Minutes of the Fly Board Meeting—07202020

1. Thanks to out-going members, welcome to incoming members

Mariana introduced the meeting, offered thanks to outgoing FlyBoard members and welcomed new members, including our President Elect Tin Tin Su. She also welcomed Denise Montell, current GSA President and Hugo Bellen, incoming GSA President, and the GSA Executive Director and Staff. She summarized the planned agenda.

Mariana then highlighted three major issues and one more minor currently facing our community. 1. NIH and NSF funding, and the importance of model organism research. 2. Future of Fly Meeting. 3. Diversity and Inclusion. 4. Adding a Trainee Representative.

2. Reports from the Bloomington Stock Center and the VDRC

Bloomington Drosophila Stock Center (Kevin Cook): The stock center is doing OK. 80 folks help run the stock center. The last few months have been dominated by the adjustment to covid-19, including the usual social distancing and other measures and planning for possible personnel absence (e.g. cross-training). Shipments were suspended for 7 weeks and orders have been down by 50%, reducing fee income (70% of operating budget comes from this source and they had to pay workers time and a half during the shutdown). This led to a significant budget deficit. They are looking at a fee increase of 25-55% for 2021 (e.g., a 25% increase would be from ~$5.50/stock to ~$7/ stock). This is not entirely due to covid-19 as there has been a long time freeze in fees (in part due to issues with grant accounting), and the Stock Center had been deficit spending. The hope is NIH may provide some additional funding, but there is no concrete information. Kevin was asked if there were fee waivers. Kevin pointed out that this had been a feature in the past but it had been abused.

VDRC (Lisa Meadows): The situation is much the same as for Bloomington. People were working shifts. No stocks were lost. They are now back to normal in shipping. Orders are way down and this is also leading to budgetary issues. They have 4 years of funding in place--It's unknown how this will be resolved. They may have to increase fees, or discard more low-use stocks.

3. Discussion from representatives of other community resources.

FlySlack: Erika Geisbrecht reports there are now 229 members. Please help advertise this to the community.
flyslackco.slack.com

FlyBase (Norbert Perrimon). The covid-19 impact eas relatively low, due to the fact that most folks could work remotely. They are in the 3rd year of a 5-year grant—this included a 5% cut. Some money has been returned from voluntary community fees (~$300K over 2 years). A small grant will be submitted to NSF. They are working on integrating single cell data.

Gene Disruption Project, BDGP, etc. (Hugo Bellen and Sue Celnicker). NIGMS refused to consider the GDP renewal grant, as they decided to stop supporting additional grants for model organisms. The GDP is collaborating with Norbert and the FlyBase folks to try the grant at ORIP - Office of Research Infrastructure Programs. Sue Celnicker noted the same issue applies to BDGP grant support.

4. FlyBoard Treasurer's report (Michelle Arbeitman)

She showed data on the funds we have invested—this was completed with the help of the GSA. The Finnerty Award could be invested as well—the new Treasurer Jessica Treisman will look into this. The Sandler fund is also administered by the community.

A small committee was convened to discuss use of the reserve funds for travel awards. Michelle then summarized their proposal, which was shared with the Board before the meeting.

In this proposal 5% of the account value would be spent per year. This would fund about eight $599 travel awards. She opened this for discussion. Mariana asked whether we would consider other awards orprograms—e.g., the proposal by Tina Tootle asking us to consider funding for a summer outreach program. Tin Tin asked how we would use these funds in the case of a virtual meeting—this could be accommodated by changing the name to funding for "Conference attendance", and making amount flexible. We could also call them "trainee" awards, or open this to including "child care awards". Debbie Andrew advocated for funding high school outreach programs, especially in these times of Black Lives Matter. Tin Tin Su and Tina Tootle summarized the Eclose Program run from Fox Chase.

https://ecloseinstitute.org

There also was a disucssion about whether some of the excess earnings on the Sandler funds could also be used for other purposes. Michelle Arbeitman and Jessica Treisman will look into whether this is financially feasible, and if so we will then consult former Larry Sandler students like Scott Hawley.

A motion was made to 1. Adjust the proposal from the Treasurer to make it clear that the funds earned by the invested Main Endowment can be used more broadly to support efforts to increase trainee attendance and diversity in our community, like the ones we discussed during the meeting. 2. Empower the Treasurers to examine the Sandler and Finnerty balances, determine whether there are sufficient earnings to fund additional uses for this fund, and then involve Scott Hawley in this decision.

25/26 voted yes, and 1 abstained.

5. GSA's roles, initiatives, and relationship to the fly community

Denise Montell (GSA President): She pointed out how the last year has re-emphasized the value of scientific societies. The successful transition of TAGC to an online was an example. 14,000 attended at least part of the meeting. The financial loss was reduced to $450K, much lower than had been anticipated and it can be considered an investment in the future. This also was illustrated in the GSA response to the George Floyd murder and the Black Lives Matter protests. The GSA acknowledged its own failings in DEI and in working in anti-racism. GSA re-opened nominations for the Board to try to find more diverse candidates. GSA also publically addressed the xenophobic policies of the current administration.

Denise then addressed the relationship between the GSA and the Fly Community. We are the largest of the model organism communities and well represented in the leadership. She noted that we have reserve funds because the GSA took on the financial risks associated with the Fly Meeting, and reminded us of the role GSA staff play in organizing our meeting. She noted the need for a "meeting organizer guide" as they turn over each year.

Tracey DePellegrin, GSA Executive Director. She began with a brief re-cap of TAGC and the lessons learned. She noted that the Drosophila community played important roles in organizing the meeting. She discussed the challenges of keeping people engaged. She shared stats—including 76,000 hours of recordings watched. Attendees emphasized that while they want to eventually return to in person meetings, they want to keep hybrid features like talk recordings. There were 3500 online registrants from the Drosophila community. Online registration increased the fraction of registrants who were graduate students. 20% reported an under-represented background.

In discussion Laura Johnston pointed out that we need to ask trainees about their feelings about our community, and ensure we listen to their thoughts about how to maximize the benefit of the meeting to trainees. Helena Araujo (Latin American Rep) pointed out that Brazilian institutions won't cover registration costs. Justin DiAngelo (PUI rep) emphasized that the same issues will apply to PUI people, this year in particular. Denise noted that they are mindful of the impact on different communities, but also noted online meetings do cost money. Lynn Cooley reiterated the value of close collaboration between the GSA and the FlyBoard. Amanda Norvell pointed out that hybrid meetings can reduce the environmental impact of the meeting.

6. Fly Meetings

TAGC 2020

Suzy Brown (GSA Meeting Organizer). The Drosophila Community was well represented at TAGC. She offered some additional statistics, and noted the work done by the scientific organizers to put together the science. She noted that the meeting offered a chance to bring in some additional international participants—particularly noted the ability to bring in participants from Iran. She pointed out the need to strengthen networking and community building in an online meeting. There was some discussion about how to improve poster presentations and scheduling them to improve interactions. Suzy noted they are checking out other online

meetings (e.g. SDB) to look for innovation. Alan Spradling noted issues with online recording, as the universal data accessibility may be a disincentive to presenters.

Dros2021 plans. Nasser Rusan
Co-chairing with Amy Kiger. The committee also includes Guy Tanentzapf, Nadia Singh, Karen Hales, and Michelle Arbeitman. Their goals—increase participation and match what the community wants. They will look for good ideas from other communities. In early meetings, they quickly chose a Keynote Speaker, Harmit Malik, who agreed. The committee soon realized planning an all on-line meeting was the best option and they got this approved at a variety of levels, with unanimous agreement. Their new goal is: How can we make the best online meeting. They are meeting in alternate weeks—the committee alone alternating with a meeting with GSA staff. They are currently focusing on abstract submission and meeting program.

Major current issues: Timeline—2 days versus week long versus longer. Currently looking at a longer meeting. Zoom fatigue is an issue and that issue is different for PIs and trainees. They are thinking about the issue of parents with kids at home. They recognize that longer timelines mean issues for colleagues with teaching responsibilities. They are accommodating times for Europe and North America participants (11AM-1PM; 2-4PM ET). They are discussing what features can be translated to an online meeting. One decision--Sessions will be assembled based on abstracts submitted rather than creating sessions and binning abstracts to them. They will keep the new PI forum, and add a similar forum for trainees, especially first timers. They are adding a dedicated Plenary on Diversity and Inclusion—this and other ideas are still a work in progress. They will try to keep the lunch network and add post-session break-out rooms. Poster format and the positioning of career development in program are still in the works.

Discussion continued. Questions about the poster sessions were raised. There was a suggestion of lightning talks as a format—this was tried at Crete and worked but it requires strict timing and it can be exhausting. Justin DiAngelo pointed out the fact that for undergraduate presenters a lightning talk would be a challenge—it was suggested that undergrad posters could be separated and it was noted that sessions were smaller and this reduced stress. Suzy Brown offered suggestions based on different options they used during TAGC. Celeste Berg pointed out the need for guidelines for actual poster format. Bruce Edgar pointed out that we should be thinking about the next transition to the 2022 hybrid format. Suzy pointed out that GSA is thinking about this across communities for future hybrid meetings. Nadia Singh discussed the issue of Award naming, and Mariana pointed out that we had discussed this with a decision that we won't name any new awards for people. Amy Kiger emphasized the fact that the committee has increasing inclusion and participation as a major goal, and emphasized that the

Future meetings: Mark Peifer reminded us we had previously discussed possible changes to two big picture issues with future meetings—timing and venue. We had opened a discussion of the possibility to go to an every-other year option (potentially alternating with the European

Meeting), and whether we want to continue an East Coast/West Coast/Midwest alternation. Suzy said she is currently negotiating the likely possibility of using the Town and Country in San Diego for Dros22. TAGC's future is still under discussion. Hugo Bellen pointed out that there is a lot of uncertainty about the impact of hybrid meeting on in person attendance. Suzy said the current research suggests that people still will go to in person meetings. She also pointed out that travel costs are often more important than registration. We ended by emphasizing the need to work with the GSA very carefully, and it means we will likely meet again as a FlyBoard in the Fall.

Tin Tin Su then raised two issues. How do we better bring trainees into the planning and discussion at all levels? Trainee board members, trainee inclusion in program planning, other ideas? Denise emphasized the idea that having people join the GSA is part of it. Bruce Edgar agrees we need trainees on the board. Should we have both a graduate student and a postdoc on the Board? There was general agreement that both was the best idea. Should we have a trainee on the Elections Committee or other committees? We'd make them different people than the board members, to avoid over-burdening indiviudals.

Tin Tin then moved to Diversity and Inclusion. Do we have data on the demographics of the Drosophila Community? Tracey told us GSA is now collecting self-reported data on GSA members. This is being provided to meeting organizers. Nasser Rusan reported that they now have this data and are using it in planning and increasing participation for Dros21. Tin Tin Su emphasized we need to use this data in all of our deliberations. She also pointed out that in thinking about helping folks participate, we need to think about their costs.

Bruce Edgar is chair of this year's elections committee. He asked us for suggestions and pointed out they will be making a "all Drosophila Community" call for nominations. He pointed out that those on the ballot last year who were not elected are open to be nominated again.

Reports begin on the following page.

FlyBoard Reports—July 2020
Each year the FlyBoard receives reports from committees involved in managing community activities (Elections and Awards) and from organizations who maintain community resources. Below are reports from these entities. Note that some of these were compiled and received before the move of Dros20 online and some were compiled and submitted afterward.

Table of Contents for Reports

Report from the Drosophila Board Elections Committee (Debbie Andrew)

The 2019 Elections Board Committee included two members from the previous year: **Tin Tin Su** and **Noah Whiteman**. During the yearly board meeting, **Laura Reed** and **Patrick O'Grady** volunteered to participate and were immediately accepted as members by the 2019 Elections Board Chair, **Debbie Andrew**. Based on communications between Iswar Hariharan and Mark Peifer regarding the poor representation of ethnic diversity on the Drosophila Board, **Iswar Hariharan** was invited to become the sixth member of the committee.

