

Genome Insider S3 Episode 5: Work With the JGI! Tips for a Winning CSP Proposal

Menaka: Imagine this:

Dan Udvary: It's a beautiful morning.

Menaka: Ooh – let's do fake yawns to set the scene

Dan Udvary: *fakes a yawn*

Menaka: *fakes a yawn*

Menaka: And you awaken from your slumber with an incredible research question.

Dan Udvary: One that's elegant, and insightful!

Menaka: Your question definitely leverages genomics —

Dan Udvary: But it could be in any number of fields – at least that relate to energy and the environment.

Menaka: Maybe you research plants that could become more sustainable fuel sources,

Dan Udvary: Or organisms that could convert those plants into fuels and products,

Menaka: Maybe you want to look at entire communities of microbes that cycle nutrients or carbon!

Dan Udvary: Or perhaps, like me, you're specifically interested in the chemicals that microbes can produce.

Menaka: The JGI supports researchers in all of these lanes, with a variety of resources.

Dan Udvary: And so today, we're talking about how researchers can work with the JGI.

Menaka: By submitting proposals — about those incredible research questions — to our Community Science Program. This is a program open to researchers at any career stage, from anywhere in the world, for free! So let's get into how that program works.

Menaka: This is Genome Insider, a podcast from the Joint Genome Institute. I'm Menaka Wilhelm, and today I'm here today with Dan Udvary, a secondary metabolite-ologist who hosts another podcast from the JGI. And we're together to chat about how researchers can submit proposals, and work with the JGI in that Community Science Program, which we often shorten to CSP.

Dan Udwy: So, this is a podcast crossover!

Menaka: It is! Of Genome Insider,

Dan Udwy: Uh huh, and Natural Prodcast.

Menaka: Yeah! So, we're excited,

Dan Udwy: I'm very excited! This is the first time we've recorded something together, so this is fun!

Menaka: Yeah it's true!

Dan Udwy: What do we want to tell our people about CSP's? I think, the question that generally comes up, there's a little bit of confusion about what a CSP is, as opposed to other different kinds of like federal grants, like through NIH or NSF that people would normally be applying for,

Menaka: Right, because usually, writing a research proposal means requesting support for a project, in the form of money. But the JGI is not a funding agency. We're a user facility – so an actual lab, in Berkeley, California, with all kinds of sequencing, and 'omics, and bioinformatics capabilities. So proposals at the JGI work a little differently –

Dan Udwy: Mm-Hmm. You know, we don't, we don't give out money. Instead we give out capacity and do the work that you need done.

Menaka: And users don't pay for that work. It's funded by the Department of Energy. And we accept new projects several times a year, in these proposal calls for the Community Science Program. A few proposal calls have deadlines coming up – in January, March, and next spring. You might call it proposal season.

Dan Udwy: and so, I thought it would be fun for us to work together to talk about this kind of stuff, and talk to a few people at JGI who can maybe give us some better answers.

Menaka: Yeah, and those people will be Tanja Woyke, who runs user programs at the JGI, and Miranda Harmon-Smith, who, as a project manager, helps shepherd CSP projects along. So they'll have lots of good details for us, and there's even more information on the JGI website, that's JGI dot DOE dot GOV. And without further ado, let's hear from Tanja!

Dan Udwy: Here we go!

Tanja Woyke: Hi there!

Menaka: And first, Tanja, if you could just introduce yourself.

Tanja Woyke: Hi there. My name's Tanja Woyke. I'm the Deputy Director for the User Programs at the JGI, and I'm also the head of the microbial genome program, and I'm a scientist here and have been for 18 years. .

Dan Udvary: All right. Easy start. What does CSP stand for?

Tanja Woyke: Ah, great question. CSP stands for Community Science Program. It used to actually stand for Community Sequencing Program, also CSP, same acronym, but as JGI as a user facility evolved and as we started to do more than just sequencing, we thought the name community sequencing program was really not appropriate anymore. And so we changed it to science program.

Menaka : How cool. And I think that's a nice segue for the next thing we wanted to ask you about, which is basically, so many projects at JGI start with sequencing — (listen to our audio tour of the JGI's sequencing lab here!), but where else do those sequences go? Like what are the other, the other scientific tools we offer people,

Tanja Woyke: We offer, in addition to sequencing, we offer metabolomics and also DNA synthesis. And we considered these capabilities a little bit more from sequence to function since we can start to tackle more functional questions. For example, we can synthesize a gene and then figure out what it might be doing through some sort of downstream functional assay.

Menaka : And are there any specific things you'd point to that researchers can do in a project at the JGI that they couldn't elsewhere?

