Genome Insider S4 Episode 7: The Megadata of Lake Mendota - Part 2: Souped Up Computing

Menaka: Over twenty years ago, Trina McMahon got interested in a freshwater lake. She's a microbial ecologist at the University of Wisconsin, Madison, and the lake is Lake Mendota. And there were other researchers who had already been studying this lake for years.

Trina McMahon: They study the fish, they study the plants, they study everything else, but the microbes. And so I come along and say, "Hey, I'd like to study the microbes," and they say "Great, just tag along." And so we have access to all their infrastructure, all their historical data, we have archives of samples going back to 1999, even for the microbes.

Menaka: So Trina started working with the JGI to sequence and analyze all of those samples. It was a twenty year collection. A ton of information. And because these are environmental samples, each sample includes many, many organisms at once. The JGI would assemble these samples into metagenomes. Computationally, that's a very complicated task. But thanks to developers and data scientists at Berkeley Lab and the JGI, there is an assembler program capable of working at this scale. It's called MetaHipMer.

Emiley Eloe-Fadrosh: I can say this, you know, very confidently, there is no other place in the world that can do these types of metagenome assemblies.

Menaka: To find out what MetaHipMer is, why it's called that, and how that assembler came to be — go back an episode. Part 1 of this series covers all of that.

This episode is part 2. And today, we'll cover the part of a project where these programs actually run. We're headed to the kinds of computers that do a MetaHipMer assembly, like that metagenomic data that Trina McMahon's lab collected from Lake Mendota.


Menaka: This is Genome Insider from the US Department of Energy Joint Genome Institute. Where researchers discover the expertise encoded in our environment — in the genomes of plants, fungi, bacteria, archaea, and environmental viruses — to power a more sustainable future. I'm Menaka Wilhelm.

And this is part 2 of a three-part series on how a giant dataset from Lake Mendota comes together. The project sampling Lake Mendota set out to get a close look at an environment, and how it's shifting in the face of climate change.

And by sampling the environment, rather than specific organisms, Trina McMahon's team built a giant, multi-year metagenomics dataset.

Her team sampled routinely for two decades – a giant timescale. With that kind of time, they get to look at lots of how lots of factors affect the lake — precipitation, invasive species, and climate

patterns.

And to handle so much information, we needed a new scale of tools, too. So — last episode was all about the assembly programs that have risen to meet these datasets at the JGI. It took roughly a decade for developers to create a parallel assembler capable of handling a dataset like the one from Lake Mendota. So that's the assembler called MetaHipMer.

And today, we get to the hardware that it takes to run this program. This is a parallel program, so it needs multiple computers working together — that means a cluster, or a supercomputer. And these can be very big. When the JGI assembled the Lake Mendota dataset, that happened at a supercomputer called Summit, at Oak Ridge National Lab, in Tennessee. Summit is as big as two tennis courts.

And while I didn't make it to Tennessee to see what two-tennis courts of computing looks like, I did make it to another cluster that sits right by the JGI. It's also capable of running MetaHipMer assemblies — ones that are on the smaller side.

This is a cluster called Dori. So we'll get to that in a bit.

But to start this episode, I want to sort out the difference between supercomputers and normal computers. The JGI's Chief Informatics Officer, Kjiersten Fagnan, helped out. We met at a whiteboard, so she could illustrate.

To explain supercomputers, she draws 9 squares on the board. Those squares are in a 3 by 3 grid. It looks… honestly a lot like a square belgian waffle. But for our purposes, each of those squares is a computer.

Kjiersten Fagnan: What you've got — is like a set of computers that you can network together that all have one CPU and one chunk of memory.

Menaka: So on the board, our waffle grid of squares is a small cluster, or a supercomputer. In supercomputer speak, each individual computer, or square on the board, is a node. It's got processing power and memory.

Kjiersten Fagnan: And so what makes a computer super, or like what makes a super computer is really the interconnect.

Menaka: So now, to illustrate that interconnect, Kjiersten draws information moving between this waffle grid of computers. It starts to look a lot like maple syrup, flowing between the squares of a breakfast waffle.

Kjiersten Fagnan: So every node is connected to every other node, every computer's connected to every other computer.

Menaka: At this point the waffle of computers on the board is totally drenched in syrup.

Menaka: So you just eventually have like free-flowing syrup if you have a super, super, supercomputer.

Kjiersten Fagnan: Totally. Yes. And that is like free flowing information between all of these nodes.

Menaka: And now, it's time for me to waffle a bit – because waffles make a fun analogy, but of course, a supercomputer is more complicated than breakfast. Nodes are connected physically, and also via specific software. And supercomputers come in different sizes — so sometimes they have hundreds, or thousands of nodes. But the main point — is that all supercomputers, or clusters, use multiple computers to do a big job.

