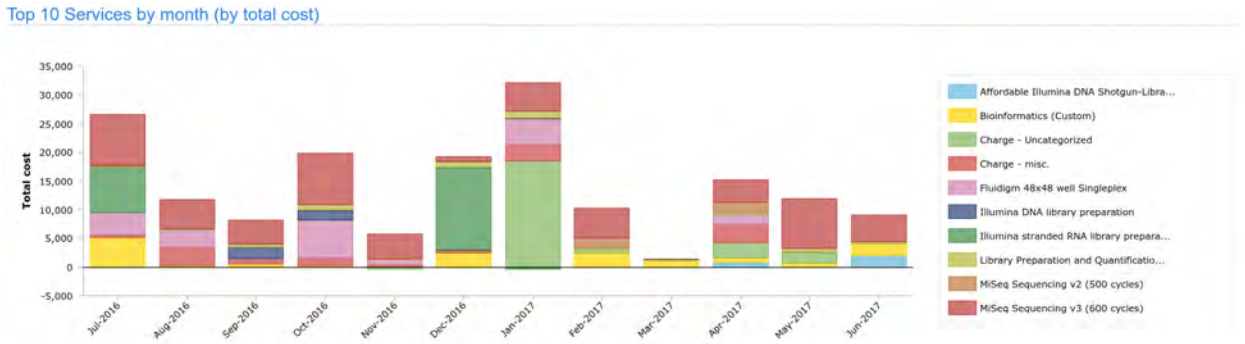


**Figure 8**. Time series of expenditures by type of service, FY2017.

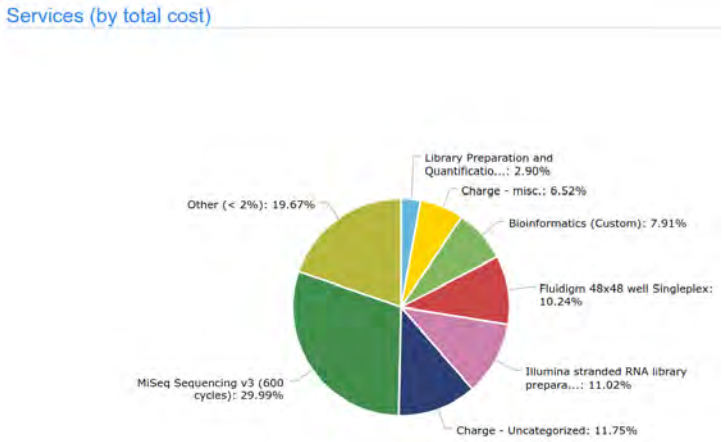


**Figure 9**. Summary of top ten types of service by cost, FY2015.

## Bioinformatics and Data Analysis

The GRC continues to increase its user base for genomic data generation and as a result has increased charges associated with bioinformatics analysis. Bioinformatics data analysis is often the most challenging aspect of any experiment, and until very recently was often overlooked in budgeting for experiments. The current system accurately tracks personnel hours on independent projects and reflects the effort that is expended for analysis.

The GRC offers bioinformatics services through staff bioinformaticians and can perform a full range of analysis tasks to address questions in areas such as population genetics, genomics, microbial community dynamics, functional genomics and systems biology. GRC bioinformaticians begin with raw output from genomics equipment and proceed through quality assurance, data processing and analysis, data interpretation and visualization. Analyses are conducted using pipelines in the public domain or those developed by Core staff members. Core personnel have developed analytical techniques and pipelines for microbial community analysis, genome assembly, transcriptome assembly, population variant analysis, SNP/INDEL detection, and RNAseq analysis. These pipelines transform and manipulate raw data into a form and format that can be mined by investigators.

Data processing occurs through a feedback loop with investigators. The GRC bioinformaticians seek feedback from investigators after preliminary data analysis, so that adjustments in output content, form, and format can be made. Data are then re-analyzed or additional analyses are performed until the project's goals are met, figures are generated, and summary tables are provided to the investigators in a form that is useful to them. The Core staff provides investigators with detailed knowledge of the laboratory protocols and bioinformatics methods used so they can be included in reports and publications as needed. As a result, core staff members are often included as co-authors on publications because of their significant intellectual contributions to research projects.

## Innovative New Methods

GRC staff has participated in the design and development of new methods and techniques for genomics research. Example projects are briefly described below.

### Assembly by Reduced Complexity (ARC)

As a part of his PhD dissertation, Dr. Samuel Hunter developed Assembly by Reduced Complexity (ARC), a software package for targeted assembly of homologous sequences. The algorithm works by comparing reads to a set of reference targets and bins them based on the results of these comparisons. Assemblies are then performed on sequences from each bin. ARC works effectively with divergent references, functions well with short sequence reads, and compares favorably to de novo assembly in terms of CPU and memory requirements.

### A Modular, Highly Multiplexed Design for Illumina Amplicon Sequencing

Dr. Matthew Settles, in collaboration with Mr. New and Dr. Gerritsen, developed a laboratory protocol and data analysis platform for performing highly multiplexed Illumina amplicon sequencing. PCR amplicon sequencing is an important tool used to query genetic variation and structure in individual samples and ecological communities. Applications range from determining the composition and structure of bacterial and fungal communities to determining allele frequencies in a set of genes across many individuals. This methodology provides a way to simultaneously sequence and analyze hundreds of samples across one or many targeted regions in the same sequencing reaction while significantly reducing experimental costs.

The analysis platform is a comprehensive application that starts with raw sequence reads and ends with abundance tables of taxonomically assigned sequences for community analysis. Additionally, the application is able to prepare reads for input into phylogenetic tree building software. The software project is ongoing, relying on user comments and feedback to continue improving the functionality and efficiency of the program.

### HTStream

HTStream is a toolset for preprocessing Illumina data to remove various forms of noise common in this type of data. It consists of a set of algorithms implemented in C++ as

standalone programs which can be easily chained together to form custom pipelines, tailored for specific project needs. HTStream was developed by David Streett as part of his master's thesis while he worked for the GRC.

## Sustainability

Service center fees are established based on the estimated costs of consumables, instrument maintenance agreements and personnel time associated with each service and updated on a semi-annual basis. Clients who request custom bioinformatics analyses or new method development are provided a cost estimate based on the amount of time expected to complete the proposed work.

During FY 2013-2014 there was a significant shift in the types of services the GRC offered. Specifically, the GRC phased out equipment for DNA microarrays (purchased 2011), DNA genotyping (purchased 2011), and Roche 454 Pyrosequencing (purchased 2009). Each of these technologies was displaced by new, less expensive technology (such as the Illumina MiSeq). These upgrades produced a 'more data for lower cost' effect, which resulted in a decrease in GRC annual revenue from $369,314 in FY2013 to $203,198 in FY2014. A slight increase in demand for services was reflected by a total revenue of $219,688 in FY2015, however the GRC was severely understaffed during much of this fiscal year, limiting the amount of time that could be spent on outreach developing new opportunities. Ongoing R&D efforts and new acquisitions continue to be targeted at decreasing reagent costs and improving efficiency to offer updated services while maintaining competitive pricing for existing and new clients. Outreach efforts and a renewed focus on identifying new opportunities for collaboration are expected to result in further improvements towards sustainability.

Additionally, GRC staff have continued to focus on supporting the generation of preliminary data, and grant writing efforts by UI Faculty. Successfully funded grants provide future work and revenue for the GRC, as well as overhead return increasing the sustainability of IBEST and ensuring continued support from IBEST staff.

### Plans

IBEST successfully completed a search for a new GRC Director in late 2015. The new director has continued to transition the Core towards a business model focused on long term sustainability. Included in this transition is continued investment in key strengths such as multi-locus targeted amplicon sequencing. Based on extensive research online, discussions with Fluidigm, clients, and other core directors, this service is offered for non-model organisms by only one other sequencing center in the country. Limited availability, in combination with presentations at conferences and publications by early adopters has led to increased demand for this service. This increased demand will be met by the acquisition of the Fluidigm Juno platform which will significantly reduce reagent cost per sample and hands-on time by quadrupling throughput as compared to our existing Access Array system (a close collaboration with Fluidigm research scientists has enabled us to maintain backwards compatibility with assays run on the Access Array improving redundancy of key service).

Another key strength of the GRC, metagenomic community analysis using 16s and other targets, has also been an area of continued focus. As well as continuing to invest heavily in training clients to prepare libraries using our custom set of dual barcode indexes, we have launched an initiative to improve reproducibility and consistency across runs by developing a set of positive controls consisting of an internal (standardized library) and external (mock community) control which will be sequenced at low depth on each amplicon run. These controls will allow us to monitor for run-to-run variations within the GRC, as well as simplify trouble shooting library prep for new clients as well as variation between library preps for existing clients.

The GRC has continued to look for new ways to expand the applications of this targeted amplicon strategy by identifying researchers at WSU and the UI who can employ this strategy to characterize mutations generated by the exciting new CRISPR/Cas9 gene editing technology. Collaborations with Drs. Jim Nagler (UI), Deb Stenkamp (UI), and Michael Varnum (WSU) were successful in all cases, producing data that was included on a poster and in two manuscripts.

Additional research and development objectives aimed at increased sustainability include adopting new protocols for decreasing costs and increasing efficiency of shotgun and RNAseq library preparation, both of which are currently in progress. Finally, increasing the user base continues to be a major objective. Efforts on campus have included consulting, sequencing, analysis and grant writing support for members of the College of Agriculture and Life Sciences as well as members of the College of Natural Resources. Fostering collaborations with off campus researchers has included collaborations with faculty at UI extension offices and USDA facilities in Idaho, as well as with researchers in neighboring states. Although “word of mouth” advertising continues to bring us many new clients, our collaboration with the Idaho Wheat Commission has also increased our visibility, opening many opportunities to work with a variety of agricultural researchers across the Northwest.

A second component of sustainability is recruitment, retention, and minimizing the damage caused by losing a staff member. The increasing popularity of next generation sequencing methods has led to high demand for experienced staff. This represents a serious concern for the GRC which has a small staff and has traditionally had little redundancy in skills and significant risk associated with losing a staff member. A major objective of the new Director has been to address these problems through a combination of strategies.

The GRC currently (and historically) has been staffed entirely by members who gained a large portion of their training in High Throughput Sequencing techniques within the GRC. Although recruiting highly skilled staff members is always a possibility, the rapid growth and popularity of High Throughput Sequencing has created consistent demand for experts in this area. Rather than relying on recruitment of such skilled staff, the GRC will continue the strategy of developing capacity by hiring entry level positions, investing in professional development and training, and promoting when appropriate. As mentioned earlier, this model was used for the current Director, Bioinformatics Scientist, and

Genomics Laboratory Manager as well as the new Genomics Laboratory Scientist (who was hired as a summer intern with no previous HTS or genomics experience).

To reduce the threat posed by loss of a staff member, redundancy in job roles has been improved by training wet-lab staff in basic bioinformatic and data delivery procedures, ensuring that a subset of services would continue to be available even in the absence of bioinformatic staff. Redundancy in wet-lab procedures has been improved by hiring a Genomics Laboratory Scientist who has been trained in nearly all critical wet-lab procedures. Redundancy in administrative tasks is ensured by continued involvement of Mr. New and Mr. Fagnan in developing new business, billing, and other administrative decisions where appropriate.

