www.a16z.com

When Software Eats Bio

by Vijay Pande

Editor's Note: Vijay Pande ("pahn-day") is Andreessen Horowitz' newest General Partner, dedicated to our new bio fund.

He first joined a16z as our first professor in residence from Stanford University, where he was a professor of chemistry, computer science, and structural biology; directed the Program in Biophysics; and ran an award-winning distributed computing lab whose work contributed to our understanding of Alzheimer's, Huntington's, and various kinds of cancer. Pande also co-founded Globavir BioSciences; was an early developer at a video game company that was sold to Sony; and is an advisor to numerous IT and bio startups.

a16z: This seems so obvious, but why a bio fund?

Vijay: There's a couple different ways I think about this. One is that we all care about human health — whether it's for ourselves, our parents, our children — and it's a big deal on a very deep, fundamental level, in terms of thinking about the meaning of life. At a much more mundane business level, there's clearly a huge market opportunity here. Just think of the marketing budget that Google can go after (with ads and such) — $200 billion. But compared to that, the U.S. healthcare budget is $2 trillion! Even tiny little sub-budgets of that are huge markets for startups to go after.

Bio is therefore an area where there's a real chance to change the world … but also a chance for really great financial returns as well. The firm has been excited about this space for a while, and we've made investments before even raising a bio fund. But we want to do something really big, and expect this to grow bigger in time, so establishing a separate fund is also about our thinking years down the road.

a16z: So why now? Areas like healthcare (and education, among a few others) have been impenetrable to disruption, despite periodic claims to the contrary. Such hyperinflated industries have always been ripe for tech, yet they've never really been remade with tech. I feel like we keep saying 'this time is different'.

Vijay: There IS a specific confluence of trends right now. On the computational side things are fundamentally different. Even though Moore's law made Silicon Valley, we still can't conceive of how exponentially the cost of computation is going down.

One of my big projects at Stanford, Folding@Home, got a Guinness World Record for the most powerful supercomputer in the world; it was the first to reach 1 petaflop. But now, that amount of compute power costs $400 a day on Amazon. That sort of "exponential decay" results in declining cost, making what used to be extraordinary and world-record making both average and cheap today.

That's what's happening in compute. But there's also a Moore's law for storage that's been exponentially decreasing as well. When you combine this "free" compute and storage and data, you get big data — which machine learning depends on — which in turn leads to deep learning.

a16z: So how do we connect these dots to bio?

Vijay: Bio has its own Moore's law. Because the cost of sensors are going to zero, the cost of things like genomic sequencing are going to zero. Actually, they're going to zero faster than with Moore's law.

The Human Genome Project was set up in 1991 and finished in 2001, for something like $3 billion. Now, it would cost $300. That's a clear exponential decay in cost. It creates an interesting situation where so much is available to us right now. What's left is the *software* to put it all together.

a16z: How can you make the claim that software connects the dots? Because when I think of bio, I think of tissue and flesh; I don't think of computation and algorithms. How do those two actually come together?

Vijay: Let's take machine learning. You can now do so much with image recognition there. And a big part of medicine involves images. Sure, when you go to your doctor, a bit of listening happens, but most of it is really about analyzing your x-rays (radiology), examining your skin (dermatology), or looking at your eyes (ophthalmology).

Of course, these doctors aren't just using their eyes; they're applying and honing decades of medical training to do the pattern recognition, which in many cases is very subtle and requires significant expertise. There's going to be many examples like this where computation can do something beyond what a human being can. It's not limited to just vision. Think of all the inputs that humans take in with their senses; each of those are amenable to machine learning and deep learning: Listening with a stethoscope. Smelling something. And so on.

In many cases, algorithms can do better than humans. Just as computer vision has had a huge impact in non-medical areas, it's now getting to the point where it can set a new gold standard. If the gold standard in radiology is to predict what radiologists would do, computers can go beyond that. In radiation oncology for example the gold standard would be to *predict* the biopsy results … without having to actually put the patient through one.

a16z: What you're describing is essentially disintermediating doctors, isn't it? What are the implications of that, more broadly?

Vijay: I don't think the goal here is to take people entirely out of the equation. It's to help the experts.

Imagine a computer algorithm that does the equivalent of what spellchecking does for writers. Similarly, instead of radiologists having to look at thousands of images, the computer vision algorithm flags only the important ones. Just as with a spellchecker. And maybe you say, wait, that's not a typo it's actually someone's name. But the ultimate judgment is for the human to make.

What I'm describing doesn't replace all radiologists and other medical specialists; it just dramatically speeds up their work and allows them to concentrate on higher-order, more complex, more important things.

a16z: It's not just about being cheaper and faster, but *better*.

Vijay: Yes, and what I've just described is actually one of the three big areas we are focusing on with this bio fund — "computational biomedicine".

Because for anything that's machine learning-based — like image recognition and computer vision as with these examples — the machine learning gets better as the cost of compute and cost of storage goes to zero. But what machine learning *really* craves is data. And the reason machine learning and medicine is a marriage made in heaven is that medicine has a ton of data. All of which can now be stored, brought to algorithms, and related to later outcomes.

