

Ginkgo Bioworks

Ginkgo Bioworks presentation delivered at the 42nd Annual J.P. Morgan Healthcare Conference on Wednesday, January 10, 2024 at 9:45 AM

Moderator: Hi, everyone. Thanks for coming. I'm from the Life Science, Tools, and Diagnostics team at J.P. Morgan. Welcome to the 2024 JP Morgan Health Care Conference.

We're about to do about 20 minutes of questions and 20 minutes of Q&A. With that, I'd like to welcome CEO Jason Kelly from Ginkgo Bioworks.

Jason Kelly: Thank you. Sure, why not? Yeah, clap it in. Appreciate that.

I'm Jason Kelly. I'm the co-founder and CEO at Ginkgo Bioworks. Super excited to talk to you today about the progress Ginkgo's made in 2023, especially the work we've been doing in biopharma. I really consider this to be kind of a breakout year for us in that market, so excited to talk about that.

Also, I want to spend a little time on why I see Ginkgo as sort of general purpose infrastructure for biopharma R&D in the second half of the talk. Before I get to that, I will be making forward-looking statements for a public company. When you have a chance, feel free to read our disclaimer here. Let's jump into it.

Ginkgo's mission is to make biology easier to engineer. I really like this photo here. This is a picture from our facility in Boston, so one of the bays. We have about 300,000 square feet of highly automated labs. This particular infrastructure, these robotic carts, we actually build these in-house.

We have about a 1,200-person company, about a 300-person software and automation team, where we're vertically integrated into hardware manufacturing to support automation.

The reason I give this example is this facility is somewhat akin to an Amazon Web Services data center. Think of us, or think of Amazon as investing a large amount of capital into a large data center, and then Amazon Web Services customers accessing it on a service basis.

That's very much what Ginkgo's trying to do in the area of cell engineering, so designing DNA to engineer cells for different applications. We're investing, in this case, hundreds of millions of dollars into a facility like this so that our customers can access it as a service.

I think that's particularly important. I know this is a sort of investor-heavy crowd. It's been a tough couple of years in the biotech capital markets. There's a lot of pressure and a lot of questions around, when do companies need to raise money? Do they have enough funding to get to the next milestone? What's the path to profitability?

I want to highlight that there is a pretty substantial distinction between a product-oriented small mid-cap biotech and a services-oriented one, like Ginkgo.

In particular, we have a much smoother path to profitability than a asset-based biotech, where you're waiting to see how is that clinical trial going to pan out. Eventually, I'm going to get to market and get to recurring revenue far into the future. Gingko's service fee has come in today. We'll talk a little bit about how we structure those deals.

I really like our position going into 2024. We're ending '23 with more than $950 million of cash and cash equivalents in the bank. We are reiterating our full year '23 revenue guidance expectations of $250 to $260 million, $145 to $150 of that's coming from cell engineering, which you're going to hear about today, and a $110 million from biosecurity.

We're expecting to also be within our previously disclosed guidance of 80 to 85 new cell programs. These are partnerships we have with external partners. We're adding 80 to 85 new ones last year to the platform.

If you think ahead and we'll talk more about this with our next earnings call into 2024, a big part of my goal is going to be keep adding new programs to the platform while keeping a lid on our spending and driving more efficiency through that automation, through that sort of data center like environment we have.

I think we're in a really nice spot to just sort of tighten that cash burn at Ginkgo and sit safely within our cash balance here so that we don't have to raise money if we don't want to. Again, I want to highlight that as a distinction between a services and an asset-based company.

That said, '23 was an absolutely breakthrough year for us in the biopharma space. Just on some highlights, we announced a deal with Pfizer in the area of RNA drug discovery. We do bring in

near-term service fees, and I'll show that on the next slide. But our programs also have milestones, kind of similar to what you would see in an asset-based company.

Where down the road, we have the opportunity if there is successful technical and clinical results in the hands of our customer -- in this case Pfizer -- we could get up to $331 million in milestones. Boehringer Ingelheim, we signed a deal with $406 million in downstream milestones. In this case, for small molecule and natural product drug discovery against undruggable targets.