Although traditionally nominations for candidates had been made by the election committee, Iswar suggested that we solicit open nominations from the entire community, an idea that we readily agreed upon, acknowledging that we might have to resort to the traditional mechanism if enough nominations were not received.

June 5: The first announcement asking for nominations from the community for Fly Board members was posted by GSA. The nominations deadline was originally set for Aug 1.

August 1: Brian Calvi sent out a request for new Fly Board nominees to all those on the Flybase Newsletter Mailing list (2224 people), since we had only two nominations for the same person

August 14: GSA ran another announcement on August 14.

August 23: No additional nominations were received.

August 23: Brian Calvi sent another email to the Flybase Newsletter Mailing List

August 23: Debbie emailed members of the Drosophila Board asking for nominations.

August 23: Several Tweets were sent out to ask for additional nominations from the community, including a Tweet from the GSA.

We got several nominations from members of the Drosophila Board. Additional nominations were from members the elections committee.

A list of potential candidates was created and candidates were prioritized, with a goal of creating a more inclusive membership.

Each committee member chose to cover an election or two. Their tasks were to recruit two candidates from the prioritized list and solicit blurbs from the candidates who agreed to serve. The blurbs were included in the election ballots. Sample blurbs were provided to candidates.

President-elect: Debbie

Treasurer: Debbie
New England Rep: Iswar
Mountain Rep: Noah
Australia/Oceana Rep: Tin Tin
PUI Rep: Laura Reed

This process began on **Oct 9** and we had our candidates by **Oct 17, 2019**.

Since some of our initial first choices for president elect declined to run and since several members of the committee thought that Tin Tin Su would make an excellent president, she resigned from the elections committee and we nominated her as a candidate for president elect on **Oct 14**.

The first candidate's blurb was received on **Oct 13** and final candidate's blurb was received on **Oct 30.**

On **Nov 7,** the following email was sent to everyone on the Flybase Newsletter Mailing list (2224 people). This emailing was facilitated by Brian Calvi and Josh Goodman.

Dear Drosophila researcher,

It is time to cast your vote for new members of the National Drosophila Board Members. The Drosophila Board plays an important role in the Drosophila research community, so please take a few moments to learn about the Board and, importantly, participate in this election. The Board's duties include overseeing community resource centers and addressing other research and resource issues that affect the fly community. The Board also administers the finances for the annual North America Drosophila Research Conference and its associated awards, and it chooses the organizers and the site of the annual meeting. The Board consists of 13 regional representatives: Eight from the U.S. and one each from Canada, Latin America, Europe, Asia and Australia/Oceania, and one representative for primarily undergraduate institutions, all of whom serve 3-year terms. The Board is led by three elected officers: a President, a President-Elect and a Treasurer. In addition, the Board has ex officio members, including past-Presidents, meeting organizers and representatives of the Drosophila community resource centers.

This year we are electing the President-elect, who will serve as President starting with the Fly meeting in 2021 and a Treasurer, who will serve from 2021 through 2024. We are also electing regional representatives from New England, Mountain, Australia/Oceana and primarily undergraduate institutions. Please participate in this election! This is your opportunity to choose the individuals who will help set priorities and secure support for community resources. In order to record your vote, please go to the following URL and follow the instructions on that page.

https://www.surveymonkey.com/r/CMF6P65

Balloting will end December 6, 2019.

Thank you,

The Drosophila Board Election Committee: Debbie Andrew (Chair) Iswar Hariharan

Patrick O'Grady Laura Reed Noah Whiteman

Dec 20: Results of elections received

Tin Tin Su - University of Colorado - was elected future president. She ran against **Cassandra Extavour** - Harvard. 367 people voted for President Elect.

Jessica Treisman - Skirball Institute - was elected future treasurer. She ran against **Rachel Cox** - Uniformed Services University. 366 people voted for Treasurer

Alexey Versaksa - UMass Boston - was elected future New England board representative. He ran against **James Walker** - Mass General/Harvard. 342 people voted for New England Board Rep

Kieran Harvey - Monash University - was elected future Australia/Oceana board representative. He ran against **Leonie Quinn** - the John Curtain School of Medical Research. 340 voted for Australia/Oceana Board Rep

Nadia Singh - University of Oregon - was elected future Mountain region board representative. She ran against **Young Kwon** - University of Washington. 332 voted for Mountain Region Board Rep

Justin DiAngelo - Penn State Berks - was elective future Primarily Undergraduate Institutions board representative. He ran against **Afshan Ismat** - University of St. Thomas. 328 voted for Primarily Undergraduate Institutions Board Rep

All of the candidates who were not elected were not only gracious about their loss but they also agreed to run again in the future. We also have some additional names on the list, including some people who wanted to postpone their candidacy. So please consider these individuals in future board elections.

Image Award (Nasser Rusan)

This year the Image Award committee has added two new member - Irene Miguel-Aliaga from Imperial College London and Toshie Kai from Osaka University, bringing the total committee members to 6 (including Nasser Rusan, Elizabeth Chen, Don Fox, and Mia Levine). Committee members are asked to serve 2-3 year terms, which will allow for steady turnover. It is therefore likely that 2 members of the current committee (Mia and Don) will be replaced for next year's competition.

This year, we have maintained our Twitter presence, especially leading up to the submission deadline.

Results of the 2019 competition

73 total submissions: 51 images and 22 videos. That is a 14% decrease from last year.

The winners this year are:

Igor Siwanowicz (image), for a depth color-coded confocal microscopy image of fruit fly reproductive organs during coitus

https://www.sciencedirect.com/science/article/pii/S0896627319303447#!

Todd Schoborg (video), for a video showing an adult female in 3D digitally sliced along the XZ and YZ axes to reveal all intact internal structures

https://dev.biologists.org/content/146/23/dev176685

The runners-up are:

Julia Sauerwald and Frank Schnorrer (image), for an image showing the developing flight muscles, branches of tracheal terminal cells invade the myotubes and ramify within them to provide oxygen supply

https://elifesciences.org/articles/48857

Wenze Li (video), for a video showing rapid movement and the neuronal activity of the crawling larva

https://www.sciencedirect.com/science/article/pii/S0960982219300892

Report on 2020 Larry Sandler Award

2020 Committee:
Barbara Mellone (Chair)
Amanda Larracuente
Erica Larschan
Brooke McCartney
Guy Tanentzapf

2021 Chair: Guy Tanentzapf

Process:
The committee received **24 nominations**. Most of the nominations came on the day of the deadline.
We ranked applicants using a 4-point voting system. Anticipating that all the nominees would be quite strong, we used the following ranking rubric: 1=outstanding, 2=excellent, 3=very good, 4=good.
The committee was instructed to review candidates in a random fashion (and not, for instance, in alphabetical order, or in the order they were listed on the scoring spreadsheet). We met via WebEx to discuss the rankings, starting with nominees that received three scores of 1, followed by nominees with two scores of 1 and any additional nominees suggested by members of the committee. We chose **four** finalists to review further by reading the full theses. After each member of the committee submitted their 1-4 rankings to the Chair, the committee met again by WebEx to decide the winner. After a stimulating discussion, we settled on a unanimous winner. We also selected **two** nominees as co-runner-ups.

2020 winner: **Dr. Balint Kacsoh**, Ph.D. Geisel School of Medicine at Dartmouth (Mentor: Dr. Giovanni Bosco).

Abstract supplied by the nominee:

To what extent does the social environment influence ones' underlying biology? This important question drove my thesis work. Until recently, the relevance of social cueing was relegated to the social sciences and those investigating fully eusocial organisms, such as ants and bees, whose eusocial nature is considered an evolutionary 'quirk' of some highly specialized animals. However, we are now beginning to see the profound impact that the environment, both physical and social, can have on a wide range of biological outputs. Genetic model systems, thus, are indispensable for understanding molecular mechanisms of causation as a function of environment.
Drosophila naturally exist in communal settings, where, throughout their lives, they encounter other Drosophilids (sisters/other species), in addition to other insects and microorganisms, in contrast to the sterile lab setting. Forays into social structure and social learning using Drosophila as a model have only recently begun to emerge. Sociality influences each life stage of Drosophila in some way, such as the larval stage, where wild-type larvae disperse evenly

when initially cohabitated, but then quickly swarm around the periphery of a food source. This exhibition of cooperative behavior by burrowing in clumps is dependent partially on *neuropeptide-F*. Sociality even influences the pupal stages, where social experience dictates fecundity. Drosophila have been recently shown to have the ability to visually perceive dead conspecifics resulting in decreased longevity via a inhibition in serotonin signaling. In food choice experiments, naïve observer flies are allowed to watch 'teacher' flies choose a colored food substrate (yellow or red). One of the substrates contains sucrose, and thus, the flies prefer that color only due to the nutritional addition. Naïve students are then given the same choice (yellow v. red food), but without the sucrose. The student flies prefer the same substrate that was chosen by the teachers, even though no sucrose reward is present. Recent work has demonstrated the importance of the communal aspect of the Drosophila life cycle in tumor growth, where different social environments had dramatic effects on tumor progression and lifespan.

By incorporating the communal aspect of Drosophila life history, and by using a natural predator of *Drosophila melanogaster*, my thesis work explores the processes by which transient adaptations occur both within and between Drosophila species. Predatory wasps are ubiquitous keystone species in many ecosystems around the world. We utilized wasps that prey on immature stages of other insects, using larva and pupa of certain species as hosts for their own offspring. Such wasps pose a serious threat to juvenile Drosophila, with infections rates as high as 90% in natural populations. Adult Drosophila have evolved complex behavioral changes to protect their offspring from these predators, including altered food preference and reduced egg laying. In my thesis work, I present two behavioral phenotypes that are elicited in response to parasitic wasps: an ethanol preference and socially communicated reduction in oviposition rate. Furthermore, I utilize these behaviors to identify novel learning, memory, and social learning genes, in addition to the underlying neural circuitry governing these behaviors.

First, I identify a unique ethanol preference following wasp exposure that is dependent on canonical long-term memory genes, by testing mutant (-/-), heterozygous (+/-), and transgenic lines expressing RNAi hairpins in the fly mushroom body. The memory is not permanent, as the oviposition preference returns to base-line after 10 days. Using this paradigm, I sought to identify the effect of age on memory decay. I found accelerated memory decay in aged animals partially dependent on changes in mushroom body morphology. Furthermore, I found a haploinsufficiency effect of PTEN and FMR1, genes whose mutations that are involved in autism spectrum disorder and fragile-X, respectively. Excitingly, when these genes are mutated in the same genetic background, the aging effect is rescued both with respect to the memory and the mushroom body morphology phenotypes, demonstrating a unique interaction between these two genes.

The second behavior I analyze is the observation that female flies reduce their oviposition rate in response to endoparasitiod wasps. I found that this reduced oviposition rate is a result of stage- specific apoptosis in the female germline. This altered germline state is socially transmitted between flies through visual cues, where a wasp exposed female fly communicates the wasp presence to a naïve observer fly, resulting in a change in the germline physiology of the observer. This communication is conserved across the genus Drosophila. Inter-(between-species) communication also exists, though the efficiency of communication is dependent on phylogenetic distance of the two species being tested. Less efficient communication can be

enhanced following a cohabitation of the two Drosophila species, termed the dialect training period. This dialect training is extremely multimodal, and I identified regions of the Drosophila brain that are both necessary and sufficient for the behavior. This is one of the few examples of socially transmitted behavior in Drosophila, and a powerful illustration of the influence of the environment and conspecifics on behavior and reproduction.

Finally, I utilize the two behaviors outlined above to explore the dynamics of gene expression through the memory formation process and social learning in a two-pronged approach: 1) by utilizing two unique machine learning approaches; and 2) by performing RNA sequencing on female fly heads at various time points. Both of these methods produced novel genes predicted to be involved in memory formation. Analysis of RNA sequencing and machine learning revealed functional enrichment of several gene clusters, including signal peptides involved in the Immune Deficiency Pathway, and genes implicated in autism spectrum disorder (*minibrain* and *Mob4*). Neuronal knock down of several of the genes identified in both the bioinformatic and sequencing approaches produce flies deficient in memory formation and maintenance, identifying several novel genes involved in learning and memory. I also identified genes involved in memory retention and social learning using these paradigms. Of particular excitement, I found a histone deacetylase involved in post-memory formation (retention) and as a mediator of social learning. These findings provide one of the first clear links between transcriptional changes in immunity pathways to learning and memory functions in addition to successfully utilizing a bioinformatic gene prediction pipeline on an organismal level behavior.