And that's key — because remember, researchers working with the JGI are often putting together giant metagenomic datasets. They need parallel computing. Because to handle assembling genomes this way, you can't really just slice up the job and divide it among computers. Multiple computers need access to the same information at once –  which is why that interconnect between nodes, or the syrup on our waffle, is so important. With that kind of connectedness, supercomputers can do all kinds of things – from modeling the coronavirus behind COVID-19, to simulating the climate, and of course, assembling these giant metagenomic datasets.

So next, I wanted to hear from someone who's run these metagenome assemblers for the JGI. More specifically, the person who took the Lake Mendota dataset from Trina McMahon's lab, and ran it through the MetaHipMer2 assembler that we covered in the last episode.

Robert Riley: I'm Robert Riley. I'm a data scientist in the Genome Assembly Group at the JGI.

Menaka: Robert has been at the JGI for fourteen years. Right now, he mainly works on assembling and analyzing genomes from fungal, algal and metagenomic datasets. He works with smaller datasets, and bigger ones.

Robert Riley: I typically get the request to do the assembly and I'll just pull all the data together from the proposal and I'll look at it and make sure it all looks good and do any  quality filtering that needs to be done. And then I'll run MetaHipMer.

Menaka:  Running MetaHipMer happens at Robert's command line, where he starts the job at a supercomputer, remotely.

Robert Riley: These days it's the NERSC Perlmutter cluster. Although sometimes we'll use the Oak Ridge Summit or  Frontier Cluster if it's a bigger data set.

Menaka: And the Lake Mendota dataset was a bigger one. So it ended up at the Summit cluster

at Oak Ridge National Lab.

Menaka Wilhelm: And if we trace that project — so the samples were collected in Wisconsin at Lake Mendota. And then samples came to the JGI in California. And then the assembly was run in Tennessee at Oak Ridge. Yes. And then analyzed - that assembly was analyzed in California at the JGI.

Robert Riley: That's right.

Menaka Wilhelm: And back to Wisconsin

Robert Riley: Where they're doing the more detailed analysis of the bins now. Yes.

Menaka Wilhelm: Yeah.

Robert Riley: It's neat. And it's a collaboration of two national labs, Oak Ridge and Berkeley, as well as the University of Wisconsin.

Menaka: So of course, I wanted to know if Robert had ever been to any of the supercomputers he calls up via the command line.

Robert Riley: Not in person. NERSC maybe, but not Oak Ridge. Yeah.

Menaka: Because when Robert and I spoke, I had recently visited that cluster I mentioned at the JGI. It's called Dori, and it's certainly not as big as Summit. It's more like a little cousin to those gargantuan supercomputers.

Robert Riley: Well actually I have run MetaHipMer now on Dori, in the last few weeks, I was able to coassemble some datasets.

Menaka: And that's neat – because MetaHipMer is an interesting assembler for JGI users to have access to, and also, because Dori is a JGI-managed cluster. And it might seem straightforward to run an assembler on a new cluster, since it's a program,  but it's a little more complicated.

Robert Riley: The software needs to be sort of fine tuned to the specific cluster that it's running on because there are a lot of in intricacies in how the compute nodes, you know, when you're assembling a metagenome on a thousand nodes, the nodes have to talk to each other and know what the other nodes are doing to some extent. So, that's very peculiar to the individual cluster. But nevertheless, we were able to assemble some, good, medium size datasets on 63 nodes of the Dori cluster.

Menaka: And – so – for a sense of the kind of computer that can run MetaHipMer, we'll head next to Dori – that's after the break!

BREAK

Allison Joy: The JGI supported this project via the Community Science Program. This program provides genomic resources for projects with Department of Energy relevance. And we accept proposals from scientists at all career stages.

Menaka: 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 Udwary: You know, we don't give out money. Instead we give out capacity and do the work that you need done, right?

Menaka: And users don't pay for that work. It's funded by the Department of Energy.

Allison Joy: You can find out more about submitting proposals to the JGI on our website, head to joint gino.me/proposals. We've also got links waiting for you wherever you're listening to this episode. Either in the episode description or the show notes.

Menaka: This is Genome Insider. So we've separated supercomputers from computers — they're dozens, to thousands of computers that can work together on a parallel program, and we've heard how data scientist Robert Riley runs the parallel assembler MetaHipMer on giant supercomputers from afar. Now, it's time to see a small version of the kind of computing cluster that can run MetaHipMer. This is a system based at Berkeley Lab.

And our tour starts with CIO Kjiersten Fagnan, the same person who helped explain supercomputers with waffles. She set up the tour of Dori that we're about to embark on — I joined a few folks from the Department of Energy and they kindly allowed me to record. But someone else is leading our walk through.

Kjiersten Fagnan: This is Georg Rath, he is our systems infrastructure lead. I don't know if you all have gotten to meet him before. I will just say he's why JGI's infrastructure is functioning. And so, I know he won't say that himself or he might, I don't know, just depends on the day of the week.