Then with Merck, we signed a deal with $490 of downstream milestones for optimizing enzymes in service of biologics manufacturing. Then what's really cool is some of our deals that I spoke probably two years ago now and talked about a new deal we had just done with Biogen at the time. Well, just a couple weeks ago, we announced the successful completion of that collaboration. This is in the area of AAV manufacturing optimization.

Again, we don't do the manufacturing. We're essentially outsourced research services here. What's great about a successful announcement, and this is another distinction between Ginkgo and sort of an asset-based company...This isn't like, "Oh, we had an asset and [indecipherable] they're taking it forward and it's off. There it goes. There isn't another chance to do it."

This is a signal -- and if there are biopharma companies in the room -- that Ginkgo is good at doing gene therapy, AAV manufacturing optimization, and they should call us if they want to use that service.

When we complete these successful programs, it's actually wind in the sails of us doing more business of that type rather than an end, like you might imagine with more of an asset-based business model.

Finally, we announced that we had a successful achievement of our first milestone with our partnership with Novo Nordisk, again an opportunity in manufacturing optimization that we're very excited about. That was great to see.

The numbers back this up as well. If you look at the active R&D programs at Ginkgo, we have about 100 of them right now. The percent that are biopharma is now 36 percent. That's twice what it was two years ago. Our cell engineering revenues from year over year have grown 50 percent in the biopharma space.

That's a little bit blocked off there, but it's 48 million, so about a third of our total cell engineering

revenue in '23 coming from biopharma. And that's up from basically non-existent in 2020. I do want to highlight that. This is a key point I'll again make about Ginkgo as a platform.

I'm going to spend most of our time today talking about biopharma, but Ginkgo also works in the industrial biotech industry, things like industrial enzymes for laundry detergents and chemical manufacturing. We work in ag biotech.

As you know, tough capital markets years, I would say certain sectors have actually gotten impacted even worse than biopharma. Industrial biotech got really squeezed by capital markets. These are people pursuing things like animal-free meats or new applications of biotech. As a broad platform, Ginkgo could pivot our sales effort, our marketing in the direction of where we still saw biotech research happening, in this case, into biopharma.

That transition happened very nicely. This is also a thing to keep in your back of your head if you're trying to understand what makes for a company that can move with changing markets and in tough conditions. A platform and services business model offers us the opportunity to do that.

You can imagine a very similar thing when it came to, for example, modalities. You could imagine a particular modality having a bad clinical result and some safety data tanking further investment in that area. If you're an asset-based company in that modality, that's a tough day for you. As a platform company, we can just move our efforts into other modality areas. This is a nice slide again for the biopharma, potential customers in the room of Ginkgo.

We have now programs in a wider range of different modalities. I mentioned at the top, you can see our discovery programs. At the bottom, manufacturing related, but in RNA and therapeutics, talk about Pfizer. Gene therapy, just announced a partnership with Arbor. More on the discovery side.

Obviously, Biogen on manufacturing, small molecule, biologics. In microbiome, a company like Synlogic is going into phase two, phase three trials with engineered microbes that we help them develop in the microbiome space. We're getting this 1,200-person company.

There's no way an asset-based company could be in all these different modalities. As a service company, the common thread between all these programs is behind all those products is a design piece of DNA. If you want to pursue your Pfizer, you want to get a better RNA therapeutic, you are designing nucleic acids, and you want to test a lot of different ones to see their performance.

If you're a bow-ringer and you're trying to discover a natural product, you want to go explore microbial genomic collections, synthesize DNA, test it in microbes, express new natural products. You have to go through thousands or hundreds of thousands of genetic designs to do that project.

The common thread across all these modalities is Ginkgo is able to help our customers essentially program their DNA better than they could themselves. That last point is very important. If I'm developing an asset, I can just decide that my technology is good at RNA drug discovery.

You have to try it. You've got to believe me or not. But as a service company, I don't get to decide that. I actually have to get Pfizer to sign a contract and pay me money and say that I'm good at doing RNA drug discovery.