In my thesis work, I utilized Drosophila to dissect previously unknown behavioral and physiological responses to environmental conditions, gleaning insight into the acquisition of memory, sociality, and neuronal plasticity. Collectively, my work presents an argument for the profound and elaborate influences of the environment on individuals. These studies were also the first of many demonstrating innate and social behaviors in Drosophila that have since birthed a larger field of social behavior studies using this model system. Sociality exists across many branches of the tree of life and is thought to encompass a spectrum of behaviors. But is there a single criterion that is truly suggestive of all forms of sociality? E.O Wilson suggested that the common denominator of sociality is the ability to communicate. Wilson states, “the terms society and social must be defined quite broadly in order to prevent arbitrary exclusion of many interesting phenomenon.” The interplay of neurobiology and the social sciences is a critical region of study, termed “sociogenetics.” E.O. Wilson accurately predicted that a split in these two disciplines would occur primarily based on mechanistic/molecular approaches and evolution/ecological approaches. It is my hope that my thesis and studies within begins to bridge the gap between multiple fields of study.

2020 co-runner-ups:
Dr. Victoria Deneke, Duke University Medical Center (Mentor: Dr. Stefano DiTalia)
Dr. Miranda Hunter, University of Toronto (Mentor: Dr. Rodrigo Fernandez-Gonzalez)

Comments and suggestions on the selection process.
The process ran similarly to previous years, with a slight change in the nomination process. In an effort to make the letters more consistent and comparable for the selection process, the 2019 Sandler committee recommended changing the format of the nomination letter. The 2020

call for nominations asked the mentors to directly answer the following questions: 1) What are the main discoveries in the nominee's thesis? What are the intellectual and experimental contributions of the nominee to the project? How does the thesis work advance what was previously known in the field and to previous work from your lab?
Although a few mentors organized their letter to address these questions explicitly, making it easier for the committee, many did not. For next year, the 2020 committee suggests adopting a template form instead of a free-form letter.

The committee discussed the gender distribution of recent winners and their mentor's. The numbers of male and female nominees in 2020 was 60% and 40%, respectively (note that gender was inferred from the nomination letters), similar to that of recent years. Only 6 of the 24 nominating mentors were women (note the caveat that gender was inferred from the mentor's first name and was assumed binary), consistent with the history of this award. We noted that historically students from female mentors are nominated less often and are rarely selected for the award (see data put together by Suzy Brown here). To take steps to increase the diversity pool of winners and their nominating mentors, the committee suggests that data not only on gender, but also on racial/ethnic background of both nominees and their mentors should be tracked in the future using optional self-reporting questions in the nomination process. An online nomination process would automate the collection of this data.

Nominees for 2020

Nominee Last Name	Nominee's first name	Gender	PI Name	Gender
Goodman	Lindsey	F	Nancy Bonini	F
Hopkins	Ben	M	Straurt Wigby	M
Frank	Dominic	M	Marco Gallio	M
Warecki	Brandt	M	Bill Sullivan	M
Hatkevich	Talia	F	Jeff Sekelsky	M
Kacsoh	Balint	M	Giovanni Bosco	M
Crocker	Kassi	F	Grace Boekhoff-Falk	F
Hope	Kevin	M	Lawrence Reiter	M
Ko	Clint	M	Adam Martin	M
Yoon	Yoseop	M	Urs Schmidt-Ott	M
Wooten	Matthew	M	Xin Chen	F
Kursel	Lisa	F	Harmit Malik	M
Courgeon	Maximilien	M	Claude Desplan	M
McGurk	Michael	M	Dan Barbash	M

Hunter	Miranda	F	Rodrigo Fernandez-Gonzalez	M
Ryu	Taehyun	F	Irene Chiolo	F
Anreiter	Ina	F	Marla Sokolowski	F
Sethi	Sachin	M	Jing Wang	M
Siddiq	Mohammad	M	Joseph Thornton	M
Deneke	Victoria	F	Stefano Di Talia	M
Liu	Yang	F	Mark Rebeiz	M
Zheng	Yiming	M	Tim Megraw	M
Manzo	Ernesto	M	Daniela Zarnescu	F
Yang	Sheng-An	M	Wu-Min Deng	M

--

Email to nominees selected to be finalists:

Dear XXX,

As you know, your advisor, Name LastName, nominated you for the prestigious Larry Sandler Award and I am excited to let you know that the award committee selected you as one of the finalists.
I am therefore requesting an electronic copy of your thesis for full review by the committee. Please send it to me as a pdf file via email or use wetransfer.com if the file is too large.

Congratulations!
Best wishes,
Barbara Mellone

Email to mentors of nominees that were not selected :

Dear XXX

Thank you for nominating your student for the 2020 Larry Sandler Award Memorial Award. We had an extremely strong pool of 24 applications this year, and it was a challenging process for the committee to choose a winner. I am sorry to let you know that your student was not selected. On behalf of the selection committee, I thank you for taking the time to support your excellent student and for helping to keep the Sandler Award selection process a true reflection of the breadth and strength of Ph.D. research in our community.

Sincerely,

Barbara Mellone (Chair)
Amanda Larracuente
Erica Larschan

Guy Tanentzapf
Brooke McCartney

Email to winner:

Dear Balint,
I am enclosing a letter awarding you the 2020 Larry Sandler Award. Congratulations! Please keep this confidential until you receive the award or an official announcement is made.
Best wishes,
Barbara

Letter to winner:

Dear Dr. Kacsoh,

On behalf of the 2020 Sandler Award Committee, I am delighted to inform you that you have been selected as the recipient of the 2020 Larry Sandler Memorial Award!

As I am sure you know, the goal for this award is to identify the "best" Ph.D. thesis in *Drosophila* research from the previous year. In this round we had 24 nominations, which made the competition extremely tight. The committee felt that your beautiful work on “Elucidating genes, circuits, and behavior in a novel *Drosophila* social learning and memory paradigm” stood out as especially significant and deserving of this recognition. It also helped that we received very strong and supportive comments from your advisor, Dr. Giovanni Bosco. Please accept our warmest congratulations on executing this spectacular set of experiments and on a superb thesis. You now join a long list of excellent scientists who have gone on to have successful careers.

As the recipient of this award, you will have the honor of presenting your thesis work at the Larry Sandler Memorial Lecture. This is currently scheduled for Thursday, April 23, 2020 at the TAGC GSA meeting in Washington D.C. In addition to sharing your work with the field, this talk is an opportunity to inspire other students just starting or in the midst of their Ph.D.s. With the uncertainty caused by the Covid-19 pandemic, I defer to Suzy Brown (cc’ed here) for all information regarding this event.

Finally, an important note: the Sandler Award winner has traditionally been presented as a surprise to the community, therefore *please wait* until you have received the award to make any public announcements. If the meeting is cancelled, please wait until a public announcement is made to the community.

Please don't hesitate to let me know if you have any questions.

Sincerely,

Barbara Mellone (Chair)
Amanda Larracuente
Erica Larschan
Guy Tanentzapf

Brooke McCartney

Email to runner-ups:

I am writing to inform you that you have been selected as one of two Runners-up for this year's Larry Sandler Memorial Award.

Although you are not the winner for this year's award, I nevertheless want to congratulate you for executing a spectacular thesis. This year's competition was intense: we received 24 nominations, several of which were truly outstanding and deserving of the Larry Sandler Award. The committee struggled to narrow this down to even a top four. We truly enjoyed reading about your work and accomplishments and have no doubt that you will continue to do superb research in the future. I should add that your advisor was extremely supportive and wrote a glowing nomination letter about you and your work.

On behalf of this year's Sandler Award Committee, we congratulate you on being selected as a Runner-up, and wish you the very best of luck for continuing success.

Report to the North American Drosophila Board,
Amanda Norvell, Finnerty Undergraduate Travel Award Committee

This year we received 31 applications for the Victoria Finnerty (VF) Undergraduate Travel Award, and 27 were reviewed (3 were incomplete, lacking a letter of support from the advisor and 1 was from a high school student as was therefore ineligible according to the standards. Following an initial round of evaluation, 14 applications were moved to a second round of consideration and were all recommended for funding. In order to maximize the number of students who received funding, money was awarded on a sliding scale, according to their ranking.

The awardees are:

Jillian G. Gomez, University of Tampa, $800
Joey Y. Wong, University of California, Santa Cruz, $600
Nicole E. Folan, College of the Holy Cross, $500
Brooke M. Allen, University of Detroit, Mercy, $400
Anastasia Welch, Rhode Island College, $400
Marta Stetsiv, Kansas State University, $400
Corinne R. Crosslyn, University of Evansville, $400
Josh Sikder, University of North Carolina Charlotte, $400
Efren Silva, University of Houston, $400
Anthony E. Ruiz, Bemidji State University, $200
Nathan Fischer, Marquette University, $200
Dailia Fainberg, CUNY Baruch College, $100
Isabella J. Hanesworth, Mercy College, $100
Jordyn A. Moaton, University of Missouri Columbia, $100
This year's selection committee was Amanda Norvell(chair), Dan Cavanaugh, Justin D'Angelo, Jennifer Kennell, Judith Leatherman, Matthew Wawersik.

Note: Due to the cancellation of the TAGC because of the coronavirus pandemic, the Victoria Finnerty Travel Awards were not given. Students were notified that they were awardees and returning undergraduates will be encouraged to apply again next year. The incoming PUI Representative, Justin D'Angelo, will note the names of the students awarded this year, so that information may be taken into account if any reapply for the 2021 Fly Meeting.

BLOOMINGTON DROSOPHILA STOCK CENTER

June 22, 2020

COVID-19 The big challenge we have faced this year is responding to the COVID-19 outbreak. In mid-February, we began planning for potential disruptions and started coordinating with university emergency response officials to assure access and services. We formalized phased plans for curtailing operations that involve lowering culture temperatures, eliminating back-up cultures, and discarding stocks. In early March, we ordered extra supplies, cross-trained personnel among the BDSC, Media Kitchen, Dishwashing Facility and research labs to ensure a reliable fly food supply, and implemented vigorous lab disinfection procedures. We are using empty classrooms and labs and scheduling shifts to maximize physical distancing of stockkeepers. All employees are required to wear masks on site. Staff scientists are working remotely and rotating shifts on campus to minimize contact. We suspended stock shipments from March 26 to May 21 during the university research shutdown, which resulted in a significant loss of fee income, and, as required by the university, we are paying stockkeepers 1.5 times their normal wages until July in recognition of potential risk. The result is that we will overspend our 2019–2020 budget by at least 20%. Our ability to compensate will depend on the availability of relief funds, the course of the pandemic, the effectiveness of our belt-tightening measures, and, perhaps most importantly, continued strong use by fly researchers. It is very likely that BDSC users will see a substantial rate increase next year or perhaps sooner. We are uncomfortable with the uncertainty, but cautiously optimistic that, with time, we will recover from the disruptions.