Tanja Woyke: Yeah, I would say a lot of things. So we like to consider ourselves not a core sequencing facility. In a core sequencing facility, you can just, you know, send your DNA and get some sequence back. We provide a lot more than that.

The first one, I would say is the complexity of the capabilities that we have, and that you can request. So you can request a mixture of, for example, metatranscriptome sequencing, metagenomics, single cell genomics. We even do the flow cytometry for that. Again, metabolomics, DNA synthesis. And then also the scale is pretty unique in that we don't just sequence one microbe for you, we sequence thousands for you, or thousands of metagenomes. And I think also the quality that we provide. So for fungal and for a lot of our eukaryotic genomes, all our eukaryotic genomes, fungal, plant, algal genomes, a lot of effort goes into generating these data sets. There's a lot of manual work in terms of the assembly, the curation, QC, annotation of these data sets for our eukaryotes, we, in addition to sequencing the genome, we always sequence RNA alongside. That really helps us to determine, you know, where genes start and end and really help improve our annotation.

So a lot of effort goes into these products that we offer that I think a core facility would not be

able to offer you. So there's a lot of, we also have a lot of analysis pipelines. Again, metagenome, metatranscriptomes for SNP analysis. If somebody has a re-sequencing project, for example, people do experiments where they evolve a population in the lab under certain stressor conditions, and then they see how does this genome change over time? And we have pipelines for that. So again, we don't just give them the data output, but we tell them which genes may be affected by that. And how, and, of course, we have tools also and Dan knows very well, about this for secondary metabolite analysis, which is really great. It just provides, again, more than just a data output. It, it helps the collaborator to interpret the data and pursue their science more efficiently.

Dan Udvary: Yeah, I'm just gonna make the point that, you know, secondary metabolism, very interdisciplinary kind of field. And so having access to all the technologies that the JGI has can, can really benefit secondary metabolism projects specifically for my audience. But yeah. So Tanja, what makes a research project a good fit for working with the JGI?

Tanja Woyke: Very important first criteria for being a fit is that it meets the DOE mission criteria. So we are funded by DOE BER and work within the DOE mission space of bioenergy, carbon cycling, biochemistry, bioproducts. So we cannot work on health related projects, we really need to fit, a project needs to fit under this umbrella. Now, this is still a pretty broad umbrella. I would say even within these broad areas, there are few areas that are less of a focus. I would say, for example marine research is not so much of a focus within BER anymore. We are really focused on terrestrial ecosystems, but we work on, in coastal environments, brackish environments, et cetera.

Menaka : And what about the scale of projects – how much work should researchers be aiming to propose?

Tanja Woyke: So for our annual CSP call, we do like to get proposals that are really multifaceted, that do request diversity of the different capabilities that we have that really exploit these capabilities for the PI's research questions. We don't wanna get a proposal that's just asking for 20 microbial genomes. That is, that is not a good fit for our annual CSP proposal call. Now we have smaller calls, a new investigator proposal call where smaller proposals may be a fit. Still not 20 microbes. We still, we work in plates, right? So just sending us 20 DNA samples is also not a good fit for JGI. Even for the New Investigator CSP, we need to work on a certain scale, right? So there's a minimum. A PI should submit a fairly complex proposal, ideally multi-investigator and generating, I think this is important too, generating data sets that don't just benefit that one PI, but that have utility more broadly, right. Across a group of scientists, across a scientific community that goes far beyond just a single investigator PI. We would like our data to get used over and over again. Looked at from different angles, different questions asked, and published over and over again. But again, with different research questions in mind.

Dan Udvary: Awesome answer.

Menaka : And who's eligible to submit CSP proposals? We talked about PIs, but,

Tanja Woyke: So for most of our calls, and I'm gonna make the exception here for the New Investigator CSP call, anybody is, is really eligible. You can be from any country, you can be in any career stage. It's important to have information that shows that you can really execute on this project. So if you're a grad student, if you don't have a track record yet, you need to build a team around this project, right? Right. So you need to have expertise on the front end and back end.

Reviewers are gonna look really closely. Are they gonna deliver? Do they have funding and do they have the expertise to actually do the sampling that they propose and go to, you know, these sometimes pretty elaborate places where you need sampling permits and significant budget to get these samples. Do they have expertise and resources to extract the nucleic acids or collect the cell material and get that shipped to JGI?

And then they look at the other end, once JGI generates the data, do they have the resources and expertise to do the bioinformatics and to synthesize these results into publication? So this is really important because we don't want the data to just sit there or we don't wanna start a proposal and then we can't get the samples in.