What gets me excited about this chart is we've been filling in all these different modalities with customer validation that the technology we have is relevant and additive to what these companies have in-house. That's very exciting, and that gives us more growth in that industry, which is why I was excited to announce just earlier this week that we've added. This is just the inaugural crop of folks joining us on our BioPharma Advisory Board.

Norbert Bischofberger, a longtime technical leader at Gilead, Jeff Lagos at Novartis, John Maraganore, founder of Alnylam, Paolo at Moderna, Mark, and Kristi from Kite.

We have this group who is able to help us make sure we are building technologies that are relevant to what senior R&D leaders are looking for in the biopharma industry. That is the job of this board, and make sure that Ginkgo keeps investing in things that are breakthroughs for folks developing new drugs.

In 2023, if we scope out beyond the biopharma industry, there was a new tech company added to the trillion-dollar tech company club in market cap, and that was Nvidia. I was really excited to see Nvidia speaking here on Monday and hear from Kimberly Powell.

I'll just say or quote, "Generative AI presents a new class of tools that will get codified into applications and new methods of discovery. In fact, it will go beyond drug discovery, evolve into design, helping create the conditions to no longer be a hit-or-miss industry, and that will be what helps build the world's first trillion-dollar drug company."

Which is a good question. Why aren't there trillion-dollar companies in biopharma? We have five trillion-dollar tech companies. This is the biopharma. We have quite a number above $100 billion, quite a spread.

What is preventing this sort of scale, not achieving this scale in the biopharma side? It's not the market. It's a $2 trillion market for medicines today, and this is data from the NIH here on the left, 95 percent of known human diseases with no treatment.

Enormous opportunity for new application development, giant existing market, great margins. It's not a market problem. It's this problem. This is a well-known problem in the industry. This is the number of new drugs per billion dollars of research spending since the '50s to 2020 going down. We're getting less effective at the development of new applications over time.

This is evidence that the development of new drugs does not benefit from economies of scale. A bigger company pursuing twice as many drug candidates doesn't just magically do it better than a smaller one.

That's not a problem exclusive to therapeutics. I mentioned Ginkgo works across biotech in ag and industrial biotech, too. If you look in the ag industry, this is the cost to bring a new genetic trait to market. It's not exactly analogous to a drug if you put a few traits together to make a new crop.

$ 100 million-plus, and it's dropped 5, 10 percent in the last 20 years, so similarly stalled out, not having these breakthroughs like we've seen in the tech industry. It also takes 10-plus years. GIGO works with the largest ag biotechs today, Syngenta, Corteva, and Bayer. We're familiar with this.

I think this point at the bottom is the heart of why this is, and I want to spend a minute on it. Each product in biotechnology is developed in a one-off process with a goal of improving the odds of success. In other words, we need a scientific breakthrough to develop that new drug.

You want the scientist to try any crazy new variable idea to try to have that breakthrough that makes the drug happen. That variability is the enemy of scale. By the way, industrializing R&D is not an idea that wasn't tried in pharma.

In the '90s, there was a big push around this in combinatorial chemistry and da-da-da, treating it

more like a funnel and making the R&D more standardized to run more through it and produce more drugs. Didn't work.

The reason is you need the scientific breakthroughs. You need the flexibility and diversity of thought of scientists to come up with these new drugs. I think the industry looks at that, and they say, "Well, we need the unique one-off approach to get a drug, and unique, one-off approaches aren't scalable. I guess we won't have trillion-dollar companies.

We won't have economies of scale in drug development. There's a fundamental tension that we just can't get through. Ginkgo believes strongly that's a false choice between scale and flexibility. We do agree that you need flexibility, that the scientific leaders, our customers, and Ginkgo need to have the ability to pursue the different trails they want to pursue to develop new drugs.

We believe that can be done in a way that gets better with scale. The reason philosophically we believe that's possible -- I'm speaking to the biotech side, we're not a chemistry shop here at Ginkgo -- biology has a common origin. Four billion years ago, life evolved, and everything, all the biology you see around you, is derivative from that.