Stock Holdings as of June 19, 2020

- 76,425 stocks with 79,707 unique genetic components
- 17,467 annotated *D. melanogaster* genes are associated with alleles, constructs or deficiencies in the collection
- 12,530 annotated *D. melanogaster* genes are associated with alleles or constructs in the collection
- 1,779 non-fly genes (1,593 human genes) are associated with constructs in the collection

2019 Use Statistics

- 204,672 samples shipped in 12,736 shipments
- 2.8 orders per stock on average with a range of 0 to 171; 62% of stocks ordered at least once, 15% ordered 6 or more times, 6 stocks ordered >100 times, the most popular stock was elav-GAL4 (#8760)
- 3,347 registered user groups, 1,948 of which ordered stocks in 2019
- 7,450 registered users, 2,727 of whom ordered stocks under their own name in 2019

Growth 5,388 stocks were accessioned in 2019:

- 1,474 guide RNA stocks for gene knockout from the TRiP
- 678 Exelixis-generated transposon insertion from Spyros Artavanis-Tsakonas
- 470 UAS-human-gene stocks from Hugo Bellen, Sue Celniker and others

- 430 guide RNA stocks for gene overexpression from the TRiP
- 372 “Chemoconnectome” allele and reporter stocks for chemosensory genes from Yi Rao
- 316 GAL4 ‘CRIMIC’ stocks made via CRISPR/Cas9 from Hugo Bellen, Norbert Perrimon and colleagues
- 305 guide RNA stocks for gene knockout from Oren Schuldiner
- 106 epitope- and fluorescent peptide-marked transcription factors from the modERN Project
- 92 lexA enhancer trap lines from Lutz Kockel (StanEx Project)
- 70 apoptosis sensors from Suzanne & Monier combined with GAL4 drivers from Norbert Perrimon
- 53 stocks for labeling or manipulating neurons from various donors
- 45 transgenes expressing GAL4 in the patterns of ionotropic receptor genes from Richard Benton
- 42 short hairpin RNA stocks for RNAi from the TRiP
- 39 guide RNA stocks for gene knockout from Chun Han
- 27 GAL80 variants or lexA expressed in chemosensory neurons from Jessica Eliason
- 27 stocks for studying human disease from Perlara, PBC
- 22 tool stocks for using QF-based split transcriptional activators from Olena Riabinina
- 820 stocks from other donors

Staff 61 stockkeepers (25.5 full-time equivalents), 6 managers/scientists and 2 research associates.

Funding We are in year 1 of a 5-year grant from NIH with $462,969 in direct funds with funds contributed by OD, NIGMS and NICHD. Years 2—5 will have funding of $445,046. We received supplemental funds from NINDS for maintenance of the split-GAL4 collection generated by the Janelia Research Campus with prorated direct funds of $28,021 this year and $86,674 in direct funds for years 2—5. Fee income covers our remaining expenses and, in recent years, has accounted for ~79% of our funding.

New Stocks We expect to add ~3,850 new stocks in 2020 (numbers may change significantly due to COVID):

- 1,500 TRiP guide RNA and RNAi stocks
- 1,000 UAS-human-cDNA stocks from Hugo Bellen, Sue Celniker and colleagues
- 800 CRIMIC stocks from the Norbert Perrimon, Hugo Bellen and colleagues
- 50 stocks expressing tagged transcription factors from the modERN project
- 500 assorted stocks from the community at large

Pruning We did not undertake systematic culling in 2019, but we lost or discarded 127 assorted stocks. In April 2020, 1,744 stocks were deliberately discarded as part of our response to COVID-19. The largest group carried transgenic constructs with GAL4 driven by regulatory fragments from genes expressed during development that had received little use.

Scientific Advisory Board

- Hugo Bellen, Baylor College of Medicine (chair)

- Nancy Bonini, University of Pennsylvania
- Lynn Cooley, Yale University
- Susan Parkhurst, Fred Hutchinson Cancer Research Center
- Norbert Perrimon, Harvard Medical School
- Benjamin White, NIH, National Institute of Mental Health

Vienna Drosophila Resource Center (VDRC), Vienna, Austria

The VDRC (**www.vdrc.at)** is part of Vienna Biocenter Core Facilities, a **non-profit** research infrastructure. Its mandate is to maintain and distribute transgenic RNAi lines and other resources to Drosophila researchers, both locally and worldwide, and to further develop and expand VDRC resources according to the emerging new technologies and community needs.
User fees are subsidized by the Austrian Federal Ministry for Science and Research and the City of Vienna.

Key changes during 2019

- Since Austria currently represents less than 2% of the total VDRC usage, (with US ~40%, Europe ~40% and Asia ~15%) our core funding has been reduced and we are required to increase the proportion of our running costs coming from user fees.
- User fee structure was adjusted in Sep 2019: prices were reduced for orders of 6 lines or less. To offset the slightly increased price for larger orders, we improved our service to guarantee that all orders up to 300 lines (previously 100) qualify for fast shipment and will be dispatched within 2 weeks.
- In order to streamline the VDRC stock collection we discarded nearly 3,500 RNAi lines (3,339 GDs and 122 KKs), aiming to keep 2 independent lines per target gene, each line with a different construct and different target sequence.
- A further 764 VT lines were culled due to very low usage.
- Incorporation of Heidelberg CRISPR Fly Design library, donated by Fillip Port and Michael Boutros. Collection includes transgenic lines with 2 sgRNAs under UAS control (~1600) targeting primarily TFs, kinases, phosphatases and fly orthologs of genes implicated in human pathologies. Toolkit stocks include UAS-Cas9 lines of various strengths and combined with common GAL4 driver lines.

Usage Statistics 2019

- **42,257** stocks delivered to **632** user groups in **1,465** separate orders.

Resources as of Mar 2020

Total stocks currently available to the community: **26,000**

- 23,425 RNAi lines (12,942 in GD, 9,680 in KK and 803 in the shRNA collection).
- 21 toolkit stocks used for the construction of the RNAi collections.
  Collectively, the GD, KK and shRNA libraries target a total 12,671 Drosophila protein-coding genes (91%). For over 8000 genes, more than one independent RNAi line is available through the VDRC.
- 1,459 UAS-sgRNA and 23 Cas9 toolkit lines for CRISPR-mediated genome engineering (Heidelberg, HD-CFD).
- 200 enhancer-GAL4 lines (VTs, Vienna Tiles). Expression patterns annotated in adult brain and embryo. Searchable databases available.
- 895 Tagged FlyFos TransgeneOme (fTRG) lines.

- A small, but growing number of plasmids and stocks made available to the community from Private Stock Collections, including mutant alleles, tagged constructs and reporters.
- 13,848 DNA constructs used for the generation of the GD collection.

Services
VDRC is open to **donations** of highly used stocks for integration into its community stock center collection, complementary to other stock centers.
In addition, we offer a Private Stock Keeping Service to maintain and distribute personal fly stock/plasmid collections on a cost recovery basis and also offer a fly food service.
See VDRC policy for stock keeping services.

Future
In addition to acquiring new stocks, we are keen to discuss involvement at an early stage to help develop new resources. In addition to stock maintenance and distribution, our team has significant experience in high throughput construct generation, Drosophila injection and transgenic production.

DRSC/TRiP Functional Genomics Resources at Harvard Medical School

Big News! The Drosophila RNAi Screening Center (DRSC) at Harvard Medical School was recently awarded NIH NIGMS P41 funding to form the Drosophila Research and Screening Center-Biomedical Technology Research Resource (DRSC-BTRR). The DRSC-BTRR will focus on technology development, including CRISPR/Cas-based technologies, in collaboration with 'driving biomedical projects' by other labs. More information about the DRSC-BTRR will be added at the DRSC/TRiP Functional Genomics Resources website and presented at TAGC 2020. Thank you to the community for your continued support of DRSC activities!

The DRSC continues to serve as a platform for high-throughput RNAi and CRISPR screens in Drosophila cell lines. For example, check out Nicholson et al. 2019 in Science Signaling to see how our platform is being used for cancer therapeutics discovery. We recently published two protocols in *Current Protocols in Molecular Biology*, describing in detail how to do fluorescent protein knock-in and CRISPR pooled screens with Drosophila cells. We also generated several GFP knock-in cell lines and made them available through the DGRC in Bloomington.

DRSC Bioinformatics launched a new specialized CRISPR sgRNA design resource, SNP-CRISPR, and released a third major update to our go-to Find CRISPRs resource for sgRNA designs. Our ortholog search tool DIOPT remains our most popular resource. Check out the Online Tools Overview page to view resources for Drosophila reagent identification, data mining, and more.

The TRiP continues to generate fly stock resources for the community. In addition to >13,500 RNAi stocks, we have deposited >4,000 CRISPR-related fly stocks to BDSC for distribution to the community. These include 'toolbox' and sgRNA fly stocks for knockout (TRiP-KO) or activation (TRiP-OE). You can search, nominate genes, and track production of sgRNA fly stocks.

FlyBase Report to the Drosophila Board
10-March-2020

For the past twenty-seven years, FlyBase has provided a centralized resource for Drosophila genetic and genomic data to enable researchers to further their research. Drosophila is one of the premier model organisms and provides cost-effective help in elucidating the etiology of human genetic diseases. FlyBase has three main goals.

1. To continue curation of literature and reagents relevant to Drosophila research, so that researchers can continue to rely on FlyBase to find the latest innovations in the field. We will prioritize curation of data sets relevant to gene expression, cellular functions, signaling pathways, and human diseases, and display the information in an intuitive, integrated, readily searchable format.

2. To improve FlyBase's utility to the human genetics and population genetics communities, by curating and integrating relevant data sets, and developing tools that enable better access to this wealth of data. As a member of The Alliance for Genomic Research (AGR), FlyBase will work closely with other Model Organism Databases (MODs) to integrate data sets and develop tools to enable cross-species analyses. This effort will have a major impact on the fly community, accelerating the development of models of human diseases.

3. To facilitate more integrative analyses and approaches, FlyBase will continue to expand its utility as a platform for integrating and displaying large-scale studies, transcriptomics and proteomics data sets. In addition, FlyBase will improve access and display of tools available within the community, and incorporate the most useful data sets and tools for visualizing complex data sets to enable more researchers to take a more global approach to their genetic research.

April 1, 2020 begins year 3 of our 5-year renewal with NHGRI. As anticipated, our budget was reduced with cuts over the 5-year period of up to ~25% (which normalize to 35%). Our necessary user-fee collection to supplement FlyBase funding continues. As of 27-February-2020 (nearly 2 years since fees were implemented), 422 labs have committed to pay ~$354,937. We have collected $307,507 of this amount.

We are grateful for the strong support from our community.

FlyBase is a mature project with an experienced staff of long-term employees and many of our activities are continuous. In this report, we include minimal descriptions of on-going activities and highlights of new or modified activities, as well as web site usage statistics.

Respectfully submitted on behalf of PIs by
Norbert Perrimon
Susan Russo Gelbart

Many literature curation and high throughput curation activities will continue unchanged. Some highlights are:

- Automated triaging pipeline: We will use the SVM system to flag disease-related papers and integrate this into our triage pipeline.
- Full genetic curation: Curation of data on DNA inserted into the genome by non-transposable element techniques is now part of our regular curation process.
- An emphasis has been placed on genome feature curation and physical interaction curation, with goals of bringing genome feature curation completely up to date, and keeping pace with new physical interaction curation while also addressing the backlog.
- Human disease model curation: Extensive curation effort has been devoted to human disease models, including creation of free-text summaries and capture of the genes, both fly and human, used in these investigations. Recently, there has been an increase in characterization of disease-implicated variants using Drosophila models; while these are highlighted in the human disease model summaries, the growing number has led us to plan for a more integrated view. We are in the process of developing tabulated displays of disease-implicated variants in the context of a specific disease or of a specific gene. These tables will include links to external resources, such as ClinVar, and links to related information in FlyBase.
- We continue allele-based curation of disease models based on the Disease Ontology (DO); this curation is compatible with the approach used by the Alliance. A key aspect of FlyBase DO curation is the capture of genetic interactions that ameliorate or exacerbate a disease-related phenotype.
- We will use available orthology data and expand our representation of orthology calls especially as they relate to the other MODs and human genes.
- Gene Group curation: We have submitted a majority of existing Gene Group data to the FlyBase database, and produced Gene Group Reports and provided searching and browsing of these data on the website. These data will continue to be curated and added to the database.
- We will curate select datasets deemed to be of highest general interest to the FlyBase user community. In the coming year, we anticipate that these will include large-scale reagent collections such as new sgRNA targeting constructs and human UAS-cDNA constructs introduced into flies; expression data, such as RNA-Seq profiles; and data that inform gene models, such as genome-wide screens for polyadenylation sites.
- We will continue to import available Graphical Abstracts in FlyBase references.
- Genome annotation has become relatively stable, however *D. melanogaster* model annotation will continue to be triggered by new available evidence.
- We will continue identification of new lncRNAs, anti-sense lncRNAs and smORFs.
- We will incorporate available transcription start site data into FlyBase using high

confidence RAMPAGE TSSs that do not correspond to annotated transcripts.