Concerns about retention have been partially addressed by somewhat reducing workload and stress through hiring additional core staff and by working with IBEST and University administration to provide above baseline salary increases for existing staff members. A recent Market Based Compensation initiative within the University has also led to significant salary increases for some GRC staff members, helping to address salary levels which were significantly below market levels. Unfortunately these changes did not go into effect until after the GRC lost one staff member to a better paying position.

## Outreach

The Genomics Resources Core engages in several outreach activities across the University of Idaho campus, the state of Idaho, across the nation. Examples of outreach activities include:

*Professional associations:* Genomics technology partnerships and consultation with service centers and researchers at the University of Oregon, University of California-Berkeley, University of California-Irvine, University of Montana, Washington State University, and the University of Washington.

*Conferences:* Travel to technological and administrative conferences developed for service centers and core facilities, including the Association of Biomolecular Resource Facilities and the Western Association of Core Directors. Novel approaches to analysis, budgeting, customer service, sustainability, and technological innovations are all topics that are encountered at these conferences.

## Opportunities

The Genomics Resources Core continues to look for opportunities for new customers and collaborations. Of interest are the potential synergies with center-type research programs. For example:

- Continued involvement with researchers working with the Idaho Wheat Commission has led to further support for original research on the phylogeography and population genomics of the Sugarbeet wireworm, *Limonius californicus*. This work has already produced a number of genomic resources which are being leveraged by the GRC, Dr. Arash Rashed's group, and Dr. Daolin Fu's group to test an RNAi strategy for controlling wireworms. Additionally

involvement with the IWC has led to an opportunity to work with Dr. Jianli Chen in an attempt at finding a mutation associated with resistance to the dwarf bunt fungus in wheat.

- The NIH Center of Biomedical Research Excellence (COBRE), called the CMCI, includes projects that will require genomics technologies and a systems biology modeling collaboratorium that will engage both the GRC's "wet" and "dry" labs. The GRC will occupy space in the same building as CMCI following a planned move to the new IRIC building in January 2017.

## Future Objectives

### Challenges

Maintaining a balance between accessibility and financial sustainability continues to be the biggest challenge for the GRC. The GRC operates under a unique structure that integrates all three phases of genomics project management - combining data generation and bioinformatics like few other facilities in the United States. This is both its greatest strength and its greatest ongoing challenge. Because the GRC is so unique, there are few (if any) other facilities that can serve as a model for growth and sustainability. In addition, the scope of research facilitated by the GRC is complex and highly varied, working with a wide variety of data types, non-model organisms, and a range of experimental protocols. This challenges staff to develop expertise pertinent to a wide range of technologies and methodologies, and can limit the ability to develop high volume standardized work flows. Despite these challenges, the integrated approach remains the GRC's signature characteristic and is a key component to continued success.

Another challenge for the GRC that is related to financial sustainability is the lack of recognition the GRC receives for molecular and bioinformatics work. Many clients assume that because GRC services are paid, the GRC staff should not receive recognition as part of the publication process. However, because of the investment of time through multiple, (typically free) consultations with researchers, the GRC staff has a significant intellectual impact on many of the projects submitted to the core. These intellectual contributions often result in completely redesigned projects which take better advantage of modern technologies and more effectively address the research question. These impacts should be attributed and will help favorably increase the core's reputation amongst the scientific community, and will also justify the continuing University investment into this shared resource. The GRC has taken steps towards addressing this problem by requesting that new and existing clients review and sign a document acknowledging these contributions, and reminding clients of the necessity of citing the COBRE grant which has supported development of the core facility.

Perhaps the most significant threat to the Genomics Resources Core continues to be its ability to hire new staff and retain them. Classification and pay scales at the UI have undergone significant revisions, largely addressing the difficulties in offering competitive

salaries. However, the bioinformatics field continues to grow, creating many opportunities for qualified staff and maintaining high demand in the job market.

## Future Directions

The IBEST Genomics Resources Core will continue to offer state-of-art services in genomics and bioinformatics that will enable University of Idaho investigators to overcome the "barriers to entry" posed by their own lack of expertise in these fields. Collaborating with the GRC will allow them to pursue new avenues of research that leverage the resources available within IBEST. The goal is to continue to provide integrated services to IBEST researchers – facilitating cutting edge research in genomics and real time evolution.

The GRC constantly evaluates the portfolio of offered services, a critical activity because the field of genomics changes quickly. New technologies emerge every year, and the capacity for data generation is outpacing the capacity to store, analyze, and interpret these data. Staying on top of emerging technologies and trends enables the GRC to continually identify novel business and collaborative opportunities while focusing on key existing services ensures financial stability, at least in the short term. Balancing these two competing interests will ensure that the GRC is sustainable, and adaptable as old technologies are replaced by new ones. The GRC's most important offering is in identifying and developing solutions to facilitate scientific discovery, particularly in areas where less integrated or more entrenched cores may be unwilling to innovate. To this end, the intellectual capital, expertise, and adaptability represented in the GRC are its biggest asset. So, while purchasing new equipment may be necessary to stay current or increase capacity, continued investment in personnel is the best way to ensure long-term viability of the core.

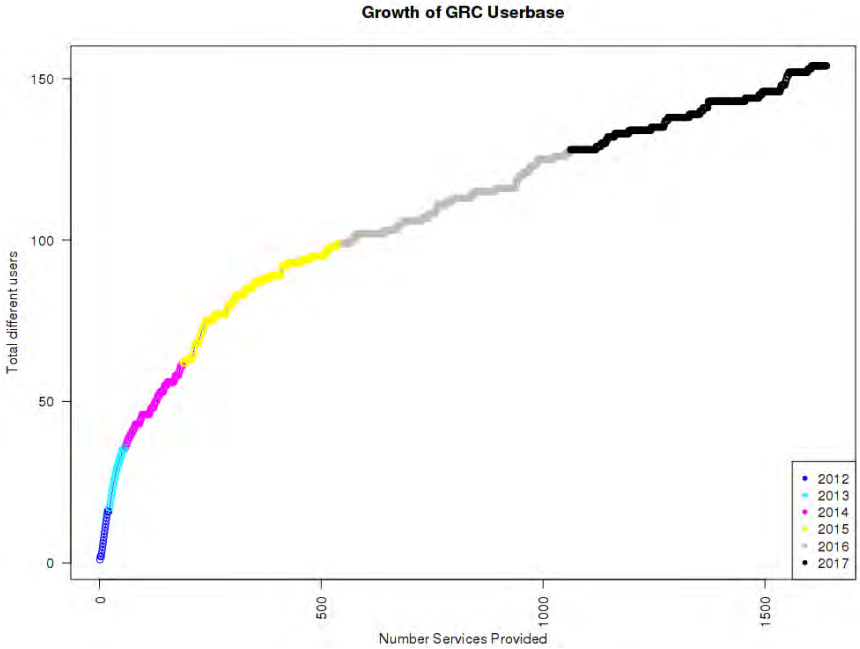


**Figure 10**. Growth in user base of the GRC since 2012.

# VII. IBEST Optical Imaging Core

The Optical Imaging Core (OIC) has been offering expertise, training and instrumentation for optical imaging and flow cytometry as a part of IBEST since 2011. These services will continue to be available to University of Idaho investigators though will not be administered by IBEST after June 30, 2018.

## Service Model

The Optical Imaging Core model offers investigators opportunities to advance their research using imaging and characterization tools in a variety of ways. The director of the OIC is available to provide a full service – data acquisition, analysis and presentation results – when an investigator requires expert assistance that may not be currently available among their existing staff, yet, most researchers choose to have their staff and students learn to work on the instrumentation independently, with occasional assistance from the director. These choices create an educational opportunity for staff and students and flexibility for the investigator. To provide adequate oversight for these sophisticated instruments and direct assistance when required for the users, the director has been available on a full-time basis. This design may no longer be a sustainable option.

## Breadth of Services

The instrumentation and expertise provided by the OIC creates options for investigators that would not be sustainable in their individual laboratories. By sharing the infrastructure investment, these research tools can be built with higher quality hardware and increased parameters for improved results while providing more services to more users. As modern optical imaging and flow cytometry tools emphasize characterization and transformation of biological structures based on fluorescent markers, the more markers that can be clearly identified creates increased depth and flexibility for any experiment. Acquiring more nuanced data from an experiment is efficient in acquisition costs, use of animals and speed of discovery.

## Infrastructure

The complex tools for fluorescent imaging, flow cytometry and analysis in the OIC have mainly been funded through instrumentation grants or program project grants. When making the final purchases we consider existing research programs and try to anticipate new directions. None of the current OIC instrumentation has active maintenance agreements and some of the instruments are quite old. The instruments below are listed in order of most recently purchased to the oldest, yet, currently active ones.

- **Analytical Flow Cytometer** – this system provides characterization of cells and particles in solution based on both fluorescent and non-fluorescent parameters. The configuration of four lasers and 13 fluorescent channels provides great flexibility in experimental design and opportunities for new

users. The instrument is simple to learn and maintain, inexpensive to run and provides a high throughput option with a plate loader.

- **Flow Cytometry Analysis** – along with the new analytical flow cytometer we were able to purchase a faster computer and new software for more sophisticated analysis.
- **Spinning Disk Confocal microscope** – this system allows for great flexibility in imaging with three cameras, improved speed of imaging over the older laser scanning confocal system, environmental conditions for live imaging and potential for optogenetics studies. Though this system cannot replace the high resolution of the laser scanning system, most users have moved to it for the improved speed and more comprehensive software.
- **Image Analysis** – the funding for the spinning disk system allowed us to purchase an updated computer, increased storage space and new analysis software that matches the acquisition software.
- **Stereo microscope and histological microscope** – a stereo microscope and traditional histological microscope share a very basic color camera. The stereoscope is also used for preparing samples for the confocal microscopes and has fluorescent viewing.
- **Fluorescence Activated Cell Sorter (FACS)** – a flow cytometer that also sorts cells is available for users that want to save a subset of their characterized cells for further processing. This is an older instrument and no longer supported by the manufacturer. It is limited by having only two lasers, yet, is still used by an occasional investigator. There are no other FACS systems on campus, though, some investigators have used systems at WSU when this system has been found to be of the wrong configuration for their experiments.
- **Laser Scanning Confocal microscope** - for high resolution 3D and second harmonic imaging, this is the only system on campus. It is an older system that is underutilized, so any major maintenance needs will have to be carefully considered.

## Potential Infrastructure

The current director of the IBEST Optical Imaging Core will be retiring in June, 2018. She will make a few recommendations for new instrumentation to consider, as listed below, yet, the OIC will no longer be a part of IBEST.

- **Multiphoton Microscope** – a carefully configured multiphoton laser system will satisfy those users that require second harmonic imaging as well as multiple wavelengths, high resolution and improved imaging depth for larger live experiments and fixed samples that require high resolution imaging. This could replace some aspects of the older laser scanning system.
- **Slide Scanning System** – newer slide scanning systems allow for high throughput imaging of histological and fluorescent based samples. A few

investigators have inquired about such as system and if in place might provide fast, overview imaging for investigators to prioritize which samples should be further imaged at higher resolution and magnification.

- **Fluorescent Activated Cell Sorter** –a modern cell sorter would provide investigators with multiple labeling options, improved sensitivity and create single cell samples for sequencing. Discussions could be considered to make improved arrangements for this service to be provided by WSU as opposed to purchasing an instrument for UI.