That turns out that something in what you can learn from an RNA project can read on an antibody project. Something you learn in protein design for agriculture can read on protein design in biologics. Having a common tool that's aggregating biological learnings across all programs in this industry would be valuable to every scientist, not just a tool specialized to their one area.

Secondly, the lab work. A huge amount of the research cost is the lab personnel, the reagents, and the equipment we all use. If you walk into an RNA discovery lab, an antibody discovery lab, or even an ag biotech lab in St. Louis or whatever, the equipment's all the same -- same centrifuge, same pipette, same mass specs, same GCs. Half the kits are the same.

What's different is the protocols that are being run, the scientific training of the person doing all of it. The hardware, the low-level infrastructure, is actually common. Why is that? Because it's all working with biology. The fundamentals, we're all working with the same substrate as we're programming these cells to do all these different things in drug development.

That means that if you could solve the problem of having your samples move between the various equipment in a custom way that scales, that's automated, it should work. You should be able to use the same automated infrastructure for all of those different projects across all of those different markets. At the end of the day, the fundamental hardware is all the same across those

different programs.

Those are the two reasons that we're confident. Fundamentally, you could get flexibility and scale if you just invested sufficiently in core platform tools. It is not a small amount of investment. It is billions of dollars in investment needed to get to that type of common infrastructure.

Why has no one done it? Look at the motivations of the various players. Academic community, very early stage, trying to do scientific discovery, don't have the kind of budgets to put billions of dollars into robotics.

Small biotech are rightly very focused on their asset and getting it as far along as they can, as cheaply as they can. They are certainly not interested in spending a bunch of money developing fundamental tools.

Traditional CROs, I think there is an opportunity for them to do something here, but these companies were built out on outsourcing standardized things. Let me take that animal study from you. Let me take that synthetic chemistry from you, sort of predictable research. They weren't really driven out of breakthrough technologies.

Large Pharma has taken a few attempts at this. Vertex acquired Aurora in the early 2000s, an automation company in San Diego. At various times, people build automation, but the major constraint is that therapeutics company just cares about their pipeline. It's either some one to three disease areas or one to two modalities.

They don't have the motivation to build something general across the whole industry, which is what you need to bend down that cost curve on the R&D per project. I think the current players aren't going to do it, which is why I'm happy we've been investing in it. I think you do need a new player who's focusing on this as what they exist to do.

Ginkgo's invested over a billion dollars and a decade building a common technology stack to work on this problem. I want to give you a little bit of insight under the hood of what I think are critical assets in that technology stack to make it work.

The first is flexible automation. I showed you that automation infrastructure at the beginning. I'll talk about it in a second. This has solved that problem of doing general lab work all on the same platform, even if scientists are asking for different work to be done.

Second, data assets. Automation will generate a lot of data. You need to actually keep it and store it in a form where it can be reused. It's not a trivial data science problem.

Then finally, thank goodness for our friends in tech who are really pushing the field forward on large neural nets and so on. We're going to benefit from all of the investment in compute, better chips, lower cost, larger neural nets, algorithm development. That's all going to feed back into making use of that data.

These three pillars are what's going to allow us to have a general purpose platform across all biotech. Let me go through each one.

If you're not an automation nerd, you might not know this. On the left-hand side, this is the general automation spectrum in the average biotech lab. On the left, you have a person with a pipette, completely infinitely flexible. They can do anything they want with that thing. They can come up with a new protocol tomorrow and try it. Super flexible, extremely low throughput, and expensive.

In the middle, you have walk-up automation or task-targeted automation, like a box that you put some plates in, and then it does a very specific thing. When that thing is done, you take the plate and you move it to either another box or back to your lab bench. It's like half flexibility, a little bit of scale.

The third one is what you might buy from an automation vendor like Hi-Res, an integrated work cell. There's an arm, and it reaches these different equipment, and it moves the plates through the equipment, so a high-throughput screening workflow in a large biopharma would run on a work cell.