- FlyBase will work with groups studying interspecific variation at the genome level in *D. melanogaster* (*e.g.,* the *Drosophila* Genetic Resource Panel - DGRP) to align key information to the reference *D. melanogaster* genome. Work continues on a FlyBase genome variation module.
- With combined funding from VFB (Virtual Fly Brain) and FlyBase, we will continue adding new anatomy terms and enhancing the existing terms by an ongoing review process, with a focus on new neuroanatomy terms and definitions.
- We will continue to review and improve the phenotypic class ontology and, focusing on terms for behavioral, learning and memory phenotypes in collaboration with VFB.
- Development of Chado modules for gene groups and human disease models will be maintained and updated as necessary. Work on new modules of the FlyBase Chado central database will continue.
- We are actively working actively with the Fly Cell Atlas and single Cell Atlas to annotate all Drosophila cell types and curate scRNAseq data sets. We will work with the Human Cell Atlas to integrate Drosophila and human scRNAseq data sets to enable cross-species comparisons.
- We will update the genomic sequences of all Drosophila Genetic Reference Panel (DGRP) strains from release assembly 5 to release assembly 6 in order to produce variant call format (VCF) files to be displayed on FlyBase.org's JBrowse instance to allow quick visualization of single nucleotide polymorphisms originating from DGRP strains in conjunction with FlyBase's extensive genomic annotations.
- We have continued to expand our Pathway page resource, using GO annotation as a basis to compile experimental evidence-weighted lists of genes that encode either core pathway components or pathway regulators. We have added new pathway reports for Activin, BMP, FGFR, PVR, Hedgehog, IMD, Notch and Toll pathways in the current reporting period, meaning we have now completed the first-pass review of 15 major pathways and thereby bringing the first phase of this project to a close. All FlyBase-curated pathway data can now be viewed in a dedicated 'Pathways' section of Gene Reports and can be searched via a new 'Pathways' tab on the QuickSearch tool. As part of the second phase of development, we have added graphical network representations to pathway pages.

Alliance of Genome Resources:
The primary mission of the Alliance of Genome Resources (the Alliance) https://www.alliancegenome.org/ is to develop and maintain sustainable genome information resources that facilitate the use of diverse model organisms in understanding the genetic and genomic basis of human biology, health and disease. This understanding is fundamental for advancing genome biology research and for translating human genome data into clinical utility.

Commented [MOU1]: Spell out

The founding members of the Alliance of Genome Resources are: FlyBase, Mouse Genome Database (MGD), the Gene Ontology Consortium (GOC), Saccharomyces Genome Database (SGD), Rat Genome Database (RGD), WormBase, and the Zebrafish Information Network (ZFIN).

FlyBase staff currently are members on several working groups within the Alliance: 'Disease and Phenotypes', 'Interactions', 'Gene Descriptions', 'Biological Function', 'Variants', 'Basic gene information', 'Orthology', and 'Literature Curation'. In addition, 1 FlyBase member serves at the Alliance Twitter Master. Developers are involved in producing and integrating data for the Alliance website members of the Architecture working group, and setting up Redux state management. Two FlyBase members have served as Alliance Data Quartermasters (responsible for overall dataset integration / liaison between working groups and developers). Significant contributions have been made to the specification and display of basic allele and phenotype data and automated gene summaries in the Alliance database/website, orthology backend (use of specific methods), and data integrity and loading, and help with specifying commonalities in the content/format of data exchange, as well as the display and searching of integrated datae in the Alliance website.

We will continue to contribute to working groups within our remit and areas of expertise.

FlyBase web site production and development will continue as planned:

- Six releases of FlyBase will be produced during the upcoming grant year.
- FlyBase will continue to obtain community input through FlyBase Community Advisory Group, feedback at the US and European Drosophila Research Conferences, input through the FlyBase help desk and from the FlyBase Scientific Advisory Board.
- FlyBase will attend conferences of other research communities (such as other model organism communities and the human genetics community) to advertise FlyBase and to get feedback on how to make FlyBase data more accessible to these communities. We will also continue the production of a series of training videos on the best methods for using, browsing and searching FlyBase.

For the next year, the website team plans to focus their development efforts on observing and listening to users to optimize the usability of the new site. Usage of the site and ideas for improvements will be captured via analytics and other existing outreach efforts. We believe that a major focus should be on improving ease of user experience rather than just adding functionality and increasing complexity. We also plan to provide support for development required by new FlyBase projects. Near term projects include migration of all species and tracks from GBrowse to JBrowse, an updated FTYP tool, and a revamped BLAST tool.

Future development goals:

- Fast Track tool improvements.
- Address server load issues by adding capacity and using a CDN.
- Network display using Cytoscape in pathway reports.
- Complete species and track migration from GBrowse to JBrowse.
- An updated BLAST tool for the latest NCBI BLAST.
- Add commentary displays to pages and enhance display on home page.
- Optimizing usability in FlyBase based on user feedback and observations.
- Provide support for new FlyBase curator projects.
- Continue to expand our use of cloud-based services where it makes technical and financial sense.
- Evaluate open source tools for automating cloud deployment and management.
- Enhancement of public programmatic endpoints (APIs) to improve data access for external collaborations (e.g. Alliance) and advanced users.
- Continue to implement GraphQL for use as an enhanced API endpoint and coordinate common schemas with Alliance.
- Investigate use of GraphQL federation for querying and data retrieval of data from the Alliance and other MOD sites.
- Continue to work with Alliance development teams.
- Continued security improvements in the face of increasing threats.

2019 FLYBASE WEB USAGE

The following are web statistics from the FlyBase website as captured by Google Analytics. Unless otherwise stated, all usage statistics in this document cover the period of Jan-Dec for the years 2015-2018 and Jan-Oct for 2019. In summary, the usage statistics, when compared to the previous year period, indicate that our overall pageviews have declined (-5%), our sessions have gone up slightly (2.8%), and the number of our users has gone up (14%). In addition, data class report and tool usage has not significantly changed from previously observed and well-established patterns.

Pageviews

Figure 1 shows FlyBase pageviews for the previously mentioned time periods. A pageview is defined as a hit to an HTML page, script output or other content that does not include non-document files (CSS, images, JavaScript, etc.). The average number of pageviews for 2019 thus far is 780k, with a high of 946k and a low of 614k. The periodic dips in this plot all correlate with expected holiday patterns. Compared to Jan-Oct of 2018, pageviews are down 5%.

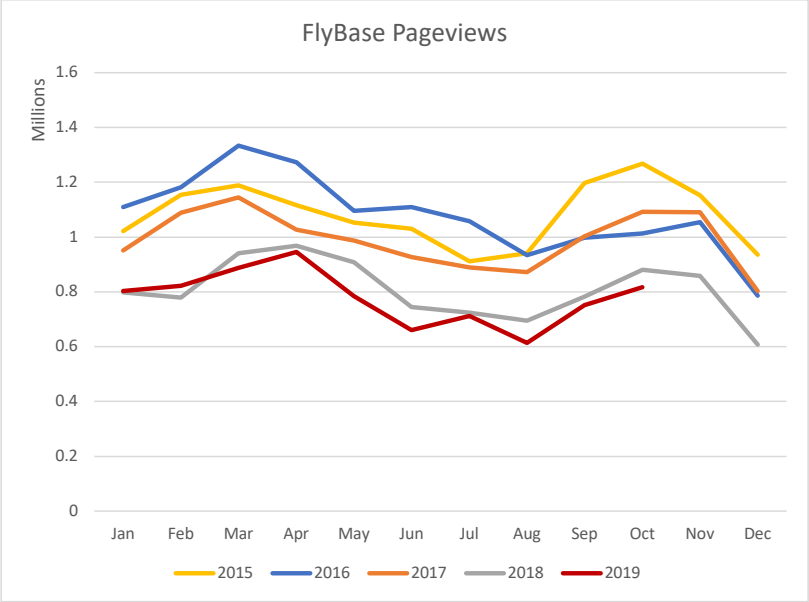

Figure 1 – FlyBase Pageviews for Jan 2015 – Oct 2019

Session: Figure 2 shows FlyBase sessions (visits) for the same period as pageviews. A session is defined as a period of activity by a unique web user. If no activity is recorded for 30 minutes, any subsequent activity is counted as a new session. The average number of sessions for 2019 thus far is 125k, with a high of 154k and a low of 103k. Compared to Jan-Oct of 2018, sessions are up 2.8%.

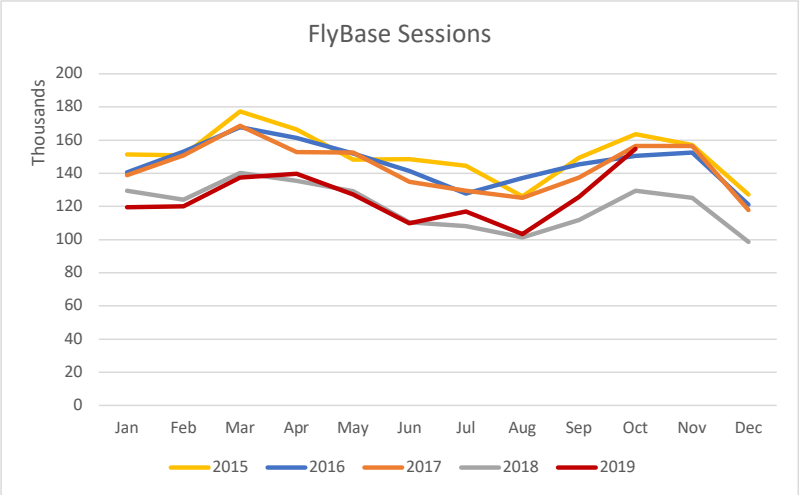

Figure 2 – FlyBase sessions for Jan 2015 – Oct 2019

Users: Figure 3 shows FlyBase users for the same period. A user is defined as a unique session ID that Google analytics generates. This value does not take into account a single user using multiple computers and/or browsers. The average number of users for 2019 thus far is 50k/month, with a high of 61k and a low of 42k. Compared to Jan-Oct of 2018, the number of FlyBase Users are up 14%.

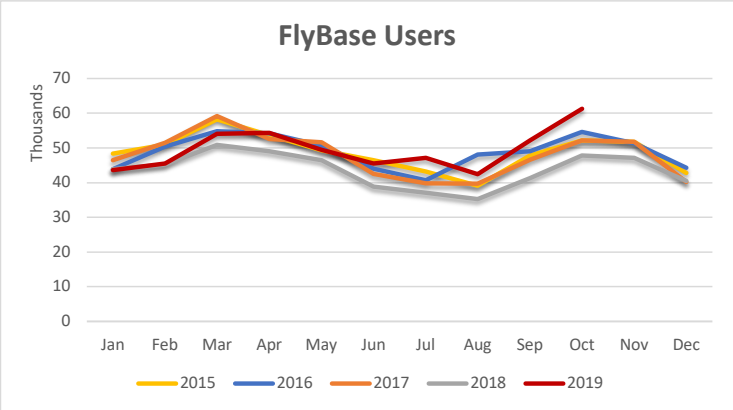

Figure 3 – FlyBase users for Jan 2015-Oct 2019

Data Class Usage: Figure 4, "FlyBase Data Class Usage by Pageviews", shows the total pageviews for FlyBase data class reports for Sep 2018 - Oct 2019. Genes and References still dominate the top two spots and alleles, insertions, constructs, and stocks have swapped positions while remaining at similar levels compared to 2018. Other notable changes include Gene Groups rising 3 spots and Human Disease Model rising 4 spots.