So that said, a lot of the proposals specifically for the annual CSP and for the FICUS call have research teams that are larger. They have a broad, broad expertise, and often they team up with others. Right. And then the team together makes a really strong proposal and a strong research team that has a high chance of success.

Now coming back to the CSP new investigator call for that, we do have a requirement that the PI who's submitting the proposal cannot have been a PI with the JGI previously. So that's really important. The whole purpose of this call for us is to bring in new users, new investigators who haven't worked with us before.

Menaka : Yeah, that makes sense. How much can proposers expect to collaborate with JGI staff?

Tanja Woyke: So I, I think it depends on the bandwidth of the staff at the time that the data analysis is needed. So we do collaborate with various PIs quite closely on their CSPs, but we can't collaborate with every CSP or FICUS user because we don't have the bandwidth for that. So I would say usually if, if we get requests for help, we, we try to help. I know it also depends on the product and what's being generated for our eukaryotic genomes, the fungal, algal and plant programs and JGI staff and HudsonAlpha staff are a lot more hands on.

Dan Udvary: Right, right. And I can speak to secondary metabolism and that you know, I know that our group has quite a bit of interaction because probably mostly because there aren't too many CSPs focused on secondary metabolism yet, because we're at a very new science program, I can promise lots of interaction if secondary metabolism proposals go through for sure.

Tanja Woyke: Yeah. Yeah. And the beauty about secondary metabolism I find is that even though it may not be the original intent in the proposal, depending on what sample you're gonna sequence and what comes out of it, and you, if you start looking at biosynthetic gene clusters, you may find a wealth and some really interesting results in the data. And there could be some really interesting research in secondary metabolites coming out without it even being the original intent of the proposal.

Dan Udvary: Exactly. Yeah. Great, great. Okay. So I guess we should talk about the proposal and submission process and maybe you'd want to give like a really high level overview of what that looks like.

Tanja Woyke: Yeah. So we have on our website, we have a lot of good information on each call. What are the deadlines, for our large call, we have letters of intent. So there's a letter of intent deadline. And then once the letter of intents are accepted, you, you get time to write the full proposal. For our smaller calls, for the CSP Functional Genomics call and the CSP New Investigator call, we don't have letters of intent, because they're generally smaller proposals, they're not as involved, I would say. And each call gets updated regularly, depending on the directions that BER is taking. Also in line with JGI strategic directions, these change once in a while. So I encourage PIs that plan to submit a proposal to really carefully read through these details.

Menaka : And at a more zoomed in level – what should people keep in mind when they're ready to head to the website and submit a proposal?

Tanja Woyke: Yeah. Really start clicking the 'Submit Here' button. And that brings you to a little interface where you can start filling in the details. The title, the PIs, the teams, definitely upload the CVs of all the team members. The reviewers, scientific reviewers look very closely at this material that's being submitted, and they're really, they're asked by us to just review these proposals based on what's uploaded. Right. Don't read between the lines. So it's really important to submit everything most completely.

And then there's different sections that describe the work's scientific merit and why it's important. The DOE mission is one of these sections, a very important section. We wanna be sure that the PIs can articulate why their work falls within one of these DOE mission areas. And that's a really important one. We wanna generate data sets that have broad utility and broad interest of the research community, and we want the PI to articulate that as well.

Menaka : Yeah, that makes sense. And to help the broadest research community, all JGI data is publicly available online. But whoever proposed a project gets access to that data first, before the world does – can you talk a little bit about how and when that data gets released?

Tanja Woyke: Yeah, absolutely. So we do have a new, or I'm gonna call it newer data release policy. And that policy is, essentially states that we provide a one year embargo to sequencing

data and metabolomics data. For example, for a microbial genome, that means an annotated genome that's integrated into our Integrated Microbial Genomes system, which is our comparative analysis system. Once that is completed, the clock starts to tick, and then for one year only the PI and whoever the PI, all the co-PIs, collaborators or who, anybody else who the PI wants access to, has access to that data. But otherwise, it's, it's hidden. Nobody sees it, nobody knows it exists, really. And then once that one year is over, the data gets released on our portals, such as IMG/M, but we also then push the data to GenBank. So the raw data goes to GenBank as well as the assembly, the annotations, et cetera, depending on the product.

Menaka : Got it. And because that data is publicly available, other researchers around the world can use it to advance other projects. Is there anything else you want researchers to know, if they're thinking about submitting proposals?