Boom, boom, boom, we're going to move the plate. It just does the same thing over and over again, but it does it a lot. Very inflexible. It's hard to make it do something new, but it does it at a high throughput.

What Ginkgo's been investing in is essentially a reconfigurable automation system, where we have each piece of equipment in a cart. It can get plugged into a magnetic track, and samples can be sent to...and you can see a little sample running on the track there, can be sent to any one of the pieces of equipment. You can swap in and out equipment, and also change protocols in software.

Again, these are built in-house, a big operation. You only need this if you're trying to do general purpose lab work. If you have only one thing you need to automate, just build a work cell.

This type of stuff is what I'm saying. If you're not trying to solve the large general problem of flexibility and scale across the whole industry, you don't make these investments, which is why we're the ones who've been making it.

Next, on the data side, two types of data I want to talk about. One is raw genomic data. This is the type of data that is present in GenBank, for example. What I want to show you here is non-human genomes. Basically, how many deduplicated genes are present in, in this case, Uniprot. It's 246 million. That's in the purple bar. Over 2 billion deduplicated genes in Ginkgo's internal database.

Substantially larger, and you can see the growth rate of our database has been quite a lot faster than Uniprot's over the last five years. That's half the battle. This is available genetic variability from nature that you could draw on to learn how to design promoters better, how to make a better enzyme, whatever it might be.

Importantly, the complement to that is then characterizing all of those different genes. We've been doing that at high throughput. Just this chart up here on the right. Each one of those lines, you can't see it clearly, but is a EC class, an enzyme class. We've tested, and this is across many, many years and lots of experiments, more than 5 million characterized data points like this that we can then use to feed the models I'm going to talk about in a second.

This summer, we announced a partnership with Google Cloud. Ginkgo's going to spend $250 million on compute training. Google is doing about $50 million in milestones back to Ginkgo.

Really nice comment from Thomas Kurian here, the CEO at Google Cloud. The strategic partnership with Ginkgo is first of its kind, underscoring our confidence that Ginkgo will play a critical and pioneering role in the life science industry. This is a type of thing that we want to invest in. We want to be able to invest in what we think are common, valuable, reusable assets across the industry.

We've started the first round of tests that we're doing on these models is, basically, to take some of that data I showed you on the last slide, where we have proprietary assets that are bigger than the public sets. Just push it into models that are very analogous to what are currently out there in the state of the art and see, how much does new data help? The answer is it helps.

I think there is an opportunity. We've seen significant outperformance, especially when it comes to classes of biology that aren't well represented in the public data sets.

I think there'll be an opportunity both for more data generation and for further algorithm innovation when it comes to improving the quality of these neural nets in the biology space.

We've got that data, so we want to push the bar and help the field understand, are we anywhere near diminishing returns on the improvements to these models from large-scale data? I'm optimistic we're not. There's a lot more room to go.

Those are the three pillars. There's the automation, the data, and this AI pillar. I want to give a few specific examples of using these in customer partnerships.

This first one's with large biopharma in the area of enzyme design. In this case, there's a catalytic activity that this customer wants to have. It was not known in the literature of enzyme to do it. What we first did was we started with a prediction from our enzyme engineers of some possible sequences from metagenomic search that could do it. Those are seed sequences.

Computationally, we say, based on those, here's 183,000 sequences that are similar from that set of 2 billion that we had. We don't want to make all those, so we computationally predict 1,000 that we think are particularly good, and we synthesize them.

We're Twist's biggest customer. We put them in, we automate, clone them in, test them, and see their performance. We find a few that work.

Those few that work become the basis for protein engineering. I want to spend a minute on the next couple boxes here because you're going to hear this again and again from us and other people in the future as they are applying AI in bioengineering.

What you're looking at there is us running our protein model to try to optimize these sequences on the axis of specificity on the y-axis, and activity on the x-axis.

We specifically told the model, "Make some variants that are more active, and make some variants that are more specific."

It made some choices, we built them, and we tested them. We found, "Here's one that's five times

better activity and 14 times better specificity." That's the first dot. That's the gen-one dot there on the far right chart.