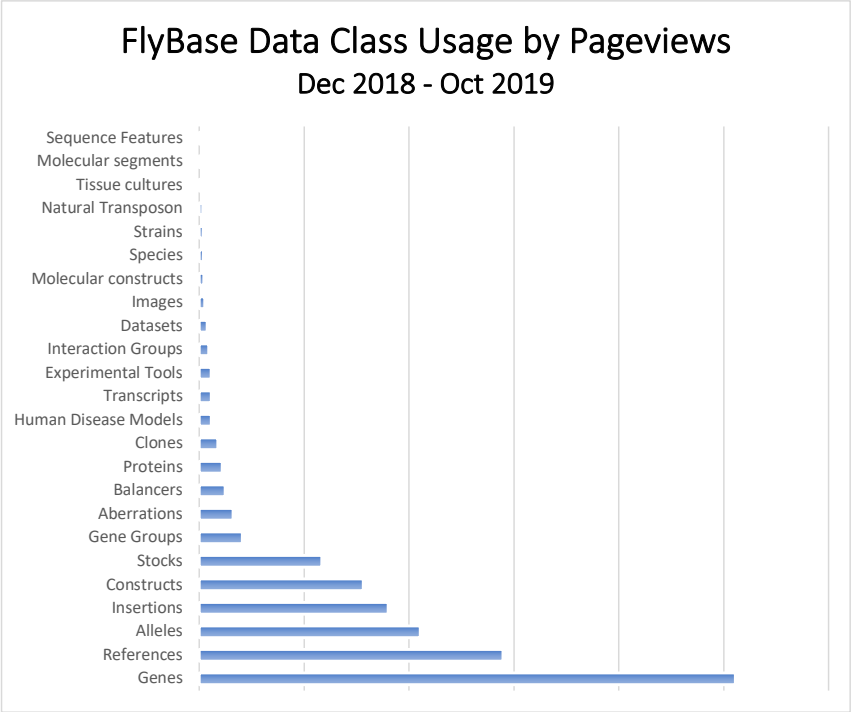

Figure 4 – Pageviews by FlyBase Data Class. Experimental Tools were first introduced on Aug 23, 2018 and fully integrated into gene reports on Aug, 22, 2019.

Tool Usage

Figure 5, “FlyBase Tool Usage”, shows that our top 5 tools are BLAST, Simple Search, Jump to Gene, GBrowse, and Sequence Downloader.

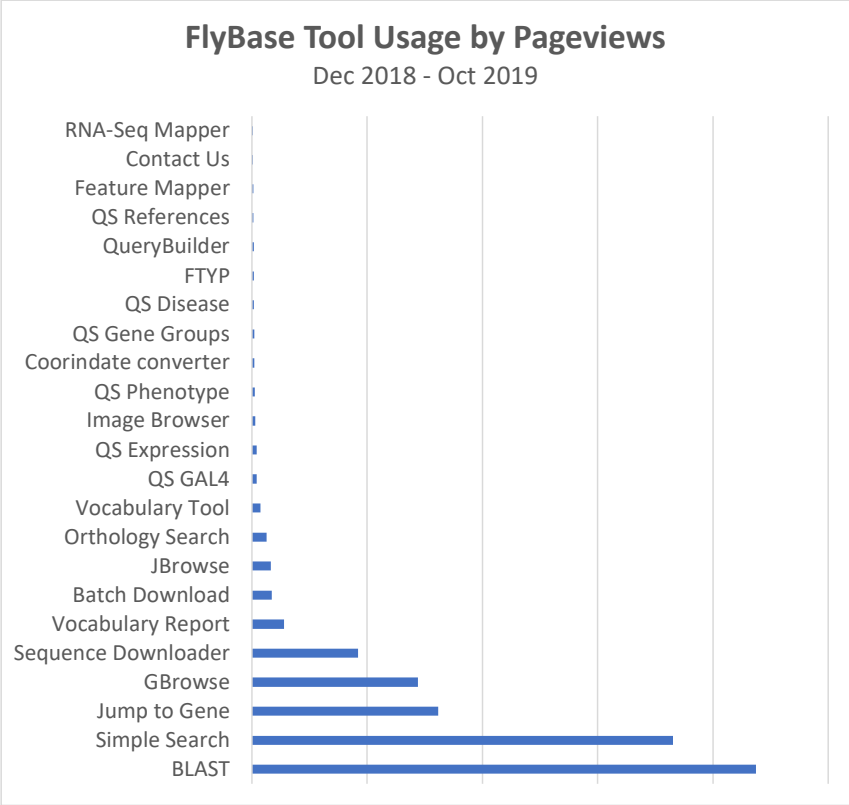

Figure 5 – FlyBase Tool Usage.

Drosophila **community database curation opportunities**

Resource	**Get started at this URL**
FlyBase ‘Gene Snapshots,’ brief summaries of gene function information	https://wiki.flybase.org/wiki/FlyBase:Gene_Snapshots

FlyBase ‘Fast Track Your Paper,’ if not already completed	https://flybase.org/submission/publication/
DRSC/TRiP RSVP database of in vivo RNAi and CRISPR phenotypes, add your data	https://www.flyrnai.org/cgi-bin/RSVP_search.pl
DRSC/TRiP FlyPrimerBank database of qPCR primers, add feedback on primers	https://www.flyrnai.org/FlyPrimerBank
RedFly asks for your regulatory data (they plan to provide additional info in future)	http://redfly.ccr.buffalo.edu/

Gene Disruption Project (Hugo Bellen)
Update of the GDP (Bellen, Perrimon, Zirin, Spradling, Levis, Kanca)

MiMICs constitute the most versatile tools for gene annotation (Venken *et al.*, 2011). A MiMIC in a coding intron of a gene can be converted into a GFP protein trap, facilitating detection of the gene product, allowing affinity purification of protein complexes that include the targeted protein and conditional knock-down of the targeted protein in any tissue in a reversible manner (Nagarkar-Jaiswal *et al.*, 2015) through the use of Recombinase Mediated Cassette Exchange (RMCE, Bateman, Lee and Wu, 2006). Alternatively, MiMICs can be converted into T2A-GAL4 gene traps that generate a strong loss of function allele that expresses GAL4 in the spatial-temporal pattern of the targeted gene (Diao *et al.*, 2015; Lee *et al.*, 2018). The GAL4 can then be used to detect the expression domain of the targeted gene using UAS-GFP and allows rescue of the GAL4-disrupted gene with UAS-cDNAs. Using UAS-cDNA of human orthologs of the gene allows us to assess whether potential human variants are pathogenic and permit to conduct a systematic structure-function analysis of proteins. Given that the MiMICs are transposable elements and their insertion in the genome is nearly random we decided to switch to CRISPR/Cas9 to insert MiMIC-like elements (CRIMIC) to tag conserved genes at precise locations (Lee *et al.*, 2018).

We reported (Lee *et al.*, 2018) the generation of a library of ~1000 *Drosophila* stocks containing *T2A-GAL4* artificial exons in coding introns of genes. ~400 of these genes were targeted by CRIMIC and ~600 were generated by converting MiMICs in coding introns of the genes into T2A-GAL4 gene trap alleles through RMCE. We showed that these strains allow numerous applications based on testing of a sample of this extensive collection. First, ~90-95% of insertions in essential genes cause a severe loss-of-function phenotype, demonstrating that the insertion of the *T2A-GAL4* cassette is an effective way to mutagenize genes. Second, 70% of lethal insertions tested were rescued with a single UAS-cDNA construct, even when many transcript isoforms are annotated in FlyBase. This permits rapid testing of variants and hence structure-function analyses. Third, loss-of-function phenotypes associated with CRIMIC *T2A-GAL4* insertions that are flanked by FRTs can be reverted by excision with *UAS-flippase.* This has many advantages: mutagenic cassette can be excised to generate a control genetic background to test whether observed phenotypes are specific to the loss of function caused by the *T2A-GAL4* gene trap, akin to use of precise excision of a P element as a genetic control of alleles generated by P element insertions. It also shows that the CRIMIC library does not contain other lethal or visible mutations introduced by CRISPR/Cas9 while generating the allele. Importantly, excision of the mutagenic construct also allows assessment of the need for the gene product in any given tissue, as the *flippase* can be expressed tissue specifically using another binary expression system. Fourth, *GAL4* driven *UAS-GFP/RFP* reports tissue and cell-type of gene expression with high sensitivity. This has allowed us to report the expression pattern for hundreds of previously unreported genes. In summary, these stocks comprise a very powerful resource and integrate many elegant features that allow assessment of gene function, protein structure, and many other features that are dependent on the *GAL4* system. Given its great versatility and usefulness, we decided to continue to expand the development of this library using a modified approach (see below).

The Drop-in technology

The main issue with our original protocol for generating CRIMIC T2A-GAL4 lines (Lee et al. 2018) is not conceptual but technical. There are two hurdles: cost and efficiency. dsDNA homology donors for insertion of large cassettes in Drosophila typically require stretches of at least 500–1000 nt of homology on either side of the SIC (Beumer *et al.*, 2013; Bier *et al.*, 2018; Diao *et al.*, 2015; Lee *et al.*, 2018; Rong and Golic, 2000; Zhang, Koolhaas and Schnorrer, 2014). The large size of the homology regions affects cloning efficiency of the donor constructs using Golden Gate cloning (Lee *et al.*, 2018). We typically achieve an overall success rate of ~50%: ~80% cloning success rate; ~80% transformation rate and ~80% rate of proper integration as assessed by PCR (0.8 X 0.8 X 0.8 = ~50%).

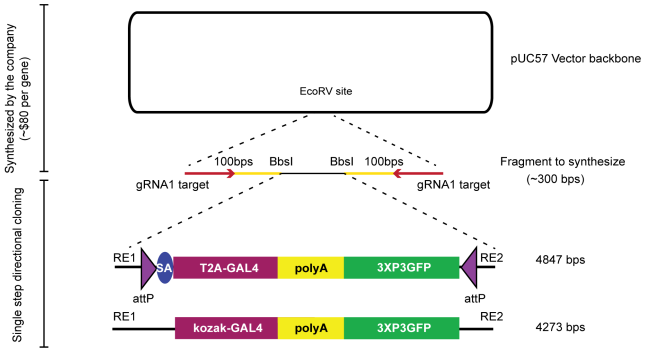

Figure 1. Single step cloning of CRIMIC donor constructs using Drop-in. An intermediate containing gene specific homology arms and a restriction cassette is commercially synthesized. A single step cloning of SIC in this intermediate is sufficient to prepare a homology donor construct. FRT sites are

We therefore dedicated a lot of time to developing an alternative strategy that makes use of commercial gene synthesis methodology to integrate full length CRIMIC cassettes of ~ 5 kb flanked by short homology arms. We first established that linearizing homology donor vectors *in vivo* increases the CRISPR mediated homologous recombination rate even with short (100 nts on each side) homology arms. However, even when the homology arms are shortened, synthesis of a full length 5 kb CRIMIC cassette is cost prohibitive (>$1,000) especially for a thousand genes.

To reduce the cost, we synthesize a gene-specific small plasmid carrying a gRNA1 target-100nt left homology arm-Restriction cassette-100nt right homology arm-gRNA1 target in a pUC57 backbone (cost of synthesis is $80). The SIC containing the *attP-FRT-T2A-GAL4-polyA-3XP3-EGFP-FRT-attP*, which is shared for all target genes, is then subcloned directionally into this plasmid in a single straightforward cloning step, replacing the restriction cassette with the SIC (**Figure 1**). We refer to these constructs as the Drop-in int100-CRIMIC constructs (Kanca *et al.*, 2019b). The construct is injected together with two gRNAs: gRNA1 to excise the double stranded donor from the donor vector and a gene specific gRNA to cut the target site in the genome (**Figure 2**).