## Innovation

The Optical Imaging Core instruments and expertise are focused on high resolution imaging and cell characterization through flow cytometry. These sophisticated tools are rarely part of any investigators daily work flow and the OIC has always seen an ebb and flow of activity. New projects have been discussed with the director, from both new and existing investigators, yet, major projects have not come to fruition this past year. Many of the heaviest users have recently graduated so their publications and new proposals will likely bring in some additional funding, especially in the neurobiology imaging arena. The director has had conversations with multiple PIs about potential collaborative projects and multi-investigator proposals that are still developing.

The Optical Imaging Core could be a place of additional services and discovery. Examples may include building instruments for preliminary studies. We have the talent and interest on our campus to create imaging options that are specific to acquiring data for modeling projects and if configured appropriately could provide preliminary data for a team of investigators. In addition, the OIC has had a number of requests for assistance on building a workable microscope for routine work in individual laboratories. The OIC now has an inventory of parts, which when combined with some new hardware, would help young investigators jump start their initial projects without investing too much money into high end systems. This could become a more routine service and marketed as part of a start-up package.

## Sustainability and Investment

Working with the Service Center Committee, two years ago, the Director developed a fee structure that included the opportunity for investigators to purchase passes for use of specific instrumentation over a 3-month period of time. Most investigators are now using this option as opposed to paying directly by the hour for use of the OIC services. Tracking of the hours of use has continued and helped inform the value of the passes for future fee determination.

The figures below show numbers related to the income and usage of the Optical Imaging Core in the past year. These numbers, when reviewed in relation to each other and to past years at the OIC, also may tell a number of stories. Figure 11 shows the OIC income for the last three years, yet, also likely reflects how that income has been affected by the recently developed pass options. The income in FY2017 is less than FY2016, yet, more

than FY2015. This income reduction is certainly partially explained by similar changes in hours of use for the same three years (Fig. 11), yet, also reflects the high number of pass purchases as compared to previous payment method of direct payment for each hour of use. For FY2018, the pass fees have been increased to better reflect their market value.

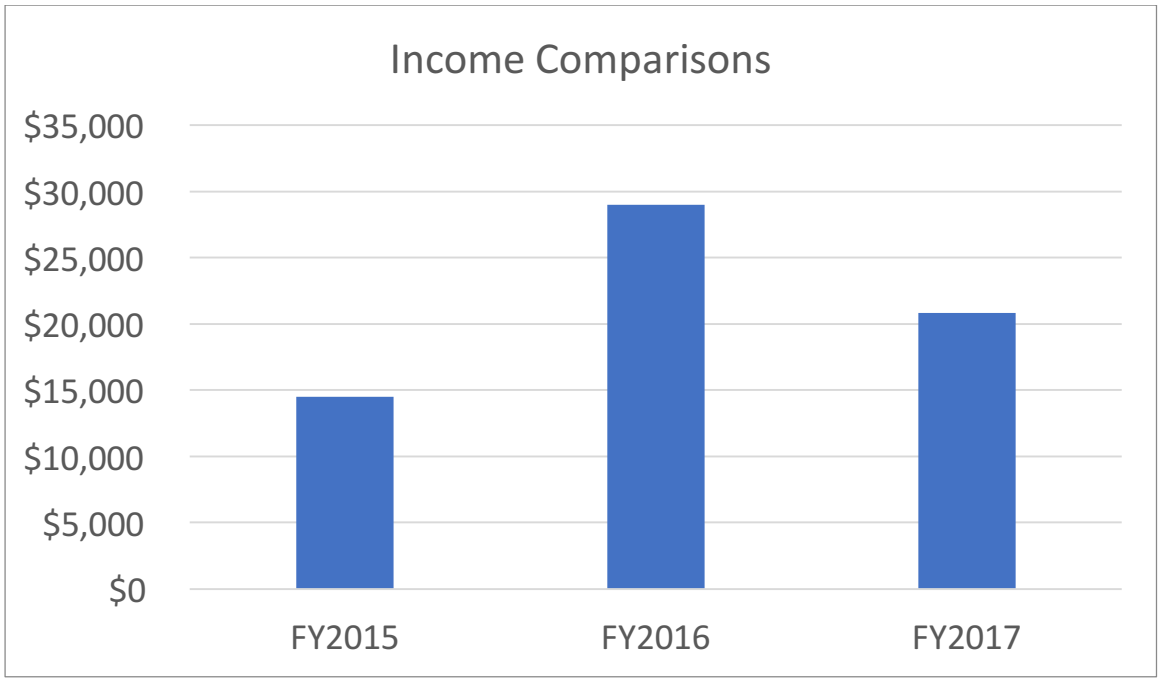


**Figure 11**. OIC Income for FY2015, FY2016 and FY2017.

Though the income for use of the OIC is generally received by the purchase of quarterly passes, the director has continued to keep a log of hours of use for each instrument and by user type. Some of this data is reflected in Figure 12. This data informs the value of each service, helps understand what type of user (Faculty, Staff, Graduate Student, Undergraduate Student) we are serving and the degree to which a particular service is used. These numbers help make important decisions about what service to invest in and what services to consider sun-setting. One noteworthy trend is that, though the overall number of hours of usage for the OIC was reduced between FY2016 and FY2017, there was a significant increase in the usage by undergraduates and slight increase by faculty directly. That is likely a reflection of the overall lack of funds available to users of the OIC in the past year, as when they are short of funds or when their graduate students complete their degrees, they use undergraduates who are often working for credit or research experience rather than wages. Also, when they lose their experienced help, i.e. graduate students, they return to the OIC to do some of the work themselves.

At least 20 new undergraduates were trained to use the OIC this past year. This is a wonderful opportunity for them to learn some hands-on skills and to go more in-depth on a project while they are attending the University of Idaho. As they are inexperienced in the nuances of working with biological samples and sophisticated, expensive

instrumentation, they do require additional oversight and efficient responses for a safe and instructive experience.

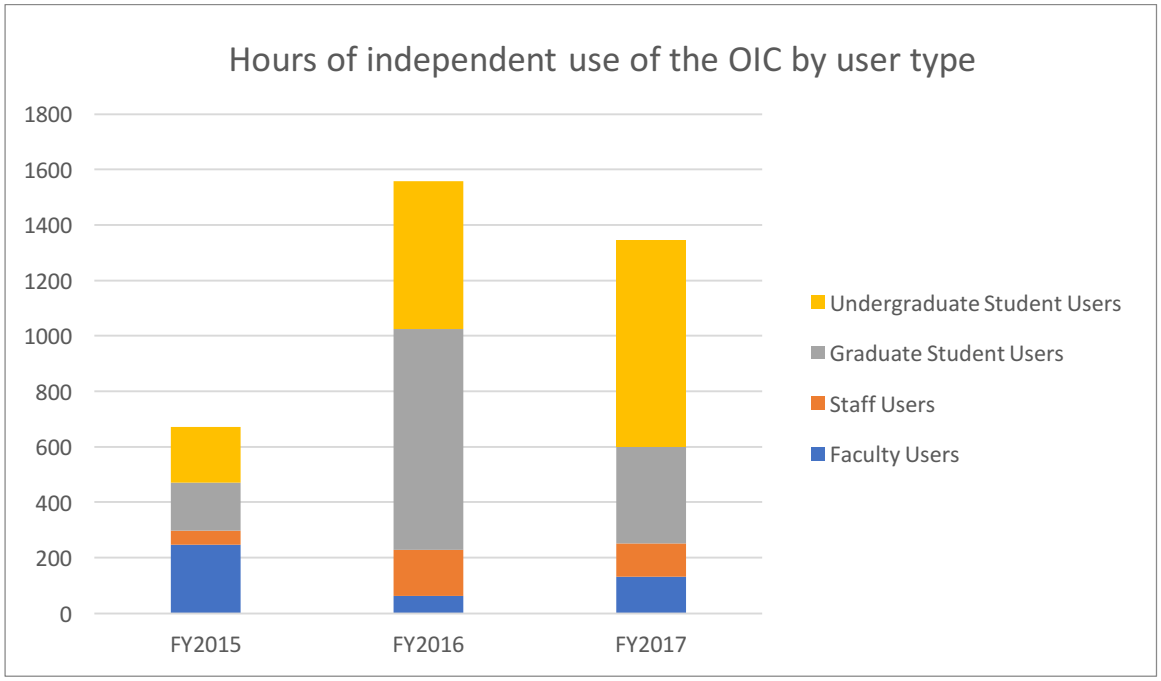


**Figure 12**. Hours of independent use of the OIC by user type during FY2015, FY2016 and FY2017.

The investment into the spinning disk confocal microscope, the new flow cytometry analyzer and associated analysis stations have proven to be very fruitful for users of the Optical Imaging Core in the past year. As the older instruments have seen reduced use, there has been a reduced requirement to spend time or money on maintaining them. In addition, likely because the main usage has been on newer instruments in the past year, there have been very few maintenance costs as these instruments have been performing well and the director has had more time to focus on their care. This has allowed FY2017 to be a year of lower operating costs.

The major operating expense of the OIC, now that there are no longer routine service agreements on the instruments, has been the director's salary. During much of FY2017, the Office of Research and Economic Development (ORED) has paid part of the director's salary as part of a direct investment into the OIC. Additional reduction in overall operating costs is that 25% of the director's time has been dedicated to other duties requested by ORED and therefore Figure 13 reflects only 75% of the director's salary and fringe costs. Most years, the director also attends training sessions and shared resource administrative meetings. This past year, all travel costs were paid for by ORED and the Association for Biomolecular Resource Facilities for the director's role as the President of the regional chapter and were focused more on the administrative functions of a core director, rather than any specific scientific or technical training.

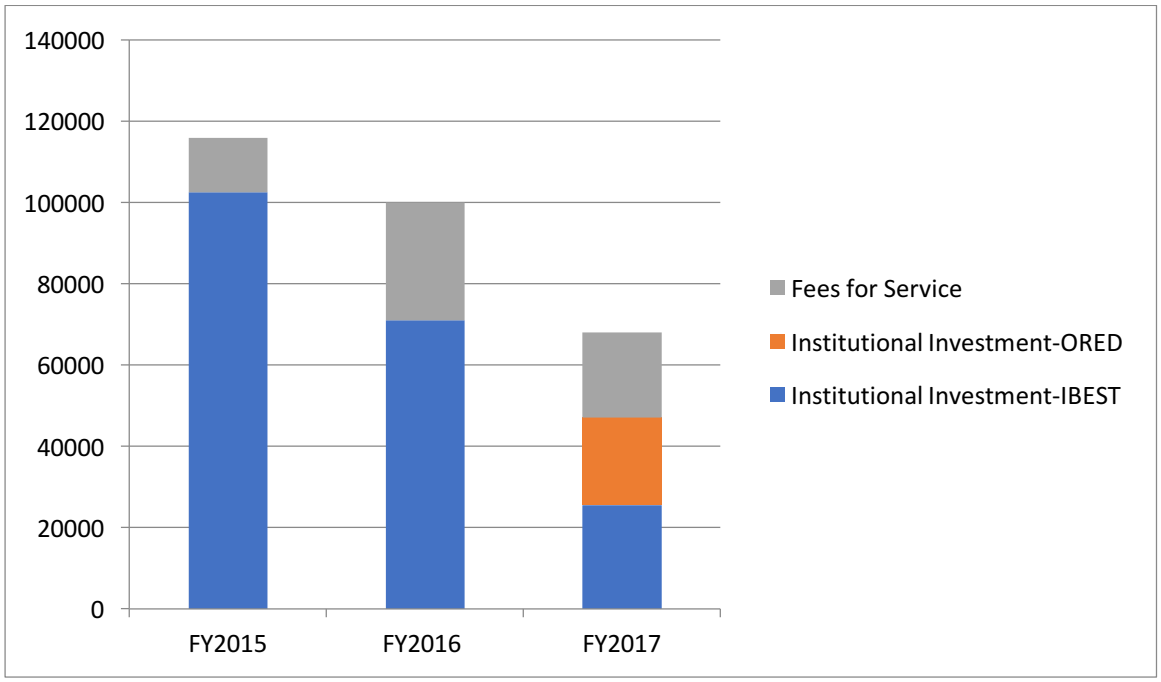


**Figure 13**. Total OIC operating revenue.

Therefore, the final operating costs for the OIC this year were $67,964. Those costs were covered approximately 37.5% by IBEST, 32% by ORED and 31% by Fees for Service (Fig. 13). The fees for service have recently been increased, which should increase the percentage of operating costs that are covered by Fees for Service in the future. It is likely that there will be increased maintenance costs in the upcoming years as the instruments age and replacements are established for those that are retired.