We can even learn new things as a result. It's amazing: We recently discovered a new piece of human anatomy due to more precise microscopes. I was shocked; I thought anatomy is one of those areas we actually had locked down!

So I think taking this data-driven, computational approach to medicine will open up lots and lots of opportunities not just to improve the accuracy and quality of medicine, but to build really big companies as a result.

a16z: This kind of machine learning and big data requires compute and storage. Does this mean we are finally at an AWS-like infrastructure moment for bio startups, much like what happened for web-based startups?

Vijay: It's an area that we've called "cloud biology" or "cloud bio". Even the name is meant to evoke cloud computing, and all the new businesses that the cloud enables.

But what's happening here is that real-life, real-world experiments can be done in a cloud-like fashion.

a16z: In a "cloud like fashion" — what does that even mean?

Vijay: Why is the cloud so important to startups? A startup in the software space would have had to spend $10-$20 million to build up a server farm, just to be able to do anything at scale in 2000. And scale is incredibly important because you can't really prove your product by only running on one or two machines. Cloud computing meant that you could later give a startup $2-$3 million, and before they came back for their next series A investment, they would have a product out there, running, with customers.

You can de-risk early. And that's a fundamental difference between bio and traditional biotech, where you often had to put in $100 million and then wait five years before there was any sort of signal for whether it was working or not. We can now give computer science grad students or MDs $2-3 million, and they can use cloud bio resources instead of having to build out the lab (which is the analog to building out a server farm).

a16z: Is all this just about achieving product-market fit faster? Or can we do more as a result of cloud computing applied to bio?

Vijay: While cloud computing leads to lower CapEx and often lower operating costs, what's nice about AWS or other cloud compute services is that if you want to spin up 10,000 cores for five minutes and spin it back down, you can do that. And so with these new cloud bio resources you can spin up experiments, whether it's *in vitro* experiments driven by robots, or animal experiments.

So no, it's not just about cost efficiencies and a more efficient pathway to market fit. You can now also do things that you couldn't do before. The elasticity the cloud provides to bio is key.

a16z: "Spin up experiments" — I love that turn of phrase. Besides being able to do that, how does cloud bio touch on the issue of reproducibility and accuracy in scientific research? I feel like we're suddenly seeing a lot more about this lately, even though the problem has been around for ages; how does this issue fit in this context?

Vijay: I think we're seeing a transformation right now, sort of like an Industrial Revolution for biology. If you look at how the state-of-the-art in biology has always been done, it reminds me of something almost pre-Industrial Revolution. It's rows and rows of people working with their hands at benches, and in an apprenticeship-like way under a master biologist (often a professor).

It's very difficult to achieve reproducibility of scientific results — which is important for advancing the field and deciding what research paths to pursue — in this context. I mean, even how you pipette can have a huge impact when you're putting reagents in test tubes! Just two weeks ago, I heard a story where what the grad student ate for lunch changed the results. (Tuna fish put amines on his breath and therefore into the reagents; it was something that was very difficult to track.) There's other stories like that out there too, like laundry dyes in a lab coat, and so on.

a16z: Ok, so all kinds of spurious variables can come in because of this mechanical fallibility. But I still don't quite get how the computational aspect helps address the problem.

Vijay: So what happens in cloud biology is not purely the computational part, but the fact that computation is driving robots who can do the experiment.

At one of the more exciting companies I've seen in this space, when you want to do an experiment, you literally write a computer program. When we say that biology is becoming a software problem, in this case it's quite literally the case. If you or anyone else wants to reproduce the experiment, you just have to get a copy of the program and rerun it.

a16z: So doing an experiment means running a computer program. Isn't this what computer simulations and modeling already do for us?

Vijay: These are real-life experiments. Simulations always have to make a trade-off for compute cost versus accuracy; that's the main issue. This is the real thing.

a16z: Ok, you've now described two areas we're focusing on so far with the bio fund — "cloud bio" just now and "computational biomedicine" earlier. Anything else?

Vijay: The other huge area of interest for us is "digital therapeutics". It's a term pioneered by Omada [which we're investors in] and others.

The way I like to think of it is this: If the first phase of medicine was about small-molecule drugs delivered intravenously, and the second phase (then led by biotech companies like Genentech) was about protein biologics, then the third phase is about *digital* therapeutics.

It seems like the holy grail of medicine is to take a pill, wait a bit, and then get better — just like magic! But there are real limits to this, especially when it comes to depression, PTSD, smoking cessation, type II diabetes, insomnia, and other behavior-mediated conditions.

I'm confident that 10-20 years from now when we look back on this phase of medicine, it's going to seem backwards and even barbaric that our solution to everything was just giving out pills.

a16z: Are you just describing preventing disease in the first place (vs. treating it after the fact)? Or are you talking about changing habits? What does a digital therapeutic actually do?

Vijay: It very much changes habits. Digital therapeutics treat what are really *behavioral* problems with a *behavioral* solution.