Georg Rath: Nah, there are a lot more people than me.

Menaka: Georg did not take credit for the entirety of the JGI's infrastructure running. But he did introduce us to Dori, the computing cluster.

Georg Rath: It has, now it has a hundred nodes and every node has 64 CPUs and half a terabyte of ram, which is quite a bit. Your laptop has 32 gigabytes, it's like 15 times your laptop.

Menaka: Originally, they actually planned for Dori to be bigger. But then, because of the kind of work Dori will do, they shifted sizing a bit. Dori is geared for assembling biological datasets. So, specific to that work, it actually needs more memory than nodes. The final Dori design has fewer nodes, more storage at each.

Georg Rath: We basically folded 200 nodes together into a hundred. They're the same size overall, but it's just fatter nodes. And that's been working pretty well.

Menaka: With a bit of background, we're ready to go see the cluster itself. It's nearby the JGI, but not inside the main JGI building.

So we're climbing a bit of a hill to get to this cluster — a little more hike than walk. Because Berkeley Lab on the whole is on a giant hill – up here, we can see a lot of the Bay Area.

Georg Rath: Take a small detour to enjoy the view. Yeah.

Menaka: There's Berkeley and Oakland down below, and then San Francisco across the bay.

Tour group: Nice view. Nice. Yeah.


Menaka: While we walk over, Georg handles a key question – where did the Dori cluster get its name?

Georg Rath: It's not the fish, it's also not the dwarf from Lord of the Rings. The reason that Dori is called Dori is because I let people vote on the name.

Menaka: It's true. This happened at a JGI All-Hands meeting a few months ago. There were many options besides Dori.

Georg Rath: We had Cluster Mcclustyface in the, in the running, but it didn't make it, what can I say?

Menaka: The votes went logical, rather than silly. There was previously a cluster called Cori. C-O-R-I. And they decided the next cluster should follow suit, as Dori.

Georg Rath: Truth is, it's Cori with a D. Mm-hmm. because it's the next step after Cori.

Ramana Madupu: That's what I thought. That's what I thought. But somebody told me it's a fish.

Menaka: That fish comment is from the JGI's program manager at the DOE, Ramana Madupu. I think she's nodding to the fish Dory, from the acclaimed children's film "Finding Nemo," as well as the sequel, "Finding Dory." Fitting for our adventure.

As we head inside, we've got to be careful to go into the right building. Because like many buildings at Berkeley Lab, this one's been around for a while, and lived through lots of different arrangements.

Georg Rath: It's a very confusing building because, it's, in reality, it's two buildings and the floor numbers don't match, between one building and the other. But, uh. Finding Dory. That's the other fish.

Todd Anderson?: See that fits!

Menaka: The person filling in the final gap on that joke is Todd Anderson. He is also at the department of energy, where he oversees several program managers including the JGI's Ramana Madupu.

So – in this basement of the correct building, we're finding our way toward Dori. You can tell we're getting close to some serious computing, because the HVAC is louder by the second. It takes quite a system to keep lots and lots of computer nodes cool – we'll hear more about that in a minute. First, we find our second tour guide.

Gary Jung: Hey!

Georg Rath: Hello.

Gary Jung: Good to see you.

Georg Rath: Likewise.

Menaka: That'll be Gary Jung, and he is the scientific computing group lead at Berkeley lab.

Gary Jung: It gets loud inside, so I wanted to do the talking out here before we walk in.

Menaka: I know it sounds loud already, but trust me – we're still in the quiet part of this basement. We're outside the doors of a big data center.

Gary Jung: This data center is an interesting thing in itself, It's a 5,000 square foot data center.

Menaka: That's a little bigger than a basketball court. We're here to see just one cluster, Dori. It's really just a little part of a big computing center.

Gary Jung: And so this is where we keep all of our institutional scientific computing. So it's a

national facility. But here we make this available for the researchers across the laboratory. And then we also run systems for research projects or facilities that need their own dedicated computations, for example, like the JGI. So this room has actually been in use since the sixties.

Menaka: And Gary's seen a fair amount of this room's long life. He was here when this room had some of the earliest supercomputers that ever existed.

Gary Jung: When I got here in 1979, there was originally a CDC 6600, a 6200 and a 7600. The 7600 came here in 1970, and so,

Menaka: Those are all computer models, from the Control Data Corporation. They were big, almost room-sized computers, and they were a very big deal.

Gary Jung: At one time, this housed, essentially, the fastest systems on the West Coast.

Menaka: Not too shabby. Gary has also seen equipment from this facility move around some. Computers have come and go between Lawrence Livermore National Lab, and another facility in downtown Oakland. And eventually, Georg Rath, our first tour guide, was also part of all of that.