We injected vectors containing the full length CRIMIC cassette in more than 20 genes and obtained an 75% success rate as verified by

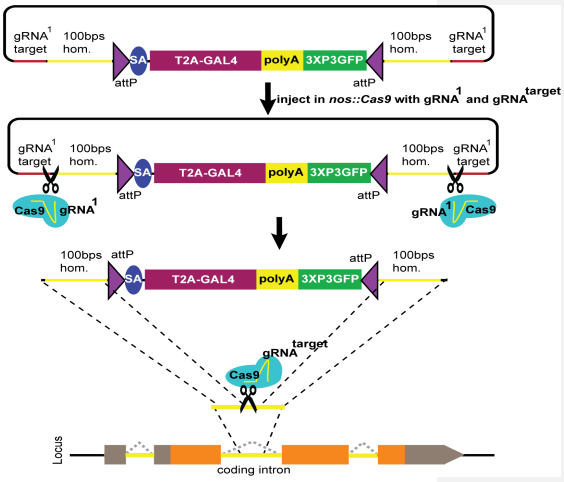

Figure 2. In vivo linearizing the homology donor permits use of short homology arms for integration of large constructs. Linearizing the homology donor construct through gRNA1, a gRNA that does not have a target in the genome, boosts homologous

PCR. Hence, 100 nt homology arms are sufficient to integrate large SICs into target sites in flies because simultaneous excising the homology donor as a dsDNA fragment *in vivo* and cutting the chromosomal target region by Cas9 strongly promotes precise homologous recombination. Note that for Drop-in constructs, the presence of short homology arms allows the donor vectors to be synthesized cheaply and introducing larger SICs is done by a single straightforward cloning step, providing a good balance between ease of construct generation and efficient *in vivo* use. Hence, we improved the overall efficiency compared to the CRISPR method described in (Lee *et al.*, 2018) by 1) shortening the homology arms to 100 nt, which allows synthesis followed by a simple cloning step, thus eliminating cloning failure and reducing the cost from ~$300 to ~$100 per construct; and 2) integration of target gRNA1 sites on either side of the SIC to linearize the donor with Cas9 in the germ cells and boost integration efficiency and precision. This reduces the total cost per gene by about 40%, from an estimated $1,500 per gene to an estimated $800-900 (Kanca *et al.*, 2019b).

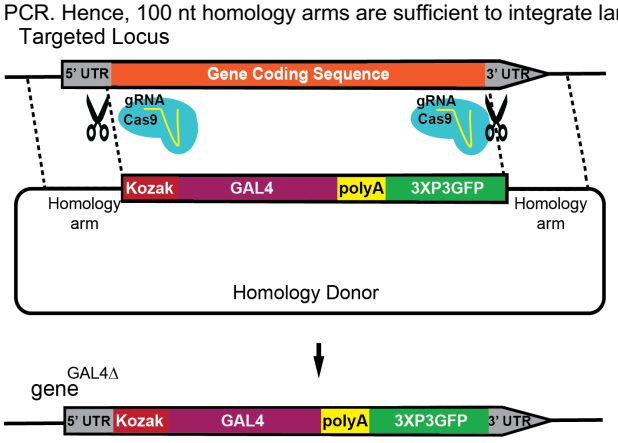

Figure 3. Kozak-GAL4 replacement of CDS of genes. Genes without a suitable intron can be targeted, using kozak-GAL4 replacement to generate a knock-out allele that expresses GAL4 in the targeted gene's expression domain

Note that 45% of the genes have no suitable introns to integrate a CRIMIC cassette (Claire Hu, personal communication). Hence, we recently adapted the Drop-in method to replace the coding region of genes that do not contain suitable introns with kozak-GAL4-polyA-3XP3-EGFP (**Figure 3**). We select a gRNA target in the 5' UTR and one in the 3'UTR of the GOI. We then replace the entire coding region of the gene with a cassette that contains a Kozak consensus sequence followed by the GAL4 ORF. We have shown that this cassette can be efficiently used to rescue the fly mutation induced by the removal of the ORF by driving a UAS-human cDNA for several of the genes tested including *Tom70* (a mitochondrial importer; (Dutta *et al.*, 2020); and *wdr37* (a cytoplasmic WD40 repeat protein; (Kanca *et al.*, 2019a). We have currently tested this strategy for 31 genes and obtained precise integrations for 22 upon a single injection. We are very confident that we can increase this efficiency by tweaking with the injection conditions and anticipate an 80% success rate in the near future. We have currently generated

7,434 MIMIC
608 MIMIC GFP
620 MiMIC T2a-GAL4
1200 CRIMICs
22 Kozak GAL4

Bateman, J. R., Lee, A. M. and Wu, C. T. (2006) 'Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange', *Genetics,* 173(2), pp. 769-77.

Beumer, K. J., Trautman, J. K., Mukherjee, K. and Carroll, D. (2013) 'Donor DNA Utilization During Gene Targeting with Zinc-Finger Nucleases', *G3 (Bethesda),* 3(4), pp. 657-664.

Bier, E., Harrison, M. M., O'Connor-Giles, K. M. and Wildonger, J. (2018) 'Advances in Engineering the Fly Genome with the CRISPR-Cas System', *Genetics,* 208(1), pp. 1-18.

Diao, F., Ironfield, H., Luan, H., Shropshire, W. C., Ewer, J., Marr, E., Potter, C. J., Landgraf, M. and White, B. H. (2015) 'Plug-and-play genetic access to drosophila cell types using exchangeable exon cassettes', *Cell Rep,* 10(8), pp. 1410-21.

Dutta, D., Briere, L. C., Kanca, O., Marcogliese, P. C., Walker, M. A., High, F. A., Vanderver, A., Krier, J., Carmichael, N., Callahan, C., Taft, R. J., Simons, C., Helman, G., Network, U. D., Wangler, M. F., Yamamoto, S., Sweetser, D. A. and Bellen, H. J. (2020) 'De novo mutations in TOMM70, a receptor of the mitochondrial import translocase, cause neurological impairment', *Hum Mol Genet.*, 29(9), pp. 1568-1579.

Kanca, O., Andrews, J. C., Lee, P. T., Patel, C., Braddock, S. R., Slavotinek, A. M., Cohen, J. S., Gubbels, C. S., Aldinger, K. A., Williams, J., Indaram, M., Fatemi, A., Yu, T. W., Agrawal, P. B., Vezina, G., Simons, C., Crawford, J., Lau, C. C., Chung, W. K., Markello, T. C., Dobyns, W. B., Adams, D. R., Gahl, W. A., Wangler, M. F., Yamamoto, S., Bellen, H. J., Malicdan, M. C. V. and Network, U. D. (2019a) 'De Novo Variants in WDR37 Are Associated with Epilepsy, Colobomas, Dysmorphism, Developmental Delay, Intellectual Disability, and Cerebellar Hypoplasia', *Am J Hum Genet,* 105(2), pp. 413-424.

Kanca, O., Zirin, J., Garcia-Marques, J., Knight, S. M., Yang-Zhou, D., Amador, G., Chung, H., Zuo, Z., Ma, L., He, Y., Lin, W. W., Fang, Y., Ge, M., Yamamoto, S., Schulze, K. L., Hu, Y., Spradling, A. C., Mohr, S. E., Perrimon, N. and Bellen, H. J. (2019b) 'An efficient CRISPR-based strategy to insert small and large fragments of DNA using short homology arms', *Elife,* 8:e51539.

Lee, P. T., Zirin, J., Kanca, O., Lin, W. W., Schulze, K. L., Li-Kroeger, D., Tao, R., Devereaux, C., Hu, Y., Chung, V., Fang, Y., He, Y., Pan, H., Ge, M., Zuo, Z., Housden, B. E., Mohr, S. E., Yamamoto, S., Levis, R. W., Spradling, A. C., Perrimon, N. and Bellen, H. J. (2018) 'A gene-specific T2A-GAL4 library for Drosophila ', *Elife,* 7:e35574.

Nagarkar-Jaiswal, S., Lee, P. T., Campbell, M. E., Chen, K., Anguiano-Zarate, S., Gutierrez, M. C., Busby, T., Lin, W. W., He, Y., Schulze, K. L., Booth, B. W., Evans-Holm, M., Venken, K. J., Levis, R. W., Spradling, A. C., Hoskins, R. A. and Bellen, H. J. (2015) 'A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila', *Elife,* 4:e05338.

Rong, Y. S. and Golic, K. G. (2000) 'Gene targeting by homologous recombination in Drosophila', *Science,* 288(5473), pp. 2013-8.

Venken, K. J., Schulze, K. L., Haelterman, N. A., Pan, H., He, Y., Evans-Holm, M., Carlson, J. W., Levis, R. W., Spradling, A. C., Hoskins, R. A. and Bellen, H. J. (2011) 'MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes', *Nat Methods,* 8(9), pp. 737-43.

Zhang, X., Koolhaas, W. H. and Schnorrer, F. (2014) 'A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in Drosophila', *G3 (Bethesda),* 4(12), pp. 2409-18.

Drosophila Humanization Project (Hugo Bellen, Sue Celniker, and Shinya Yamamato)
A collaboration with Coral Warr and Travis Johnson in Australia and Toshiyuki Takano-Shimizu in Japan .

Our labs have been generating UAS-human cDNA clones and corresponding transgenic lines for the past four years. These human cDNAs have been very useful to determine if putative pathogenic mutations in human genes found in rare disease patients affect protein function. The Model Organisms Screening Centers (MOSC) of the Undiagnosed Diseases Network consisting of worm, fish and fly labs is tasked with performing functional studies of genes and variants identified in probands who are screened in 12 academic clinical sites throughout the US to aid in the diagnosis of some of the most difficult to solve medical mysteries and to the discovery of new human genetic diseases. Whole Exome Sequencing (WES) and WGS (Whole Genome Sequencing) typically identify several candidate variants in a substantial number of probands but providing compelling evidence that they are pathogenic is not trivial. In the past three years the MOSC Drosophila Core has provided critical functional data for more than 30 novel human diseases in collaboration with the UDN as well as other clinical research consortiums such as the Centers for Mendelian Genomics and the CHARGE Consortium (Cohorts for Heart and Aging Research in Genomic Epidemiology). Publicly available reagents we and many others generated, especially mutants, MiMIC and T2A-GAL4 CRIMIC insertions, RNAi collections, BAC and fosmid constructs, and UAS-fly and human cDNA stocks have all played very prominent roles in the success of these and many other collaborative endeavors in *Drosophila* labs around the US and the world. These studies have provided the first insights into the *in vivo* molecular functions of evolutionarily conserved genes and allowed us and others to unravel the pathogenic mechanisms underlying some rare and common human diseases using flies.

Functional replacement of a fly gene with a human ortholog (humanization) has been successful for many genes, allowing disease gene discovery, determination of allelic severity, and in some cases leading to the identification of drugs that are being tested in clinical settings. A simple approach to humanize a fly gene is to insert an artificial exon consisting of a SA(splice acceptor)-T2A-GAL4-polyA in an intron flanked by two coding exons shared by most or all transcripts of the fly ortholog of a human gene of interest (GOI). Alternatively we replace the entire ORF (open reading frame) with a Kozak-GAL4 constructs. Both typically disrupts the function of the gene severely and lead to the expression of GAL4 in the proper spatial and temporal expression pattern. These GAL4 lines now allow two key experiments; A) determining the expression pattern of the gene using UAS-nls::mCherry (nuclear) or UAS-CD8::GFP (membrane), and B) assessing rescue using the UAS-fly or human cDNA. If the latter experiment succeeds, patient variants can be tested for functionality in the context of the human protein. Another approach is to immediately test if an RNAi knockdown of the fly GOI can be rescued by co-expression of the reference UAS-human cDNA or variant forms, which are resistant to knock down due to primary sequence variation between human and fly genes. A third approach is to take advantage of available mutants and rescue their phenotypes by ubiquitous or tissue specific expression of the human protein.

The limiting step of humanization experiments is often the lack or high cost of full length UAS-human cDNA constructs and generation of fly strains. Centralized creation of this library and corresponding transgenic flies greatly facilitates these experiments and drives fly biologists to work on human genes. In addition, variants previously identified in patients can be quickly tested *in vivo* if the plasmids are readily available for end users. This library will also allow to systematically establish gene function conservation based on experimental data, probe if other human genes with related functions rescue the mutant phenotypes, provide functional data about human proteins for other model organisms, allow structure-function studies, validate the specificity of RNAi reagents, permit use of the cDNAs in other species, and allow testing if overexpression of the human proteins are toxic in flies.