## Vision

Investments into high resolution imaging and flow cytometry are core to a research based university. Careful attention to the specific needs of the university's researchers, awareness of and investment in new techniques and technologies related to these powerful research tools as well as the continuation of a safe and inviting environment for all users to succeed with these tools should be part of the continued research mission at the University of Idaho. Efforts to find funding for replacement instrumentation is lagging, though, the possibility of on-campus built systems will help to bridge that gap and provide learning opportunities for students in many disciplines.

# VIII. Education and Training

## Bioinformatics and Computational Biology Graduate Program

Research and graduate education are symbiotic; IBEST has therefore provided extensive support for the graduate program in Bioinformatics and Computational Biology. The BCB program plays a unique role within the university because it prepares graduates who are at the forefront of a booming field, that of bioinformatics and computational biology. The major challenge today for mathematicians, statisticians, computer scientists and biologists is to develop innovative ways to analyze and interpret extremely large data sets in ways that will allow new discovery. In 2017, IBEST provided two BCB fellowships and we have committed $115,000 for BCB fellowships in 2018. We consider this a valuable investment because the vast majority of BCB students work in the laboratories of IBEST faculty and it therefore directly supports interdisciplinary research on evolutionary science. The BCB program currently has 22 students: 6 MS candidates and 16 PhD candidates. Since our last report, 2 students have graduated with an MS and 4 students have graduated with PhDs.

In addition, we continued our support for the IBEST/BCB Seminar Series, which, once again, attracted excellent scientists from across the nation and world to the campus of the University of Idaho. These formal seminars and informal interactions expose IBEST personnel to the research interests, ideas, and expertise of leaders in the field. Over the years we have realized an indirect benefit of our seminar series in that invited speakers return to their home institutions and spread the word about the impressive research done at the University of Idaho and the collegial and collaborative atmosphere within IBEST. This has bolstered our reputation in the scientific community, helped us recruit students, and attracted new customers to our core facilities. See Appendix IV for a listing of seminar speakers and topics in 2017.

These seminars (about four per semester) have been used as a core element of a graduate seminar course (BCB 501) and are open to the public. Often more than 50 people attend them. The scientist invited typically spends two days on campus meeting one-on-one with faculty members or small groups of students and postdocs. The graduate students of the Bioinformatics and Computational Biology (BCB) program choose and invite speakers for the seminar series and organize their itinerary. Beginning in FY 2017, we are merging our seminar series with that of CMCI as part of efforts to ensure that IBEST and CMCI are synergistic rather than competitive.

## Business for Scientists

In 2017, we offered the “Leading and Sustaining Your Research program” workshop for the third time. This workshop has been designed in partnership with the University of Idaho College of Business and Economics and teaches skills that are not part of traditional graduate and postdoctoral programs. These include project management, budget development, human resources administration, strategic planning, risk

assessment, team building, communication with stakeholders and especially lay people. The course was advertised across the Moscow campus, with postdoctoral fellows eligible, and this year we extended eligibility to colleagues at the fellow BEACON institutions. There is no cost for attendance and we used BEACON funding to fly in and lodge participants from outside the Palouse.

This year, we had a full roster of 19 attendees from 9 departments and representing 4 colleges (College of Sciences, College of Agricultural and Life Sciences, College of Natural Resources, and College of Engineering) at UI, one attendee from WSU and four from Michigan State University (Appendix V). As was the case last year, the workshop was delivered in 5 morning sessions over the course of a week. Our assessment from this year indicated that the workshop was very well received again. Attendees rated the overall effectiveness of the workshop 4.2 (out of five), and provided suggestions that will allow us to improve the 2018 edition of the workshop.

## Inland Northwest Genomics Research Symposium

The fifth annual Inland Northwest Genomics Research Symposium (INWGRS) was held in May 2017. As in the past, the symposium was a one-day event, however this year we focused on the local community of 70 evolutionary scientists and addressed the theme "Ecology, Evolution, and Human Health and Welfare." This decision was made to initiate local enthusiasm to pursue an NSF Science and Technology Center grant proposal centered on this topic and speakers (listed in Appendix VI) were invited to describe the alignment of their research programs with this theme. It was successful in coalescing the community, and IBEST is now developing a pre-proposal for this program.

**Figure 14**. BCB students in the Bioinformatics classroom.

# IX. Transitioning to a Stable Business Plan Post-COBRE

The most immediate challenge to IBEST is the imminent sun-setting of the NIH COBRE award (to PI Larry Forney) at the end of January, 2018. This is a double-edged reduction; direct cost in that award is used for salary support (ca. $123 K per year) and for investing in GRC and CRC major equipment (among other things), and the COBRE accounts for ca. 38% of our returned overhead. We have a plan in place to deal with such budget reduction that includes cost savings (revising the IBEST administrative structure) and generation of new revenues through additional funding sources.

## Administrative Structure

In order to deal with salary issues, we are restructuring the IBEST Administrative staff by deciding to not replace the Communications Coordinator, who left recently to take a hard money position in CALS. Fifty percent of this salary line was paid from the COBRE, with the requirement that 50% of the position description be allocated to serve as administrative assistant to the COBRE PI. Since those requirements sunset with the COBRE, we will not need to back-fill that 50% F.T.E. and we are reallocating the other 50% of the position description among retained administrative staff. We are working on the new organization chart and that will be presented in the 2018 Annual Report. In addition, the GRC and CRC have been increasing the revenues they generate, and this will permit a gradual shift of scientific staff salaries to internal core facility budgets. The goal is to reduce dramatically the portions of IBEST staff salaries that are paid from soft money.

## Pursuit of New Funding Sources

A number of major funding opportunities are being pursued to fill other needs that the COBRE sunset will generate.

First, Major Research Instrumentation grant proposals are planned to replace COBRE funds allocated to core facility equipment. These are generated on an as-needed basis because proposals are evaluated based on specific needs. One such proposal in the works is for large-scale SNP generation, for use in conservation and wildlife genomics, and IBEST will invest the required institutional matching funds out of our returned overhead.

Second, as described in section VIII, IBEST provides the single largest investment into BCB graduate fellowships. This can continue in the near future (out of our current F&A returns), but a long-term plan is being implemented to secure an NIH T32 training grant. One T32 proposal was submitted in 2017; it was declined the first time around, and the single biggest criticism was insufficient institutional support; we will need to address this for the resubmission in May 2018.

Third, large scale center and program project grant proposals are in progress. For example, Harmon, Foster, and Sullivan have been successful in ORED's internal competition for limited submission to the NSF Science and Technology Center program. Our proposed center will be called BioDaHWN (the Biodiversity and Human Health and Welfare Network) and will coalesce scholars at UI, Iowa State University, Ohio State University, and North Carolina A&T to address evolutionary and ecological drivers of human health and welfare. Other such large, programmatic opportunities include a new NIH program from NIGMS, RM1 - Collaborative Program Grant for Multidisciplinary Teams. IBEST has budgeted teaching release funds for PIs to prepare such large scale grant proposals.

# Appendix I. IBEST Publications

## 2016

Agashe, D., Sane, M., Phalnikar, K., Diwan, G. D., Habibullah, A., Martinez-Gomez, N. C., .Marx, C. J. (2016). Large-Effect Beneficial Synonymous Mutations Mediate Rapid and Parallel Adaptation in a Bacterium. *Molecular Biology and Evolution, 33*(6), 1542-1553. doi:10.1093/molbev/msw035

Anderson, C.J.R., Harmon, L.J. (2016). Ecological and mutation-order speciation in digital organisms. *American Naturalist.*

Ahrabi, A. F., Handa, D., Codipilly, C. N., Shah, S., Williams, J. E., McGuire, M. A., Schanler, R. J. (2016). Effects of Extended Freezer Storage on the Integrity of Human Milk. *J Pediatr*. doi:10.1016/j.jpeds.2016.06.024

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics, 17*(2), 81-92. doi:10.1038/nrg.2015.28

Baker, C. W., Miller, C. R., Thaweethai, T., Yuan, J., Baker, M. H., Joyce, P., & Weinreich, D. M. (2016). Genetically Determined Variation in Lysis Time Variance in the Bacteriophage phi X174. *G3-Genes Genomes Genetics, 6*(4), 939-955. doi:10.1534/g3.115.024075

Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J. A., McMullin, B., sWhite, R. (2016). Defining and simulating open-ended novelty: requirements, guidelines, and challenges. *Theory in Biosciences, 135*(3), 131-161. doi:10.1007/s12064-016-0229-7

Baumgaertner, B., & Holthuijzen, W. (2016). On nonepistemic values in conservation biology. *Conserv Biol*. doi:10.1111/cobi.12756

Baumgaertner, B. O., Tyson, R. T., & Krone, S. M. (2016). Opinion strength influences the spatial dynamics of opinion formation. *The Journal of Mathematical Sociology*, 1-12. doi:10.1080/0022250X.2016.1205049

Benestan, L. M., Ferchaud, A. L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J., Baums, I. B., Luikart, G. (2016). Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Mol Ecol, 25*(13), 2967-2977. doi:10.1111/mec.13647

Borgogna, T. R., Borgogna, J. L., Mielke, J. A., Brown, C. J., Top, E. M., Botts, R. T., & Cummings, D. E. (2016). High Diversity of CTX-M Extended-Spectrum beta-Lactamases in Municipal Wastewater and Urban Wetlands. *Microbial Drug Resistance, 22*(4), 312-320. doi:10.1089/mdr.2015.0197

Brown, C. J., Quates, C. J., Mirabzadeh, C. A., Miller, C. R., Wichman, H. A., Miura, T. A., & Ytreberg, F. M. (2016). New Perspectives on Ebola Virus Evolution. *PLoS One, 11*(8). doi:ARTN e0160410

Cerling, T.E., Barnette, J.E., Bowen, G.J., Chesson, L.A., Ehleringer, J., Remien, C.H., Patrick, S., Tipple, B., West, J.B. (2016). Forensic Stable Isotope Biogeochemistry. *Annual Review of Earth and Planetary Sciences,* 44(1):175-206.