Tanja Woyke: Yeah, I think one thing I want users to know is that if they have questions, please come talk to us. We are a user facility, we're very user oriented, and we wanna help enable your science. So if you have questions, be it about the DOE mission, is, is my project DOE mission relevant? Please come talk to us. There's information on the website. I'm always available if there are technical questions on the proposal submission or if, if something is unclear, contact us. We're, we're super happy setting up a quick zoom call, hop on the call with you and, and discuss your project and, and help you. And also down the road, if you are a user. And there are questions, we're always here. Each project has assigned a project manager who shepherds the projects through and they're there for, for the users, for you, to help you.

Menaka : And next, we'll be talking to a project manager who does just that! Miranda Harmon-Smith will fill in more details about proposal review, and how projects work once they're approved. First, a quick break – for some acronym action.

Menaka: Let's meet our brave contestant, who we'll be quizzing on a bit of JGI jargon.

Simon Roux: Hi, I'm, uh, Simon Roux, research scientist in the metagenome program and lead of the viral genomics and VEGA group. We are trying to better understand what viruses are out there and what they're doing.

Menaka: Excellent – and we're got, really, more of an abbreviation today – it's EcoFAB. So it's spelled E-C-O-F-A-B, and the options are:

Ecologically fabulous
Fabricated Ecosystem
Eukaryotic Communities Facilitating Adaptation Biotically

Simon Roux: I would say it's two, but I wish it was one. I think one was great and, and we should, we should think about this.

Menaka: Great answer. Let's meet our EcoFAB explainer.

Amber Golini: My name is Amber Golini, I'm a research associate in Trent Northern's metabolomics lab, and I work on sample processing and extraction for metabolomics research.

Menaka: Ok Amber, take it away!

Amber Golini: Yeah, so EcoFABs are short for fabricated ecosystems. So basically there are small single-plant-scale growth chambers. Picture a small rectangular plastic box, so probably about, a little bit smaller than the average smartphone. And so it's a clear plastic box, inside that chamber you have a single plant growing.

Menaka: And here's why you'd want a single plant to grow in a tiny plastic box.

Amber Golini: These lab scale devices have been developed to study environmental microbiomes and to further understand soil-plant microbe interactions in a reproducible and controllable environment.

Menaka: So the idea is that when researchers use a controllable chamber to grow and study plants, they'll be able to compare experiments and share methods, even when they're working in different labs around the world. And the JGI can send these EcoFAB chambers around, to researchers working to understand plants, microbes and soil. Fabulous! Even though that's... not part of their name.

Menaka : Alright – back to our chat about the Community Science Program. As a refresher, I'm here with Dan Udway -

Dan Udway: Hey, Menaka!

Menaka : And next, we're chatting with project manager Miranda Harmon-Smith.

Miranda Harmon-Smith: So I've been a project manager for 10 years, but I've worked at the JGI for going on 20 years.

Dan Udway: Excellent. Yeah. What were you doing before you were a project manager?

Miranda Harmon-Smith: So when I joined the JGI, I started out working in the lab doing Sanger sequencing.

Dan Udway: Oh, really? Okay.

Miranda Harmon-Smith: Okay. Then when Sanger sequencing went away and 454 sequencing went away and we were just, you know, focused on, on PacBio and some other Illumina, I transitioned over into the project management team.

Menaka: Awesome – well let's get into CSPs. We heard from Tanja about the beginning of a CSP proposal, and now let's talk about the next phase – what happens after researchers hit Submit.

Dan Udvary: Yeah. So who reviews proposals and, and what scoring criteria do they use?

Miranda Harmon-Smith: Sure. So we have scientific reviewers that are generally current or past users. So they've had a successful proposal, so they've been through the process. We also have reviewers who have not had any proposals with us. And that's okay because they have expertise in, you know, sequencing or synthesis or some of the other capabilities. And based on that and sort of their expertise, we match them up with proposals to review. And we assign three reviewers. It generally takes about a month before we review the proposals. So we give the reviewers a chance to, you know, read them and make their comments.

So during the review process, again they're assessing the projects for, for scientific merit, for DOE mission relevance, technical feasibility, the capability of the applicants themselves. And then they give it a, a final score. And the scoring is between a one or a five. So one is the best, and we take the average of the three reviewers as the final score. And in general, we accept proposals that are in the range of 1 to about 2.5. And it works out to be about 40% of the proposals that are submitted.

And once we get back the final approval from the DOE, then we accept them in our, in our database. And that sends out a notification to the submitting PIs that their proposal has been accepted or rejected. And then and then if it's been accepted, there's an additional internal review that happens for the proposals that requested synthesis.

Menaka: Great. And requesting DNA synthesis is particularly relevant, because when this episode publishes, our next proposal call is the Functional Genomics call – where researchers might be requesting DNA synthesis. Can you talk a little bit about what kinds of work researchers could request in those submissions?