Georg Rath: Yes. And to close that circle, I moved the last equipment out of, of Oakland to NERSC when I started here six years ago.

Gary Jung: All right, this is great. I like this. I like the tag team. This is perfect.

Menaka: So now, this data center is set up to give researchers a way to do scientific computing that's more demanding than what a laptop could handle, but not giant enough to warrant a really big national lab supercomputer like Summit or Perlmutter.

Gary Jung: So, we'll walk inside, we'll take a look.

Menaka: Get ready – we're entering into the full force of the data center's hum.

Gary Jung: Yeah, yeah.

Tour guides Gary Jung (left) and Georg Rath (right) show the tour group the cabinets holding compute nodes for a cluster like Dori.

Menaka: As we walk in, we're looking at giant, refrigerator-sized black cabinets. These are all filled with chassis that are home to a bunch of computer nodes.

Gary Jung: But essentially you're looking at each chassis that's about this high has four compute nodes in it. And then in total, there's, across these three racks, there's a hundred compute nodes and each compute node has 64 cores. So just in these three racks, you're

looking at 6,400 compute cores.

Menaka: For context, my laptop has 4 cores. And there are many, many racks in this room. You'd think it would be hot – but that would be terrible for all of these systems. There's a super dialed-in set up where air and water keep everything cool. That starts at the floor, which is actually raised 18 inches above ground level. And in that foot and a half, air is constantly being cooled.

Gary Jung: So all the cold air is pressurized in the floor and comes up in front of the systems, and then the systems take the air in into the front, and then they go through the systems, a cooling rear door, and then it goes up into the ceiling.

Menaka: So that's the whole room. And each rack that's full of computers also has its own water cooling heat exchanger.

Gary Jung: The hot air comes from the back of the computers and it goes through essentially what's like a radiator. And when you run treated water through these, and they have fans, essentially it's just like a radiator on your car and it takes the heat off. So you can actually put your hands on the back of this and you could feel how cold, how cool it is. And then what we could do is that we can actually open it up and then you could feel how warm it is in the back.

Menaka: And it sounds like a wind tunnel if you're right in front of it.

Gary Jung:Yes! Yeah.

Menaka:  But the goal is for this data center to do its job with as much energy efficiency as possible. Lots of these ideas came from researchers at Berkeley Lab who work on this kind of problem.

Gary Jung: And so we actually pioneered a lot of these energy efficient technologies for data centers in this data center.

Menaka: So – that's the data center where Dori lives. It's big, and loud, and historical, in a way. Dori sits where many very famous supercomputers once sat. But remember, in supercomputing terms, Dori is a very small cluster. For really big jobs, data scientists call in other supercomputers, like Perlmutter or Summit.

Those supercomputers handle terabytes of data, and that makes it possible to understand microbial populations in totally new ways.

This kind of work — assembling large-scale metagenomes — is unique, because it gives you a shot at knowing what you didn't know, that you didn't know. Here's Robert Riley again, the data scientist we heard from earlier.

Robert Riley:Yeah, it helps us find the blind spots that are there in, you know, the normal metagenome assembly workflows that are usually available to us.

Menaka: And Robert told me that might be the coolest part about assembling this kind of dataset.

Robert Riley: I'm just surprised by how many things we can find that possibly no one has ever seen before. Microbes, viruses, eukaryotes. And those kind of surprises are just really neat and really gratifying to find. And I think that's one of the best things about MetaHipMer.

Menaka: So – to get a dataset through MetaHipMer, first, it took developing the assembler. And now, it takes a lot of computing power, with plenty of set up, and careful analysis and work from data scientists like Robert Riley. After all of that, it's possible to discover entirely new organisms, and learn more about how they operate in their ecosystems.

But of course — before any of that analysis happens, you've also got to collect the samples. And so that's where we're headed next. Episode 3 of this series takes us to the water. We're off to Lake Mendota in Wisconsin, for an up close look at collecting these microbes. That's in two weeks. See you on the dock.


Menaka: So again, that was Trina McMahon from the University of Wisconsin at Madison, Emiley Eloe-Fadrosh, Kjiersten Fagnan, Robert Riley and Georg Rath from the JGI, and Gary Jung from Berkeley Lab. We also had quick cameos this episode from Ramana Madupu and Todd Anderson at the US Department of Energy.

This episode was written, produced and hosted by me, Menaka Wilhelm. I had production help from Graham Rutherford, Allison Joy, and Massie Ballon.

This episode featured music from JGI data scientist Robert Riley, with drums by John Messier and Joaquin Spengemann. We also had music in the middle of this episode by Cliff Bueno de Mesquita, who's a postdoc at the JGI.

If you liked this episode, help someone else find it! Tell them about it, email them a link, or leave us a review wherever you're listening to the show. And don't forget to subscribe!

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!