Now here's the important part. You take that data, the middle plot, you go back to your ML AI model, and you do what's basically called reinforcement learning on the tech side, or closed loop -- has become sort of term of art in biotech, biopharma -- you give that data back to the model, and you tell the model to get better based on it.

You say, "Now that you know that, why don't you recommend a new set of designs for me to test?" That's gen two. You can see in gen two, now we're at 10 times better and 25 times better on those two axes.

You take that data, and you give it back to the model, and you tell the model, "Improve yourself based on this data." That's the loop. You tell it to recommend again, and that's gen three. In gen three, we get 110-fold improvement in activity and 21-fold improvement in selectivity.

We also now are sitting on a model that is real good at designing this class of enzymes. If a new customer needed this class of enzyme, we wouldn't need to go all the way back to gen zero, that's point number one, but this general loop of "start with a generic protein engineering model," understand what you want it to do.

Remember, this is my point, you have the flexibility as a scientist. This is a very specific thing they're asking. Generate some data in the direction of your particular interest, tune up the general asset, and it will get better at your particular thing. That is flexibility with scale. Not flexibility against scale.

This is the point I really want to make. We have done similar in a totally different area. This is data before our Pfizer. It's not our work for Pfizer, but we have this partnership with Pfizer and RNA.

One of the things we've been developing in the RNA space is circular RNA. This is a way to make RNA hang around longer in the cell if you want to express things. Vaccines are great because you don't need a lot of protein to trigger the immune system, but if you wanted to express more protein biologics using RNA, you've got to have it stick around. So circular RNA lasts longer.

The way you make it, you have an intron in there that's basically doing some fancy stuff to circularize the RNA. Different designs of the introns, different levels of circularization. We ran our

process, and we've been able to find introns with about 90 percent circularization efficiency. That's better than state-of-the-art, and we think we can keep driving that loop to make it better.

The other thing people are asking for is expression in different cell types. You see that on the right. Again, think of the loop. You get some data on how things perform in these cell types.

That goes back to feed into a model that understands, starts to learn what it is about the design. In this case, it's the design of the I-res's. That affects their expressibility on a differential basis in different cell types.

I'm going to use the next one more example of this. This is in the area of CAR T. This is some work we've been doing internally. We generated, in this case, a library of 10,000 intracellular domains for a CAR construct. That's over on the left.

A year ago, I showed us putting that into T cells and showed some nice data on in vitro improvements in the lab for exhaustion, so making these cells last longer.

We're now taking that work into animals in partnership with [indecipherable] , and hopeful to see that data soon, but that was the first line of work. Now what's great is we could take that same library, put it into NK cells, and screen for NK cells that had a strong immune activation.

You can see there's, in that lower right corner, three different metrics that we're using. I'm not going to get into the details to measure that immune activation. We 8 out of the 10,000 that had high activity.

What's really exciting is we found differences. You can look at the heat map of all the different designs and how they affected activity. That's the exact kind of data. It's not just the data about the eight winners. It's the data about the 9,992 losers that you also want to feed into the model so it learns what designs affect good and bad on this particular access you're trying to achieve.

That's, again, not a custom model. It's a general model that's being tuned up with custom data. It is flexibility on top of scale. Does that make sense? This general closed loop is something you're going to see again and again as we have platforms that apply automation, large-scale data, and general-purpose AI.

One last area I want to mention, Ginkgo does not have, I don't have programs here yet, but I want to highlight that I think you should also, if you're following AI in the biotech space, expect to see

this, so there's a pretty cool announcement.

Recursion and TEMP has talked about, in licensing, a bunch of human genomic data around basically patient genomes associated with different diseases. That's the top left box.

You're going to have people use these large models, parse more and more human genome data, have predictions of potentially new targets. Target discovery. OK, that's this category. New target prediction. Now the question is, is it real? Is it a real target?

This would traditionally enter the world of systems biology, where we'd say, "OK, great. Let's go look at it," and start trying to understand it as systems biologists. What the hell does this gene do that keeps showing up in these patient samples?