Chikh-Ali, M., Alruwaili, H., Pol, D. V., & Karasev, A. V. (2016). Molecular Characterization of Recombinant Strains of Potato virus Y From Saudi Arabia. *Plant Disease, 100*(2), 292-297. doi:10.1094/Pdis-05-15-0562-Re

Douglas, S. M., Chubiz, L. M., Harcombe, W. R., Ytreberg, F. M., & Marx, C. J. (2016). Parallel Mutations Result in a Wide Range of Cooperation and Community Consequences in a Two-Species Bacterial Consortium. *PLoS One, 11*(9), e0161837. doi:10.1371/journal.pone.0161837

Epstein, B., Jones, M., Hamede, R., Hendricks, S., McCallum, H., Murchison, E. P., Storfer, A. (2016). Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nature Communications, 7*, 12684. doi:10.1038/ncomms12684

Espíndola, A., Ruffley, M., Smith, M., Carstens, B. C., Tank, D. C., Sullivan, J. (2016). Predicting cryptic diversity from phylogeographic, climatic and taxonomic data. *Proceedings of the Royal Society of London B*, 283:1-12. PMC5095379

France, M.T., Mendes-Soares, H., Forney, L.J. (2016). Genomic Comparisons of and Reveal Potential Ecological Drivers of Community Composition in the Vagina. *Applied and Environmental Microbiology*, 82:7063-7073.

Funk, W. C., Lovich, R. E., Hohenlohe, P. A., Hofman, C. A., Morrison, S. A., Sillett, T. S., Andelt, W. F. (2016). Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Mol Ecol, 25*(10), 2176-2194. doi:10.1111/mec.13605

G, H. (2016). Teaching Philosophy by Designing a Wikipedia Page. In J. Oxley & R. Ilea (Eds.), *Experiential Learning in Philosophy*: Taylor & Francis.

Garrett, A. M., Tadenev, A. L., Hammond, Y. T., Fuerst, P. G., & Burgess, R. W. (2016). Replacing the PDZ-interacting C-termini of DSCAM and DSCAML1 with epitope tags causes different phenotypic severity in different cell populations. *Elife, 5*. doi:10.7554/eLife.16144

Gerritsen, A. T., New, D. D., Robison, B. D., Rashed, A., Hohenlohe, P., Forney, L., Settles, M. L. (2016). Full Mitochondrial Genome Sequence of the Sugar Beet Wireworm *Limonius californicus* (Coleoptera: Elateridae), a Common Agricultural Pest. *Genome Announc, 4*(1). doi:10.1128/genomeA.01628-15

Gillies, K., Krone, S. M., Nagler, J. J., & Schultz, I. R. (2016). A Computational Model of the Rainbow Trout Hypothalamus-Pituitary-Ovary-Liver Axis. *PLoS Comput Biol, 12*(4), e1004874. doi:10.1371/journal.pcbi.1004874

Greene, C. S., Foster, J. A., Stanton, B. A., Hogan, D. A., & Bromberg, Y. (2016). Computational Approaches to Study Microbes and Microbiomes. *Pac Symp Biocomput, 21*, 557-567.

Hagey, T. J., Puthoff, J. B., Crandell, K. E., Autumn, K., & Harmon, L. J. (2016). Modeling Observed Animal Performance using the Weibull Distribution. *Journal of Experimental Biology, 219*(11), 1603-1607. doi:10.1242/jeb.129940

Harcombe, W. R., Betts, A., Shapiro, J. W., & Marx, C. J. (2016). Adding Biotic Complexity Alters the Metabolic Benefits of Mutualism. *Evolution, 70*(8). doi:10.1111/evo.12973

Hickey, R., & Forney, L. (2016). *Gardnerella vaginalis* Does Not Always Aause Bacterial Vaginosis. *Journal of Infectious Diseases*.

Hunter, S. S., Settles, M. L., New, D. D., Parent, C. E., & Gerritsen, A. T. (2016). Mitochondrial Genome Sequence of the Galapagos Endemic Land Snail *Naesiotus nux*. *Genome Announc, 4*(1). doi:10.1128/genomeA.01362-15

Jayaram, A., Witkin, S. S., Zhou, X., Brown, C. J., Rey, G. E., Linhares, I. M., Forney, L. J. (2016). The Bacterial Microbiome in Paired Vaginal and Vestibular Samples from Women with Vulvar Vestibulitis Syndrome. *Pathog Dis, 72*(3), 161-166. doi:10.1111/2049-632X.12197

Jensen, P. K., Wujcik, C. E., McGuire, M. K., & McGuire, M. A. (2016a). Validation of reliable and selective methods for direct determination of glyphosate and aminomethylphosphonic acid in milk and urine using LC-MS/MS. *J Environ Sci Health B, 51*(4), 254-259. doi:10.1080/03601234.2015.1120619

Jensen, P. K., Wujcik, C. E., McGuire, M. K., & McGuire, M. A. (2016b). Validation of reliable and selective methods for direct determination of glyphosate and aminomethylphosphonic acid in milk and urine using LC-MS/MS. *Journal of Environmental Science and Health Part B-Pesticides Food Contaminants and Agricultural Wastes, 51*(4), 254-259. doi:10.1080/03601234.2015.1120619

Kim, M. S., Hohenlohe, P. A., Kim, K. H., Seo, S. T., & Klopfenstein, N. B. (2016). Genetic diversity and population structure of *Raffaelea quercus-mongolicae*, a fungus associated with oak mortality in South Korea. *Forest Pathology, 46*(2), 164-167. doi:10.1111/efp.12263

Li, H., Xu, Z., Yang, S., Li, X., Top, E. M., Wang, R., & Jiang, Y. (2016). Responses of Soil Bacterial Communities to Nitrogen Deposition and Precipitation Increment Are Closely Linked with Aboveground Community Variation. *Microb Ecol, 71*(4), 974-989. doi:10.1007/s00248-016-0730-z

Li, S., Mitchell, J., Briggs, D. J., Young, J. K., Long, S. S., & Fuerst, P. G. (2016). Morphological Diversity of the Rod Spherule: A Study of Serially Reconstructed Electron Micrographs. *PLoS One, 11*(3), e0150024. doi:10.1371/journal.pone.0150024

Li, S., Woodfin, M., Long, S. S., & Fuerst, P. G. (2016). IPLaminator: an ImageJ plugin for automated binning and quantification of retinal lamination. *BMC Bioinformatics, 17*, 36. doi:10.1186/s12859-016-0876-1

Li, X., Wang, Y., Brown, C. J., Yao, F., Jiang, Y., Top, E. M., & Li, H. (2016). Diversification of broad host range plasmids correlates with the presence of antibiotic resistance genes. *FEMS Microbiology Ecology, 92*(1). doi:10.1093/femsec/fiv151

Loftie-Eaton, W., Yano, H., Burleigh, S., Simmons, R. S., Hughes, J. M., Rogers, L. M., . . . Top, E. M. (2016). Evolutionary Paths That Expand Plasmid Host-Range: Implications for Spread of Antibiotic Resistance. *Molecular Biology and Evolution, 33*(4), 885-897. doi:10.1093/molbev/msv339

Magalhaes, I. S., D'Agostino, D., Hohenlohe, P. A., & MacColl, A. D. (2016). The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Mol Ecol, 25*(17), 4319-4336. doi:10.1111/mec.13746

Marx, H. E., Giblin, D. E., Dunwiddie, P. W., & Tank, D. C. (2016). Deconstructing Darwin's Naturalization Conundrum in the San Juan Islands using community phylogenetics and functional traits. *Diversity and Distributions, 22*(3), 318-331. doi:10.1111/ddi.12401

Mendes-Soares, H., Suzuki, H., Hickey, R. J., & Forney, L. J. (2016). Comparative Functional Genomics of Lactobacillus spp. Reveals Possible Mechanisms for Specialization of Vaginal Lactobacilli to Their Environment. *Journal of Bacteriology, 196*(7), 1458-1470. doi:10.1128/Jb.01439-13

McGuire, M. A. M., Price,W.J, Shafii, B., Carrothers, J.M. , Lackey, K.A., Goldstein, D.A., Jensen, P.K., & Vicini, J.L. (2016). Glyphosate and aminomethylphosphonic acid are not detectable in human milk. *The American journal of clinical nutrition, 103*, 1285-1290.

Michener, J. K., Vuilleumier, S., Bringel, F., & Marx, C. J. (2016). Transfer of a Catabolic Pathway for Chloromethane in Methylobacterium Strains Highlights Different Limitations for Growth with Chloromethane or with Dichloromethane. *Frontiers in Microbiology, 7*. doi:ARTN 1116

Miler, C. R., Johnson, E. L., Burke, A. Z., Martin, K. P., Miura, T. A., Wichma, H. A., . . . Ytreberg, F. M. (2016). Initiating a watch list for Ebola virus antibody escape mutations. *Peerj, 4*. doi:ARTN e1674

Miller, C. R., Nagel, A. C., Scott, L., Settles, M., Joyce, P., & Wichman, H. A. (2016). Love the one you're with: replicate viral adaptations converge on the same phenotypic change. *Peerj, 4*. doi:ARTN e2227

Murdoch, B. M., & Murdoch, G. K. (2016). Genetics of Prion Disease in Cattle. *Bioinformatics and Biology Insights, 9*, 1-10. doi:10.4137/BBi.s29678

Murdoch, B.M., Murdoch, G.K., Greenwood, S., McKay, S. (2016). Nutritional influence on epigenetic marks and effects on livestock production. *Frontiers in Genetics*, 7:182.

Myrvold, K.M., Kennedy, B.P. (2016). Juvenile steelhead movements in relation to stream habitat, population density and body size: consequences on individual growth rates. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(10): 1520-1529.

Nardin, L.G., Miller, C.R., Ridenhour, B.J., Krone, S.M., Joyce, P., Baumgaertner, B. (2016). Planning horizon affects prophylactic decision-making and epidemic dynamics. *Peer J*, 4:e2678.

Nayak, D. D., Agashe, D., Lee, M. C., & Marx, C. J. (2016). Selection Maintains Apparently Degenerate Metabolic Pathways due to Tradeoffs in Using Methylamine for Carbon versus Nitrogen. *Current Biology, 26*(11), 1416-1426. doi:10.1016/j.cub.2016.04.029

Nuismer, S. L., & Dybdahl, M. F. (2016). Quantifying the coevolutionary potential of multistep immune defenses. *Evolution*, 70, 282-295. doi:10.1111/evo.12863

Nunn, K.L., Forney, L.J. (2016). Unraveling the Dynamics of the Human Vaginal Microbiome, 89:331-337.

Paff, M. L., Nuismer, S. L., Ellington, A., Molineux, I. J., & Bull, J. J. (2016). Virus wars: using one us to block the spread of another. *Peerj, 4*. doi:ARTN e2166

Ravel, J., Brotman, R. M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D. W., . . . Forney, L. J. (2016). Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome,* 1, 29. doi:10.1186/2049-2618-1-29

Rezamand, P., Hatch, B. P., Carnahan, K. G., & McGuire, M. A. (2016). Effects of alpha-linolenic acid-enriched diets on gene expression of key inflammatory mediators in immune and milk cells obtained from Holstein dairy cows. *Journal of Dairy Research, 83*(1), 20-27. doi:10.1017/S0022029915000709

Ridenhour, B. J., and E.M. Top. (2016). Plasmid driven evolution of Bacteria.

Robison, B. D., Soule, T., Streett, D., Mirabzadeh, C. A., & Wood, N. (2016). *Implementing evolution in video games*. Paper presented at the Games Learning Society, Madison, WI.

Roe, S., Baumgaertner, B. (2016). Extended Mechanistic Explanations: Expanding the current mechanistic conception to include more complex biological systems. *Journal for General Philosophy of Science*, 48: 517-534.