Miranda Harmon-Smith: Right. So the capabilities that we offer are synthesis of single and multiple genes and pathways for functional characterization, synthesis of combinatorial pathways, libraries for metabolic engineering, synthesis of single guide RNA libraries. We do strain engineering using CRAGE compatible microbes. They can request data mining. So if they're not quite sure of what genes to request synthesis for, we can help them select the target genes through mining our IMG database, our MycoCosm database for fungi, Phytozome for plants. So we have a group that will work with users to identify target genes. If they need help with that, they can also request metabolomics, so they can send polar and non-polar metabolites to the JGI. (Find out more in a webinar here.)

Dan Udvary: You didn't mention sequencing,

Menaka : Of course sequencing, and what else

Miranda Harmon-Smith: , I'm thinking. Yeah, mapping of transcription factor binding sites. So that's sort of a new capability that's been offered. So we synthesize the transcription factors and then hand them off to the DAP-seq group. And then there is RNA sequencing that can be requested just for the purpose of testing gene function. So it's not like the normal sequencing requests that can be requested through the other CSP proposals where you're, you know, sending metagenomes to be sequenced or, you know, various microbes or fungi to be sequenced. This is specifically RNA sequencing. And then there's a new capability, the EcoFAB devices that the plant program can send out to users so that they can do their plant microbiome interaction studies and then they can send, send those back to the JGI and, you know, we'll run the, the tests and, and provide them the data. So yeah, those are all the capabilities. And again, they're on the jgi.doe.gov website.

Menaka : Cool.

Dan Udway: Maybe this is a more scientific question or maybe a little more secondary metabolism focused. Is there a size limitation on how big a piece of DNA we can sequence or synthesize rather?

Miranda Harmon-Smith: The product offerings that we have are we can do less than 5 Kb genes. And the timeline for those are about 90 days. So once we order, it goes through our process, we can have it delivered to you within 90 days, about three months. We offer five to 10 Kb. We offer greater than 10 Kb, we've done up to 100 Kb. And it basically adds on like a month to the cycle time as you go up in size.

Dan Udway: So after all that, your project gets accepted, when does the work on a project begin?

Miranda Harmon-Smith: Right, so as soon as the proposal has been accepted and you've been notified, we work with you to set up an initiation call. During the initiation call, you're gonna meet with the program leads, you're gonna meet with the project managers, and we invite the submitting pi. Any of the collaborators or co-PIs are welcome to join that. And so that's really the time when we're going to develop the statement of work, which is really just defining the entire project scope from beginning to end. We'll talk about project requirements, expectations on the user's end, on the JGI's end, an overall plan of action, the type of deliverables, the timeline and schedule, communication plan, and then what closing out the project looks like. But throughout the whole life cycle of the proposal, it's a very collaborative process where we're in, you know, constant communication to let them know what stage their project is at, answer any questions that they have. We maintain copies of everything that we complete so we can always test it again on our end resequence it to verify it was the right thing.

Menaka : Cool. Can you give us a sense of the range of a life cycle for a project? Like if you submit at the end of January a CSP, what kind of timeline would you expect for the life of a project?

Miranda Harmon-Smith: A proposal is approved for three years. So you have three years to send in your sequences or send in your DNA or RNA, whatever you're sequencing. On average projects generally are a year to two years long. There's variations within proposals where, you know, people ask for single guide RNA libraries and single gene constructs. Obviously those are gonna take longer. And then if you ask for other capabilities outside of synthesis that include metabolomics. So now with all these other capabilities that have been offered, most of the proposals coming in, they're asking for, you know, all of the above .

Menaka: And if someone is wrapping up a CSP proposal but interested in doing more work with the JGI, can they submit another proposal while they're still involved in an existing CSP project?

Miranda Harmon-Smith: Oh yeah, absolutely. And we have a lot of repeat users and you know, we, we welcome that. And if you have a proposal that was just approved and you have another idea for a second proposal and you wanna submit that during the next cycle, that's fine too.

Menaka: So don't hold back.

Miranda Harmon-Smith: . Yeah, no, absolutely.

Menaka: So that wraps up this chat about the CSP program – but there's plenty more information online. We've got links in the show notes, and you can head to JGI dot DOE dot GOV as well.

This episode was written and produced by me – Menaka Wilhelm, with production help from Massie Ballon, Allison Joy, and Ashleigh Papp.

If you liked this episode, help someone else find it! Tell them about it, send a link over, or leave us a review wherever you're listening to the show.

Genome Insider is a production of the Joint Genome Institute, a user facility of the US Department of Energy Office of Science located at Lawrence Berkeley National Lab in Berkeley, California.

Thanks for tuning in – until next time!