I think the new way you're going to see people do that is they're going to take a human cell. In this case, we can do greater than 20,000 perturbations this way in a high throughput, but they're going to make a bunch of changes to the genome of that cell with high throughput CRISPR. They're going to have this big array of variants around that target they're trying to explore.

Then they're going to test all of those with high throughput omics all over the place. You're going to have imaging, but you're also going to have transcriptomics, metabolomics, and so on.

You're going to get this giant data dump across 20,000 different variants across that particular target you're interested in, and you're going to feed all that data into a model, and it's going to learn what the heck is going on.

That is going to go back to the scientist who is trying to identify that target, and they're going to work with it and say, "I could use a little more information here," and they're going to go around that loop again.

Ginkgo is particularly good at those two green boxes on the right. If this becomes more of the status quo for how people do drug discovery, I think we're going to end up target discovery. I think you will see us playing more and more in that space. I'm pretty excited about this. I think it's a real opportunity.

I want to end on this. One of the other reasons the tech industry has been so successful is that these trillion-dollar companies build on top of each other. Nvidia, the largest chip company by market cap today, does not make its own chips. It manufactures the chips at Taiwan

Semiconductor. Think about that level of dependency in this industry.

Netflix, Salesforce, Twitch, these big SaaS companies that exploded in the 2000s and 2010s, built on top of Amazon Web Services, Google Cloud, the cloud providers that were going to do all that compute so that they didn't have to build that infrastructure.

All your apps that you get in the App Store on either Android or iPhone, all that distribution's handled by Apple and Android for those developers. Hugely enabling. Let's them grow away faster.

I'll give the specific example of OpenAI. I think people are familiar with OpenAI's relationship with Microsoft, which was, "I'm going to take all the data from the Internet, common crawl, 10 terabytes, feed it through a ton of compute, and build this foundation model," which they did.

That was not ChatGPT. ChatGPT needed additional work, so they called a company called ScaleAI. They told ScaleAI, "I need a ton of people to put queries into my language model and train it. Say if it's a good answer, a bad answer, why it's a good answer, why it's a bad answer, and do this human reinforcement line.

That same closed loop I mentioned earlier that we're doing in the lab. They paid scale to generate that data from humans talking to the thing. That made ChatGPT. That's what made it so conversational and so good.

It was necessary. They built that on top of scale. Scale is basically like...Tesla works with them on image analysis to circle. "Hey, that's a cone. That's a dog." That's a stop sign to help train their self-driving car algorithms. It's a data broker. It's growing into quite a big company.

Ginkgo could very much play that role in the biotech space. You're building an AI algorithm. You need more data. Call us. We could provide that data scale. If you're a large farmer, you want to put your toe in the water, and you're trying to build out data set you don't have to get that closed loop, why build it?

Why build it? It would be better to leverage infrastructure that's already in place, let us grow, and invest at bigger scale to bring down the cost of that data generation.

I think this type of betting on each other and leveraging each other's investments across the industry is missing today. I think there's a cultural change needed to do it.

Ginkgo is very happy to lead the way. I think if we lean into this as a sector, biotech should be bigger than tech. It really should. Ginkgo is hoping to lead the way there. Very happy to do it with you all. There's my email up there if you want to talk more about it. I'm happy to take questions. Thank you very much.

[applause]

Moderator: Great. Awesome. Just as a reminder here, feel free to submit questions...

Jason: Yeah, and I'm going invite Anna Marie up also.

Moderator: on our website here, as with any presentation. It seems like you flagged some really good updates this morning, including an in-line quarter, as well as great partnership momentum here.

We've seen some of these chunkier partnerships, for example, a multi-target RNA drug discovery collaboration with Pfizer, which could earn potentially up to $331 million in fees and milestones. It's hard for investors to understand these partnerships from a financial level.

How should investors think about the chunkier partnerships like this heading into this year, next year? How do you think about your partnerships progressing through your pipeline in general?