Rondon, S. I., Roster, M. S., Hamlin, L. L., Green, K. J., Karasev, A. V., & Crosslin, J. M. (2016). Characterization of Beet curly top virus Strains Circulating in Beet Leafhoppers (Hemiptera: Cicadellidae) in Northeastern Oregon. *Plant Disease, 100*(8), 1586-1590. doi:10.1094/Pdis-10-15-1189-Re

Rutledge, L. Y., Devillard, S., Hohenlohe, P. A., & White, B. N. (2016). Considering all the evidence: a reply to Sefc and Koblmüller (2016). *Biology Letters, 12*(2). doi:10.1098/rsbl.2015.1009

Shen, J., Song, N., Williams, C. J., Brown, C. J., Yan, Z., Xu, C., & Forney, L. J. (2016). Effects of low dose estrogen therapy on the vaginal microbiomes of women with atrophic vaginitis. *Scientific Reports, 6*. doi:ARTN 24380

Shen, J., Song, N., Williams, C. J., Brown, C. J., Yan, Z., Xu, C., & Forney, L. J. (2016). Effects of low dose estrogen therapy on the vaginal microbiomes of women with atrophic vaginitis. *Scientific Reports, 6*, 24380. doi:10.1038/srep24380

Simmons, A. B., Bloomsburg, S. J., Billingslea, S. A., Merrill, M. M., Li, S., Thomas, M. W., & Fuerst, P. G. (2016). Pou4f2 knock-in Cre mouse: A multifaceted genetic tool for vision researchers. *Mol Vis, 22*, 705-717.

Singer, M. L., Oreschak, K., Rhinehart, Z., & Robison, B. D. (2016). Anxiolytic effects of fluoxetine and nicotine exposure on exploratory behavior in zebrafish. *PeerJ Preprints, 4*, e1718v1712.

Soule, T., Robison, B. D., & Heckendorn, R. B. (2016). *Co-evolution of Sensor Morphology and Behavior*. Paper presented at the Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion, Denver, Colorado, USA.

Stockmann, C., Ampofo, K., Pavia, A.T., Blaschke, A.J., Mason, E.O., Presson, A.P., Forney, L.J., Byington, C.L. (2016). Clinical and Epidemiological Evidence of the Red Queen Hypothesis in Pneumococcal Serotype Dynamics. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 63:619-626.

Sukeena, J.M., Galicia, C.A., Wilson, J.D., McGinn, T., Boughmann, J.W., Robison, B.D., Postlethwait, J.H., Braasch, I., Stenkamp, D.L., Fuerst, P. (2016). Characterization and Evolution of the Spotted Gar Retina. *Journal of experimental zoology.* Part B, Molecular and developmental evolution, 326:403-421.

Taylor, T.N., Myrvold, K.M, Kennedy, B.P. (2016) Food habits of sculpin spp. in small Idaho streams: no evidence of predation on newly emerged Steelhead alevins. *Northwest Science*, 90(4): 484-490.

Trumbo, D. R., Epstein, B., Hohenlohe, P. A., Alford, R. A., Schwarzkopf, L., & Storfer, A. (2016). Mixed population genomics support for the central marginal hypothesis across the invasive range of the cane toad (*Rhinella marina*) in Australia. *Mol Ecol, 25*(17), 4161-4176. doi:10.1111/mec.13754

Uribe-Convers, S., Settles, M. L., & Tank, D. C. (2016). A Phylogenomic Approach Based on PCR Target Enrichment and High Throughput Sequencing: Resolving the Diversity within the South American Species of *Bartsia* L. (Orobanchaceae). *PLoS One, 11*(2), e0148203. doi:10.1371/journal.pone.0148203

Uribe-Convers, S., & Tank, D. C. (2016). Phylogenetic Revision of the Genus *Bartsia* (Orobanchaceae): Disjunct Distributions Correlate to Independent Lineages. *Systematic Botany, 41*(3). doi:10.1600/036364416X692299

Villasante, A., Powell, M. S., Murdoch, G. K., Overturf, K., Cain, K., Wacyk, J., & Hardy, R. W. (2016). Effect of anthocyanidins on myogenic differentiation in induced and non-induced primary myoblasts from rainbow trout (*Oncorhynchus mykiss*). *Comp Biochem Physiol B Biochem Mol Biol, 196-197*, 102-108. doi:10.1016/j.cbpb.2016.03.004

Yano, H., Wegrzyn, K., Loftie-Eaton, W., Johnson, J., Deckert, G. E., Rogers, L. M., . . . Top, E. M. (2016). Evolved plasmid-host interactions reduce plasmid interference cost. *Mol Microbiol, 101*(5), 743-756. doi:10.1111/mmi.13407

## 2017

Acharya, S., Stark, T.D., Oh, S.T., Jeon, S., Pak, S.C., Kim, M., Hur, J., Matsutomo, T., Hofmann, T., Hill, R.A., Balemba, O.B. (2017). Manniflavanone Protects Proliferating Skeletal Muscle Cells against Oxidative Stress and Stimulates Myotube Formation. *J Agric Food Chem*, 65:3636-3646.

Basinski, A.J., Varrelman, T.J., Smithson, M.W., May, R.H., Remien, C.H., Nuismer, S.L. (in press). Evaluating the Promise of Recombinant Transmissible Vaccines. *Vaccine*.

Baumgaertner, B. (in press). Models of Opinion Dynamics and Mill-style Arguments for Opinion Diversity. *Historical Social Research.*

Baumgaertner, B., Holthuijzen, W. (2017). Non-epistemic Values and Concerns about Evolutionary Mindsets in Conservation Policy: A Response to Kareiva and Fuller's `Beyond Resilience'. *Global Policy*, 10.1111, 1759-5899.12469.

Botts, R.T., Apffel, B., Walters, C., Davidson, K., Echols, R., Geiger, M., Guzman, V., Haase, V., Montana, M., La Chat, C., Mielke, J., Mullen, K., Virtue, C., Brown, C.J., Top, E., Cummings, D. (2017). Characterization of Four Multidrug Resistance Plasmids Captured from the Sediments of an Urban Coastal Wetland. *Frontiers in Microbiology*, epub ahead of print, doi: 10.3389/fmicb.2017.01922.

Boucher, F., Démery, V., Conti, E., Harmon, L., Uyeda, J. (2017). A General Model For Estimating Macroevolutionary Landscapes. *Systematic Biology*. epub ahead of print, doi.org/10.1093/sysbio/syx075

Bruce, F.M., Brown, S., Smith, J.N., Fuerst, P., Erskine, L. (2017). DSCAM promotes axon fasciculation and growth in the developing optic pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 114:1702-1707.

Bull, J.J., Smithson, M.W., Nuismer, S.L. (in press). Transmissible Viral Vaccines. *Trends in Microbiology*.

Calderon, D., Bhaskar, A., Knowles, D., Golan, D., Raj, T., Fu, A., Pritchard, J. (2017). Inferring Relevant Cell Types for Complex Traits by Using Single-Cell Gene Expression. *The American Journal of Human Genetics*, 101:686-699.

Catchen, J., Hohenlohe, P.A., Bernatchez, L., Funk, W.C., Andrews, K.R., Allendorf, F. (2017). Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Molecular Ecology Resources*, 17:362-365.

Chapalamadugu, K., VandeVoort, C., Settles, M., Robison, B., Murdoch, G.K. (2017). Maternal Bisphenol A Exposure Impacts the Fetal Heart Transcriptome. *PLoS ONE*, (9).

Cleary, K., Waits, L.P., Hohenlohe, P.A. (2017). Development and characterization of fourteen novel microsatellite markers for the chestnut short-tailed fruit bat (), and cross-amplification to related species. *PeerJ*, 4.

Cornejo, O., Hickey, R., Suzuki, H., Forney, L. (in press). The effect of vaginal microbial communities on colonization by Staphylococcus aureus with the gene for toxic shock syndrome toxin 1 (TSST-1). *Pathogens and Disease*.

Damase, T., Miura,T.A., Parent, C., Allen, P. (in press). Application of the Open qPCR instrument for in vitro selection of a DNA aptamer against Drosophila C virus (DCV). *ACS Combinatorial Science*.

Damase, T., Spencer, A., Samuel, B., Allen, P.B. (2017). Biomimetic Molecular Signaling using DNA Walkers on Microparticles. *Scientific Reports*, (7).

Douglas, S., Chubiz, L., Harcombe, W., Ytreberg, F., Marx, C. (2017). Parallel Mutations Result in a Wide Range of Cooperation and Community Consequences in a Two-Species Bacterial Consortium. *PloS One* (11).

Duchen, P., Leuenberger, C., Szilágyi, S., Harmon, L., Eastman, J., Schweizer, M., Wegmann, D. (2017). Inference of Evolutionary Jumps in Large Phylogenies using Lévy Processes. *Systematic Biology*, 66, 950-963.

Foster, J.A. (2017). Taking "biology" just seriously enough: Commentary on "On the mapping of genotype to phenotype in evolutionary algorithms" by Peter A. Whigham, Grant Dick, and James Maclaurin. *Genetic Programming and Evolvable Machines*, 1-4.

Fu, A.Q., Pachter, L. (2017). Estimating intrinsic and extrinsic noise from single-cell gene expression measurements. *Statistical Applications in Genetics and Molecular Biology*, 15:447-471.

Gomulkiewicz, R., Krone, S., Remien, C. (2017). Evolution and the duration of a doomed population. *Evolutionary Applications*, 10:471-484.

Gosselin, E.N., Lonsinger, R., Waits, L. (online early). Comparing Morphological and Molecular Diet Analyses and Fecal DNA Sampling Protocols for a Terrestrial Carnivore. *Wildlife Society Bulletin*.

Green, K.J., Brown, C., Gray, S., Karasev, A. (in press). Genetic diversity of Potato virus Y (PVY): Sequence analyses reveal ten novel PVY recombinant structures. *Archives of Virology*.

Green, K.J., Brown, C.J., Gray, S., Karasev, A. (2017). Phylogenetic study of recombinant strains of Potato virus Y. *Virology*, 507:40-52.

Hagey, T., Harte, S., Vickers, M., Harmon, L., Schwarzkopf, L. (2017). There's more than one way to climb a tree: Limb length and microhabitat use in lizards with toe pads. *PLoS ONE*, 12.

Hagey, T., Uyeda, J., Crandell, C., Cheney, J., Autumn, K., Harmon, L. (2017). Tempo and mode of performance evolution across multiple independent origins of adhesive toe pads in lizards. *Evolution*, 71. 2344-2358.

Hand, B.K., Devlin, S., Mulfeld, C.C., Frissell, C., Kennedy, B.P., Crabtree, R.L., Flint, C., McKee, A., Luikart, G. (in press). A Socio-Ecological Systems approach for Riverscape Sustainability in the Columbia River Basin. *Frontiers in Ecology and Evolution*.

Harmon, L.J. (2017). Evolution: Contingent Predictability in Mammalian Evolution. *Current biology*, CB, 27.

Harrington, K., Rueden, C., Eliceiri, K. (2017). FunlmageJ: a Lisp framework for scientific image processing. *In Bioinformatics*, btx710.

Hegg, J.C., Middleton, J., Robertson, B.L., Kennedy, B.P. (in press). The sound of migration: Exploring data sonification as a way of analyzing multivariate salmon movement datasets. *Heliyon*.

Hendricks, S.A., Epstein, B., Schonfeld, B., Wiench, C., Hamede, R., Jones, M., Storfer, A., Hohenlohe, P.A. (2017). Conservation implications of limited genetic diversity and population structure in Tasmanian devils (*Sarcophilus harrisii*). *Conservation Genetics*, 18:997-982.