The Celniker lab analyzed ~8,000 expression-ready Gateway compatible human cDNA clones that were derived from the Mammalian Gene Collection (MGC) to assess whether the clones contained the canonical full-length isoform of the human GOI. We obtained 6,654 clones representing 5,878 human genes. These clones were graciously provided by Drs. Marc Vidal and David Hill at Harvard University. Based on this analysis, we selected full-length cDNA clones that correspond to 4,336 genes. More recently, the Celniker lab obtained the entire hORFeome 8.1 collection (http://horfdb.dfci.harvard.edu/) which consists of 12,846 clones, and an additional 3,423 clones from the hORFeome Collaboration collection (http://www.orfeomecollaboration.org/) representing 12,103 genes. Moreover, the Bellen and Yamamoto labs obtained a collection of ~16,000 human ORF clones that that are part of the Ultimate ORF clone collection assembled by Invitrogen (now ThermoFisher)) to complement the MGC and hORFeome clones from the late Dr. Kenneth Scott at Baylor College of

Medicine. The Ultimate ORF clone collection are also Gateway compatible, and similar to the clones from MGC all clones have been fully sequence validated (www.thermofisher.com/us/en/home/life-science/cloning/clone-collections/ultimate-orf-clone-collection.html). MGC clones are “open clones” that lack their endogenous stop codons that allow C’ tagging when the expression vector contains an epitobe tag sequence prior to the translational and transcriptional termination sequences, whereas the Ultimate ORF clones carry stop codons. All these clones have been replicated and shipped to the Celniker lab to be analyzed and added into the cloning pipeline. For our initial studies we used the pUASg-HA.attB vector to generate ~200 expression clones. However, relatively early in the project we moved to the pGW-HA.attB vector (GenBank: KC896837.1; Bischof et al. (2013)). This allows the users to exchange the 5’ and 3’ sequences that flank the ORFs (Bischof et al., 2013). We proposed to generate 8,500 human UAS-cDNA clones that corresponds to the 10,500 fly genes that have human homologs based on DIOPT (https://www.flyrnai.org/cgi-bin/DRSC_orthologs.pl). We have currently cloned more than 5,000 of the 8,500 gene.

Of the constructs generated thus far the Bellen and Yamamoto labs generated and balanced ~2,600 transgenic lines. >1,500 lines have been donated and available from Bloomington Drosophila Stock Center, and the rest are in the process of validation or transfer. While performing this work in the USA, we became aware that Dr. Toshiyuki Takano-Shimizu’s lab at Kyoto Institute of Technology (KIT) had also initiated a similar project: cloning human cDNAs in UAS vectors and creating transgenes with support from a Japanese funding agency. Dr. Takano-Shimizu is also involved in the rare disease research initiative in Japan called IRUD (Initiative in Rare and Undiagnosed Diseases), using Drosophila to assess variant function like the MOSC. Although their scale was much more limited, we decided to collaborate and they agreed to use their resources to enhance the production of the transgenic flies.

We now share constructs that have been subcloned by the Celniker lab and >1,000 transgenic line are already available from the Kyoto Stock Center. Finally, our collaborative international consortium also included Dr. Coral Warr and Dr. Travis Johnson’s labs of Monash University in Melbourne, Australia. We established a collaboration as they had also initiated a similar project. They generated about 500 UAS-cDNA constructs as well as the corresponding transgenic stocks. They decided to stop the project when we approached them. In sum, nearly 4,000 transgenic stocks have been generated and are in the process of being deposited or are already available. We plan on generating another 3,000 transgenic stocks at BCM and Takano-Shimizu at the Kyoto Stock Center will continue to in inject plasmids that are being generated at LBNL, with our aim to generate a total of ~7,500 lines for the fly community..

The stocks that have been shipped to BDSC from BCM are searchable through the following website.
http://flypush.imgen.bcm.tmc.edu/humanfly/
BDSC also has assembled a webpage that catalogs all human gene related resources.
https://bdsc.indiana.edu/stocks/uas/uas_hsap.html
A specific page on the DGRC website to search for human cDNA clones is under construction in the “Collections” page.
https://dgrc.bio.indiana.edu/clones/Catalog#
The stocks that are available from Kyoto Stock Center are searchable from the following website.
https://kyotofly.kit.jp/stocks/documents/Humanized_fly_lines.html

Berkeley Drosophila Genome Project (Susan Celniker, Ann Hammonds, Ken Wan, Erwin Frise)

A. Introduction

The BDGP was established in 1992 to sequence the Drosophila melanogaster genome. We've continued to expand activities with the goals of improving the functional annotation of the genome and expanding community resources. Our microbiome genome sequence of the L. plantarum genome and four associated plasmids is available at NCBI and published in Genome Announcements. The L. brevis genome strain, BDGP6 is now publicly available at NCBI (CP024635.1) and is considered the RefSeq (Dec 2019). We continue to characterize the transcriptome (smORFs). We are also continuing the modENCODE project rebranded as modERN to map transcription factor binding sites and transcription factor knock-downs using RNAi following by RNA-seq. The data will be available from the ENCODE DCC. Finally we continue to use the cDNAs to generate resources for proteomics studies and as templates for probes to determine spatiotemporal gene expression patterns in the embryo.

B. Reference Genome sequence

After completion of the Release 6 genome sequence, our efforts to improve the genome are centered on incorporating the PacBio long-read whole genome shotgun assembly (MHAP) into Release 6 with the goal of producing an integrated consensus assembly that will become Release 7 with improvements to the heterochromatin and the Y chromosome. There is currently no budget for these studies and they have not progressed since reported in 2019.

C. Reference Microbiome Genome sequence

As part of an LBNL funded program we sequenced the microbiome of the reference genome strain, y;cn, br, sp. These are complete genomes sequenced using the PacBio platform and include conjugative plasmids and virions. They were automatically annotated using the RAST and GenBank annotation pipelines. We cataloged protein-coding genes, RNA genes including rRNA operons, tRNAs, pseudogenes and prophages. It would be valuable to consider having them at FlyBase. We suggested this last year but there has been no

movement as of yet.

D. cDNA Clone Resources

We maintain our clone resources which have not substantially changed from the 2018 report as a collection available for DGRC to request if they need back-ups and occasional fill requests for clones not yet available from the DGRC. The exception is the production of a human ORF collection for expression in flies. We are working with Dr. Hugo Bellen's group on this resource.

Table 1. Summary of Human Expression Clones.

Table 1. Summary of Human Expression Clones.

Collection	Vector	Promoter	N-term Tag	C-term Tag	ORF Stop Codon?	System	Past year (3/2019-3/2020)	Total
hGUHO	pUASg-HA.attB	UAS	--	3xHA	No	Gal4-UAS	0	153
hGUHO	pGW-HA.attB	UAS	--	3xHA	No	Gal4-UAS	1849	4173

D. Embryonic Gene Expression

We continue to collect embryonic spatiotemporal gene expression data from high throughput in situ hybridizations using the Gold Collection clones as templates for RNA probes. Annotations assigned by stage to each gene are now included in the FlyBase gene reports. In addition to the wild type gene patterns, we are collecting expression patterns for selected CRM-driven reporter constructs from the Rubin/Janelia collection and additional constructs generated as part of our collaboration with the Berkeley Drosophila Transcription Network Project. We incorporate the CRM experiments into the public database (http://insitu.fruitfly.org) with links to the FlyBase sequence feature reports for these constructs. Our homepage includes a separate browse tab for the CRM experiments to improve accessibility. Our improved gene reports include graphical summaries of the stage specific organ system annotations and a graphical representation of the associated modENCODE RNA-seq data. The updated version also allows searches by all known gene name synonyms and human ortholog names. We continue to add new search

and discovery tools based on computational image and annotation analysis. We published an advanced
method for modeling spatially local gene interactions and networks with our dataset. An interactive
viewer based on the annotated patterns of 708 site-specific transcription factor genes, using selforganizing
maps to show relationships among transcription factor expression patterns in the context of
organ system development, can be accessed at http://insitu.fruitfly.org/som. We are active participants in
the development of image analysis within the open source image analysis platform FIJI (fiji.sc). We are
starting to use our recently finished open source microscope automation software for automated slide
loading and imaging with commodity hardware. To date we have completed and annotated experiments
for 8538 genes and 336 CRMs documented with over 182,432.

E. ENCODE model organism Project - modERN (Bob Waterston, Susan Celniker, Kevin White,
Valerie Reinke and Mark Gerstein)

The modERN (model organism Encyclopedia of Regulatory Networks) project is an independent R01
submitted to complete the study of fly and worm transcription factors (those defined as having a currently
recognized DNA-binding domain) determining their genomic DNA binding sites in animals using the
ChiP-Seq assay as was perfected in ENCODE. To date the Celniker lab has produced 383 transgenic
GFP tagged-TF fly lines. They are deposited at the Bloomington Stock Center. The White Lab has
performed ChiP-Seq for 538 lines. The data is being processed through the ENCODE pipeline and is
being distributed through the ENCODE DCC. In addition, we produced TF knock down RNAi followed
by RNA-seq experiments for a number of TFs [43 sequenced (~1000 RNA samples)]. The validated
RNA-seq files have been submitted to the ENCODE DCC and are in their process to be made available to
the community. We hope to have a manuscript soon describing the putative targets of these TFs.

F. Other Resources

In an effort to improve the quality of our web-based user support, we continue to make changes to our
website (http://www.fruitfly.org) including: updated FAQs, updated protocols and an updated design to
make it easier for users to navigate to the relevant information.

We continue to work with FlyBase to improve gene and transcript annotations. We submit clones to the
DGRC molecular stock center for distribution to the community.

G. Technology
cDNA and expression clone sequencing continues to rely heavily on the ABI3730xl capillary sequencer.
Characterization of the transcriptome as part of the modENCODE project has primarily been on the
Illumina GAII and HiSeq platforms. We note that sequencing technology continues to evolve rapidly, and
access to the latest instruments is essential to our mission. LBNL's Life Sciences Division owns a MiSeq,
which is located in our lab, providing us with an R&D platform. We have the Oxford Nanopore platform
and software running in the lab and it was used to sequence some of the microbes from the Drosophila gut
microbiome. We have access to the latest Illumina machines through the UCB QB3 sequencing core.

H. Funding
The BDGP is funded almost exclusively by NIH grants (NIGMS). An R01 (SEC) funds the spatiotemporal
gene expression studies and was renewed in 2019 for two years. The mechanism to fund such resource
grants has changed and we need to figure out what type of grant we should write. A RO1, "Systematic,
Genome-Scale Functional Characterization Of Conserved smORFs" (Celniker, PI and Perrimon co-PI) was
obtained to functionally characterize genes that may or may not be coding proteins that have small open
reading frames (<100 aa) and are conserved from flies to humans. Manuscripts for this work are in
preparation and we have asked for a NCE and plan to resubmit next year. We are also funded under
subcontracts from the University of Washington (R. Waterston, PI, Celniker and White, co-PIs) to
participate in a consortium performing ChIP-seq analysis of transcription factors and RNAi knockdown in
embryonic development and from Baylor College of Medicine (Bellen, PI, Celniker, co-PI) to construct
human ORF clones for expression in flies.

DIS Report (Jim Thompson)

Drosophila Information Service volume 102 was published in early January with a large number of reports submitted in calendar year 2020. Since first being published in 1934, we welcome research reports, new mutants, teaching exercises, large data archive records, and other reports annually. DIS (cited in bibliographies as Dros. Inf. Serv.) is freely available at www.ou.edu/journals/dis. Although we publish one annual issue at the end of each calendar year, submissions are accepted at any time. The firm submission deadline is 31 December for each calendar year volume. Manuscripts are preferred electronically in MSWord and can be sent to jthompson@ou.edu. James N. Thompson, jr., Department of Biology, University of Oklahoma, Norman, OK 73019.