Jason: I think one thing that, again, people don't fully appreciate about a deal like that Pfizer deal. You often see these kind of deals announced between a smaller, mid-sized biotech, where they've got some assets in their own pipeline.

They're partnering with a large pharma on a two or three additional asset deal with $400 or $500 million of milestones or more, and dah, dah, dah, dah. That small or mid-size company would do one deal like that a year. It would be negotiated directly by the CEO. It's a big decision. It's about out-licensing in a certain area, a lot of exclusive tie-up, all these things.

These deals we're getting at Gingko, it's a decent amount of milestones. Not quite as much as if you're doing the whole asset, but it's not bad. I'm not doing the deals. They're being done by our enterprise sales organization.

One of the things I'd like to point out is these shouldn't be like chunky, one-off deals. In my view, if

we start to move to a model where companies are leveraging large-scale, multibillion dollar platform investments of other companies and generating data, these sorts of deals should just be coming off the pipe. This is just how we should be doing business as an industry.

There are thousands of drug development projects happening, and I'm announcing three deals. There is a lot of room here.

Because I'm not, this is my point again about being an asset company, it's not like I've got one narrow asset that the only thing I could partner is geography or something. I've got automation, large scale data pile, and AI models.

What I want to do is make this into more of a straightforward thing for people to do, but while still having some, and it doesn't need to be huge, but some amount of participation in the success so that we're aligned with our customers.It makes the economics of the whole industry work correctly.

Moderator: That's great. You recently announced, just speaking of making the economics of success work, the completion of your program with Biogen. You noted in 2021 this could be a $120 million partnership. How should we think about the financial impact of a partnership like that completing?

How do you think that Ginkgo, as it closes out some of these bigger programs, is able to replace those large programs with additional good ones? We've noted from other companies at this conference that there's been some biotech RFP pressure. Is it focusing on key partners, or how do you think about it?

Jason: Yeah, I'll answer the second question and then the first question. Well, I'll go in order, actually. When it comes to milestones, we'll often have the bigger numbers are going to be tied into...and I'll just speak generally because I think we haven't disclosed anything specifically about Biogen...often you'll have milestones tied into commercial progress and things like that that take longer.

I think that's just going to be the reality. I think that's healthy because this is an industry that there is still fundamental risk in developing drugs. A lot of them are going to pan out to be zero. You could create a ton of value, put a lot of the technology into a drug, but if the thing ends up being worth nothing, the customer feels pretty damn bad about paying you a lot of money for that.

I think it is quite aligning to have big checks when there's big commercial progress. I really like that model. It just takes time. It's just a thing you got to wait for. The reality is we'll keep stacking those up. Ginkgo will also focus on, again, compressing cash burn by increasing the service fee aspect of our deal so that we make it to all those royalties, so just from a strategic standpoint. Does that make sense?

Moderator: Yeah.

Jason: The second question was about the refilling the pipe. Yeah, and also should we double down on our current big customers?

The other thing that I'm excited about with Ginkgo today is we have basically not penetrated the biopharma industry in any serious way.

We were talking. We go out still today. We were at JP Morgan the last three days having nonstop meetings. We will meet people, and they're like, "I thought Ginkgo did Ginkgo the logo or something." They're still just like, we're early in the transition. You saw my bubble revenue going up from basically nothing in 2020.

This is the most recent market for us to apply this general platform. We started in industrial and agriculture, in part because the bar was lower for their capabilities in R&D. You guys are better. And so I'm selling a platform.

I still remember I told you I got to convince a company to use me. I can't just say I'm in gene therapy or RNA. I got to convince someone that I'm better than their infrastructure in-house. That took longer for me to get to that scale. Took me about eight years of really having funding and growing to get there for biopharma.

Now that we're here, we're here. We're going to be here for the next few years. It's going to be the focus. Hopefully, for me anyway, no, there's no reason to double down on the few we have today, because there could be someone out there who's a much better customer that we just haven't even met.

Moderator: Awesome. We're out of time. Really appreciate you guys coming.

Jason: Thanks a lot. Appreciate everyone's time.

[applause]