Hohenlohe, P., Rutledge, L., Waits, L., Andrews, K., Adams, J., Hinton, J., Nowak, R., Patterson, B., Wydeven, A., Wilson, P., White, B. (2017). Comment on "Whole-genome sequence analysis shows two endemic species of North American wolf are admixtures of the coyote and gray wolf". *Science Advances*, 3: e1602250.

Hohenlohe, P.A., Rutledge, L., Waits, L., Andrews, K., Adams, J., Hinton, J., Nowak, R.M., Patterson, B.R., Wydeven, A.P., White, B.N., Wilson, P. (2017). Whole-genome data do not support recent hybrid origins of red wolves and eastern wolves (response). *Science Advances*, (3) e1602250.

Landis, J.B., Soltis, D., Zheng, L., Marx, H.E., Barker, M., Tank, D., Soltis, P. (in press). Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany*.

Latvis, M., Mortimer, S., Morales-Briones, D., Torpey, S., Uribe-Convers, S., Jacobs, S., Mathews, S., Tank, D. (2017). Primers for *Castilleja* and their utility across Orobanchaceae: I. Chloroplast primers. *Applications in Plant Sciences*.

Leizer, J., Nasioudis, D., Forney, L.J., Schneider, G.M., Gliniewicz, K., Boester, A., Witkin, S.S. (2017). Properties of epithelial cells and vaginal secretions in pregnant women when *Lactobacillus crispatus* or *Lactobacillus iners* dominate the vaginal microbiome. *Reprod Sci*, Jan 1:193371911769858 [Epub ahead of print].

Loftie-Eaton, W., Bashford, K., Quinn, H.,Dong, K.,Millstein, J., Hunter, S., Thomason, M., Merrikh, M., Ponciano, J., Top, E. (2017). Compensatory mutations improve general permissiveness to antibiotic resistance plasmids. *Nature: Ecology and Evolution,* 1:1354-1363.

MacPherson, A., Nuismer, S.L. (2017). The probability of parallel genetic evolution from standing genetic variation. *Journal of Evolutionary Biology*, 30:326-337.

MacPherson, A., Otto, S.P., Nuismer, S.L. (in press). Keeping pace with the Red Queen: identifying the genetic basis of susceptibility to infectious disease. *Genetics*.

Marx, H.E., Dentant, C., Renaud, J., Delunel, R., Tank, D., Lavergne, S. (2017). Riders in the sky (islands): Using a mega-phylogenetic approach to understand plant species distribution and coexistence at the altitudinal limits of angiosperm plant life. *Journal of Biogeography*.

McGuire, M.K., Meehan, C., McGuire, M.A., Williams, J., Foster, J., Sellen, D., Kamau-Mbuthia, E., Kamundia, E., Moore, S., Kvist, L., Otoo, G., Brooker, S., Price, J., Shafii, B., Placek, C., Lackey, K., Robertson, B., Manzano, S., Ruiz, L., Rodriguez, J., Pareja, R., Bode, L. (2017). What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically. *American Journal of Clinical Nutrition*. 105: 1086-1110.

Miller, C.R., Van Leuven, J., Wichman, H., Joyce, P. (in press). Selecting among three basic fitness landscape models: additive, multiplicative and stickbreaking. *Theoretical Population Biology*.

Miller, E.T., Wagner, S., Harmon, L., Ricklefs, R. (2017). Radiating despite a Lack of Character: Ecological Divergence among Closely Related, Morphologically Similar Honeyeaters (Aves: Meliphagidae) Co-occurring in Arid Australian Environments. *The American Naturalist*, 189. E14-E30.

Morales-Briones, D.F., Romoleroux, K. Kolár, F., Tank, D. (in press). Phylogeny and evolution of the neotropical radiation of Lachemilla (Rosaceae): uncovering a history of reticulate evolution and implications for infrageneric classification. *Systematic Botany*.

Myrvold, K.M., Kennedy, B.P. (in press). Shifts in Great Blue Heron foraging locations following nest site usurpation: Implications for salmonids. *American Midland Naturalist*.

Myrvold, K.M., Kennedy, B.P. (in press). Warming temperatures exacerbate the impacts of density dependence in steelhead. *Canadian Journal of Fish and Aquatic Sciences.*

Myrvold, K.M., Kennedy, B.P. (in press). Indexing of salmonid abundance in small streams using reduced effort electrofishing. *Northwest Science*.

Myrvold, K.M., Kennedy, B.P. (in press). Size and growth relationships in juvenile steelhead: the advantage of large relative size diminishes with increasing water temperatures. *Environmental Biology of Fishes*.

Nasioudis, D., Forney, L., Schneider, G.M., Gliniewicz, K., France, M., Boester, A., Sawai, M., Scholl, J., Witkin, S. (2017). Influence of Pregnancy History on the Vaginal Microbiome of Pregnant Women in their First Trimester. *Scientific Reports*. 10201.

Nasioudis, D., Forney, L.J., Schneider, G.M., Gliniewicz, K., France, M.T., Boester, A., Sawai, M., Scholl, J., Witkin, S.S. (2017). The composition of the vaginal microbiome in first trimester pregnant women influences the level of autophagy and stress in vaginal epithelial cells. *J Reprod Immunol*, 123:35-39.

Nuismer, S., Jenkins, C., Dybdahl, M. (2017). Identifying coevolving loci using interspecific genetic correlations. *Ecology and Evolution*, 7:6894-6903.

Nuismer, S.L. (2017). Rethinking conventional wisdom: are locally adapted parasites ahead in the coevolutionary race? *The American Naturalist*, 190:584-593.

Nyavor, Y.E., Balemba, O.B. (2017). Diet-induced dysmotility and neuropathy in the gut precedes endotoxaemia and metabolic syndrome: the chicken and the egg revisited. *J Physiol*, 595:1441-1420.

Plowright, R.K., Manlove, K., Besser, T., Paez, D., Andrews, K., Matthews, P., Waits, L., Hudson, P., Cassirer, E.F. (in press). Age-structured variation in infectious period explains epidemiological features of pneumonia in bighorn sheep. *Ecology Letters*.

Ridenhour, B.J., Brooker, S.L., Williams, J.E., Van Leuven, J.T., Miller, A.W., Dearing, M.D., Remien, C.H. (2017). Modeling Time-Series Data From Microbial Communities. *ISMEJ*, 11:2526-2537.

Ridenhour, B.J., Metzger, G., France, M., Gliniewicz, K., Millstein, J., Forney, L., Top, E. (2017). Persistence of antibiotic resistance plasmids in bacterial biofilms. *Evolutionary Applications*, 10:640-647.

Rosenblum, E., Parent, C., Diepeveen, E., Noss, C., Bi, K. (2017). Convergent Phenotypic Evolution despite Contrasting Demographic Histories in the Fauna of White Sands. *The American Naturalist*, 190:S44-S56.

Ross, C., Austad, S., Brasky, K., Brown, C., Forney, L., Gelfond, J., Lanford, R., Richardson, A., Tardif, S. (in press). The development of a specific pathogen free (SPF) barrier colony of marmosets (*Callithrix jacchus*) for aging research. *Aging*.

Ruffley, M., Smith, M.L., Espíndola, A., Carstens, B., Sullivan, J., Tank, D. (in press). Combining allele frequency and tree-based approaches improves phylogeographic inference from natural history collections. *Molecular Ecology*.

Ruiz, L., Espinosa-Martos, I., García-Carral, C., Manzano, S., McGuire, M. K., Meehan, C. L., Williams, J., Foster, J.A. (2017). What's Normal? Immune Profiling of Human Milk from Healthy Women Living in Different Geographical and Socioeconomic Settings. *Frontiers in Immunology*, 8. doi.org/10.3389/fimmu.2017.00696

Sarver, B. A., Demboski, J., Good, J. M., Forshee, N., Hunter, S.L., Sullivan, J. (2017). Comparative mitochondrial phylogenomic assessment of introgression among several species of chipmunks (*Tamias*). *Genome Biology and Evolution*. 9:7-19. doi: 10.1093/gbe/evw254.

Shine, C.L., Robbins, C.T., Nelson, O.L. and McGowan, C.P. (2017). Grizzly bear (*Ursus arctos horribilis*) locomotion: forelimb joint mechanics across speed in the sagittal and frontal planes.
*Journal of Experimental Biology*, 220:1322-1329.

Simmons, A.B., Bloomsburg, S.J., Billingslea, S., Merrill, M., Li, S., Thomas, M.W., Fuerst, P. (2017). Knock-in Cre mouse: A multifaceted genetic tool for vision researchers. *Molecular Vision*, 22:705-717.

Smith, M., Ruffley, M., Espíndola, M., Tank, D., Sullivan, J., Carstens, B. (2017). Demographic model selection using random forests and the site frequency spectrum. *Molecular Ecology*, 26:4562-4573.

Soule, T., Heck, S., Haynes, T., Wood, N., Robison, B. (2017). Darwin's Demons: Does Evolution Improve the Game? *European Conference on the Applications of Evolutionary Computation*, 435-451.

Stalder, T., Rogers, L., Renfrow, C., Yano, H., Smith, Z., Top, E. (2017). Emerging patterns of plasmid-host coevolution that stabilize antibiotic resistance. *Scientific Reports*, (7).

Thomas, C., Thomson, N., Cerdeño-Tárraga, A., Brown, C., Top, E., Frost, L. (2017). Annotation of plasmid genes. *Plasmid*, 91:61-67.

Uyeda, J., Harmon, L., Blank, C. (2017). A Comprehensive Study of Cyanobacterial Morphological and Ecological Evolutionary Dynamics through Deep Geologic Time. *PLoS ONE*, 11.

Van Leuven, J.T., Ridenhour, B.J., Gonzalez, A.J., Miller, C.R., Miura, T. (2017). Lung epithelial cells have virus-specic and shared gene expression responses to infection by diverse respiratory viruses. *PLoS One*, 12: e0178408.

Weber, M., Wagner, C., Best, R., Harmon, L., Matthews, B. (2017). Evolution in a Community Context: On Integrating Ecological Interactions and Macroevolution. *Trends in Ecology & Evolution*, 32. 291-304.

Wells, K., Hamede, R.K., Kerlin, D.H., Storfer, A., Hohenlohe, P.A., Jones, M.E., McCallum, H.I. (2017). Infection of the fittest: devil facial tumor disease has greatest effect on individuals with highest reproductive output. *Ecology Letters*, 20:770-778.

Zhbannikov, I., Hunter, S.S., Foster, J.A., Settles, M. (2017). SeqyClean: a pipeline for high- throughput sequence data preprocessing. *ACM Conference on Bioinformatics and Computational Biology*.

Zuehlke, A., Reidy, M., Lin, C., LaPointe, G., Alsomairy, S., Lee, D., Rivera-Márquez, G., Beebe, K., Prince, T., Lee, S., Trepel, J., Xu, W., Johnson, J., Masison, D., Neckers, L. (2017). An Hsp90 co-chaperone protein in yeast is functionally replaced by site-specific posttranslational modification in humans. *Nature Communications*, 8:15328.

# Appendix II. Proposal Submitted Through IBEST

18 proposals were submitted through IBEST between 09/01/2016 and 12/31/2017. These were submitted to 11 funding bodies: 4 federal (NSF, NIH, USDA, BPA), 3 state universities (as sub-awards), one Idaho SBOE, and 3 non-government organizations.

| | PI | COLLEGE/DEPT | SPONSOR | REQUEST |
|---|---|---|---|---|
| 1 | Parent, Christine* | COS/Biology | National Science Foundation | 1,093,877 |
| 2 | Top, Eva M.* | COS/Biology | NIFA/USDA | 500,000 |
| 3 | Tank, David C. | COS/Biology | National Science Foundation | 426,924 |
| 4 | Marx, Christopher | COS/Biology | Michigan State University | 305,385 |
| 5 | Tank, David C. | COS/Biology | Soc. of Systematic Biologists | 1,300 |
| 6 | Tank, David C. | COS/Biology | Soc. of Systematic Biologists | 1,300 |
| 7 | Hohenlohe, Paul | COS/Biology | National Science Foundation | 203,666 |
| 8 | Forney, Larry J. | COS/Biology | The University of Michigan | 92,817 |
| 9 | Kennedy, Brian P. | CNR/Fish & Wildlife | Bonneville Power Administration | 129,541 |
| 10 | Foster, James A. | COS/Biology | Dept. Health and Human Services | 1,073,360 |
| 11 | Robison, Barrie D. | COS/Biology | Idaho State Board of Education | 67,600 |
| 12 | Hohenlohe, Paul | COS/Biology | Washington State University | 474,844 |
| 13 | Top, Eva M. | COS/Biology | J. Craig Venter Institute | 92,676 |
| 14 | Harmon, Luke J. | COS/Biology | National Science Foundation | 641,442 |
| 15 | Uyeda, Josef C. | COS/Biology | National Science Foundation | 136,657 |
| 16 | Top, Eva M. | COS/Biology | Dept. Health and Human Services | 1,948,581 |
| 17 | Marx, Christopher | COS/Biology | National Science Foundation | 377,419 |
| 18 | Hunter, Samuel S. | ORED/IBEST | Idaho Wheat Commission | 26,463 |
| | | | | **7,593,852** |

*Notification of intention to fund has been received by Parent and Top, but funds have not yet been awarded.

# Appendix III. Funds Awarded To Grants Submitted Through IBEST

32 new awards were made between 09/01/2016 and 12/31/2017 on grants submitted through IBEST. These were funded by to 11 funding bodies: 3 federal (NSF, NIH, BPA), 5 state universities as sub-awards or BEACON awards (those from MSU), one Idaho SBOE, and 2 non-government organizations.

| | PI | COL/DEPT | SPONSOR | AWARD |
|---|---|---|---|---|
| 1 | Kennedy, Brian P. | CNR/FWR | Bonneville Power Administration | $129,522 |
| 2 | Forney, Larry J. | COS/Bio Sci | University of Maryland | $54,565 |
| 3 | Foster, James A. | COS/Bio Sci | Michigan State University | $22,788 |
| 4 | Foster, James A. | COS/Bio Sci | Michigan State University | $27,965 |
| 5 | Foster, James A. | COS/Bio Sci | Michigan State University | $19,776 |
| 6 | Foster, James A. | COS/Bio Sci | Michigan State University | $17,243 |
| 7 | Hohenlohe, Paul | COS/Bio Sci | National Science Foundation | $699,389 |
| 8 | Hohenlohe, Paul | COS/Bio Sci | Colorado State University | $110,604 |
| 9 | Hohenlohe, Paul | COS/Bio Sci | Washington State University | $24,098 |
| 10 | Robison, Barrie D. | COS/Bio Sci | Idaho State Board of Education | $67,600 |
| 11 | Forney, Larry J. | COS/Bio Sci | DHHS-NIH | $94,020 |
| 12 | Parent, Christine | COS/Bio Sci | National Science Foundation | $12,000 |
| 13 | Foster, James A. | COS/Bio Sci | Michigan State University | $75,248 |
| 14 | Forney, Larry J. | COS/Bio Sci | University of Maryland | $63,731 |
| 15 | Forney, Larry J. | COS/Bio Sci | University of Maryland | $28,502 |
| 16 | Wichman, Holly A. | COS/Bio Sci | Michigan State University | $24,630 |
| 17 | Heckendorn, Robert | COS/Comp Sci | Michigan State University | $25,210 |
| 18 | Stenkamp, Deborah L. | COS/Bio Sci | Michigan State University | $11,683 |
| 19 | Top, Eva M. | COS/Bio Sci | Michigan State University | $18,209 |
| 20 | Wichman, Holly A. | COS/Bio Sci | Michigan State University | $37,396 |
| 21 | Wichman, Holly A. | COS/Bio Sci | Michigan State University | $38,152 |
| 22 | Wichman, Holly A. | COS/Bio Sci | Michigan State University | $47,933 |
| 23 | Marx, Christopher | COS/Bio Sci | Michigan State University | $12,101 |
| 24 | Hunter, Samuel S. | ORED/IBEST | Idaho Wheat Commission | $37,578 |
| 25 | Brown, Celeste J. | COS/Bio Sci | Texas Biomed. Research Institute | $93,098 |
| 26 | Stenkamp, Deborah L. | COS/Bio Sci | National Science Foundation | $409,998 |
| 27 | Nuismer, Scott L. | COS/Bio Sci | National Science Foundation | $83,533 |
| 28 | Forney, Larry J. | COS/Bio Sci | NIH COBRE Administrative Core | $296,254 |
| 29 | Forney, Larry J. | COS/Bio Sci | NIH COBRE Computational Core | $138,563 |
| 30 | Forney, Larry J. | COS/Bio Sci | NIH COBRE Genomics Core | $228,313 |
| 31 | Forney, Larry J. | COS/Bio Sci | NIH COBRE Pilot Project | $77,441 |
| 32 | Forney, Larry J. | COS/Bio Sci | NIH COBRE Pilot Project | $105,608 |
| | | | | **$3,132,749** |

# Appendix IV. BCB Seminar Series

# Seminar Series
## Fall 2017

***Co-Sponsored by IBEST and CMCI***
***Seminars are Thursdays at 12:30 in EP 122***

| | |
|---|---|
| August 31 | Dr. Kathleen Kay, UC Santa Cruz<br>Pollinator driven speciation in plants — insights from Neotropical gingers and California Clarkias |
| September 14 | Dr. Laura Kubatko, Ohio State University<br>Coalescent-based phylogenetic inference using site pattern probabilities |
| September 28 | Dr. Kalin Vetsigian, Wisconsin Institute for Discovery<br>The dynamics of microbial interactions in natural and synthetic microbial communities |
| October 12 | Dr. Zachariah Gompert, Utah State University<br>Causes and constraints on the evolution of ecological specialization in herbivorous insects |
| October 19 | Dr. Forest Rohwer, San Diego State University<br>How Phage Create an Immune System |
| November 09 | Dr. T. Trevor Caughlin, Boise State University<br>Quantitative Spatial Ecology |
| November 30 | Dr. Miriam Barlow, University of California Merced<br>Making statistically powerful evolutionary predictions using an antibiotic resistance model |

***The Institute for Bioinformatics and Evolutionary Studies***
ibest.uidaho.edu

***The Center for Modeling Complex Interactions***
cmciuidaho.org

ibest

# Appendix V. 2017 Business for Scientist Attendees

| | | |
|---|---|---|
| Rowley, P. | Biological Sciences/COS | Assistant Professor |
| Laarman, A. | Ruminant Nutrition and Metabolism/AVS/CALS | Assistant Professor |
| Magolan, J. | Chemistry/COS | Assistant Professor |
| Lee, J. | Biology/ Natural Resources | Postdoc |
| Andrews, K. | LEECG/FWS/CNR | Postdoc |
| Chikh Ali, M. | Plant, Soil, and Entomological Sciences/CALS | Research Support Scientist |
| Ferguson, J. | CMCI/COS | Postdoc |
| Bazurto, J. | Biological Sciences/COS | Postdoc |
| Johnson, T. | Fish & Wildlife Sciences/CNR | Assistant Professor |
| Hemati, S. | Electrical and Computer Engineer/COEng | Assistant Professor |
| Bruger, E. | Biological Sciences/COS | Postdoc |
| Margres, M. | Biological Sciences/COS | WSU- Postdoc |
| Suchar, V. | MILES/CMCI/Stat/COS | Postdoc |
| Sheneman, L. | Northwest Knowledge Network/ORED | Tech and Data Manager |
| Friesen, M. | Plant Biology | Michigan State University - Asst Prof |
| Jack, C. | Plant Biology | Michigan State University - Postdoc |
| Gandomi, A. | BEACON | Michigan State University - Postdoc |
| Patel, J | CMCI | Postdoc |
| McLachlan, E. | BEACON | Michigan State University - Lab Mgr |

# Appendix VI. Inland Northwest Genomics Research Symposium

| | | |
|---|---|---|
| **8:30 - 8:50** | **Welcome Address** | |
| **8:50 - 10:00** | **Lightning Talks** | |
| | 8:50 | *Bert Baumgaertner "Social Influence and Infectious Disease"* |
| | 9:00 | *Donna Holmes "An evolutionary medical view of aging in the new millennium"* |
| | 9:10 | *Tanya Miura "Are two viruses better than one? The impact of respiratory viral co-infection on disease severity."* |
| | 9:20 | *Luke Harmon "Make way for the robobabies: Human Health and Welfare over Billion-year Timescales."* |
| | 9:30 | *Barrie Robison "Modeling biological processes in video games"* |
| | 9:40 | *Omar Cornejo "Selection on DARC in humans was soft, yet strong"* |
| | 9:50 | *Andreas Vasdekis "A single-cell walk through metabolism, as a matter of fat."* |
| **10:00 - 10:30** | **Break** | |
| **10:30 - 11:30** | **Expanded Talks** | |
| | 10:20 | *Shirley Luckhart "Parallels, opportunities, and shared challenges across plant, animal, and human diseases: developing novel strategies for intervention based on multiple scales of biology."* |
| | 10:50 | *Kyle Harrington "The Resilience of Regenerative Systems: Complementarity of Ecology and Development"* |
| | 11:10 | *Dilshani Sarathchandra: "A Survey Instrument for Measuring Vaccine Acceptance."* |
| **11:30 - 12:30** | **Breakout Session** | |
| **12:30 - 1:30** | **Lunch** | |
| **1:30 - 2:30** | **Lightning Talks** | |
| | 1:30 | *Jodie Nicotra "Concepts of 'Normal' in Microbiome Research."* |
| | 1:40 | *Brian Kennedy "Understanding finer aspects of biodiversity and its relevance for stability in river ecosystems"* |
| | 1:50 | *Janet Rachlow "Multisensory concealment and detection: predation, crime, and aesthetics"* |
| | 2:00 | *James Van Leuven "Attenuation and epistasis via genome recoding of phiX174"* |
| | 2:10 | *Cory Gall "Manipulation of the tick microbiome: a novel tool for combating tick-borne diseases on Native American reservations."* |
| | 2:20 | *David Tank "How I study the dimensions of biodiversity in plant systems and why it matters"* |
| **2:30 - 3:00** | **Break** | |
| **3:00 - 4:00** | **Expanded Talks** | |
| | 3:00 | *Kelly Brayton "Tick-borne pathogens in the Mnisi ecosystem"* |
| | 3:20 | *Sarah Hendricks "A novel system for the genetics of inflammation-induced cancer"* |
| | 3:40 | *Larry Forney "Tracking Woozles in Women's Health Research"* |
| **4:00 - 4:30** | **Breakout Session** | |
| **4:30 - 5:00** | **Final Recap** | |