To give you an example of what a digital therapeutic would actually do: Let's say I'm borderline for type II diabetes. I could pay someone $100,000-$200,000 a year to follow me around 24/7, like a personal trainer, making me do pushups to build muscle mass and knocking doughnuts out of my hands every time I reach for one. And sure, that would work. It's just really expensive for most of us. Behavioral therapies essentially do the equivalent type of motivation and coordination — and still have a human-touch element through coaches, messaging, social networks — but do so in a way that can scale such that costs are dramatically lower.

Because there *are* existing approaches that have already shown quite good efficacy in this space — they're just expensive and don't scale. A great example is Stanford's sleep clinic or its pediatric obesity clinic, both of which do amazing things but cost a lot and can only take in a small number of (often privileged) people a year. Yet there are millions of people with type II diabetes …it's an epidemic.

Digital therapeutics allow such successful approaches to become cheaper and to scale. And they have no toxic side effects, which is very appealing from a drug point of view; what we don't like about investing in traditional biotech is the risks due to side effects, additional regulatory issues, and so on.

a16z: You've brought up a few times now how what we're doing with the bio fund is so different from traditional 'biotech'. Why? How?

Vijay: The bio fund is really about funding *software companies in the bio space*. Whereas traditional biotech has very little software in it, at its core.

I began by talking about Moore's law. There is an analogous law on the drug design side, Eroom's Law, with "Eroom" being "Moore" spelled backwards. Where Moore's law is about the exponential decrease in cost, Eroom's Law is about exponential *increase* in cost. And over the last four decades, drugs have been exponentially increasing in cost.

In terms of our investment thesis, when we said we're not going to do biotech, we basically said we're not going to do anything associated with Eroom's. And we're still saying that.

a16z: What makes something Moore's (vs. Eroom's) Law? How can you tell?

Vijay: Anything that is *driven by the declining cost of computation*.

Earlier, I mentioned how cloud bio is one of the big differences between traditional biotech and what we think of as bio. So something that's heavily computer-driven and software-driven will go on the Moore's law curve.

a16z: Is a natural consequence of this difference between traditional biotech and bio that the ideal entrepreneur for us is *not* a medical student?

Vijay: No, I don't think that's true.

One of the most exciting things about the opportunity right now is that med students 20 years ago were very different than they are now. Today, they're very computer-savvy; some have been programming since they were teenagers. (At Stanford, roughly 80% of students take a programming class.)

Even if they don't code, these students with medical degrees, these students of biology and chemistry, can talk in a very deep way about computer science. They may not be the CTO, but they have a tech-savvy mindset.

a16z: Flipping the previous question around for a moment, how does it apply to you? You weren't trained as an MD, right? Does that mean you view bio companies differently than a doctor would?

Vijay: Most current doctors haven't been immersed in machine learning let alone been researchers. The perspective I'm coming from works because this new wave of bio companies is far more software-like than bio-like — though there is of course bio at the core, too.

When I'm talking to entrepreneurs I like going deep with them not just on protein biology but on machine learning, distributed systems, infrastructure — or even just general issues with healthcare and medicine. These are all things that are very familiar to me and that I have done either as a startup founder or in my 15 years at Stanford and Stanford Medical School.

a16z: Going back to the beginning here — my point about healthcare being resistant to disruption — why can't the incumbents do some of this if conditions are so different now? They certainly have the data. They know the space inside out. Shouldn't they have a home-court advantage? This is not a case, as with Google back in the day, where a little startup is coming into a space where no entrenched alternatives existed before. This is a deeply entrenched industry with tentacles everywhere.

Vijay: And yet there are plenty of examples where incumbents just didn't have it in their cultural corporate DNA to do that something. You could argue that IBM had everything necessary to build a social network. Even Google, once a startup itself, built a social network that didn't compete nearly as well with Facebook.

It's about different corporate cultures, different styles, and completely different operations. And bio x computer science is actually something very different. It would be very expensive and difficult for hospitals to do since they don't have the infrastructure startups have; it would be like reinventing the wheel. Not to mention the cultural clash in absorbing the results even if they did get that far.

The existing pharma companies and other incumbents are very good at what they do. But from a tech point of view, healthcare companies and institutions are living in the 1980s — it's "fax machine medicine". Just as Google transformed multiple industries, or Uber and Lyft are changing the taxi and car industry, there are similar opportunities here due to tech at the core.

a16z: Clearly the culture and the operations matter. Speaking of, you mentioned the reasons why we're doing a separate bio fund earlier. But how will that work, logistically?

Vijay: If you're looking from the outside, from the entrepreneur's point of view, I don't think you'd even be able to tell that there is a separate fund in terms of how the firm operates. The separate fund simply emphasizes our dedication and excitement about the space, drawn from LPs who are committed to the vision as well.

But in terms of how pitches and everything we do would go, it's all the same. We've got the full team involved here. Besides myself, there are a number of other people vetting these deals. And then all the other general partners are fully involved too — whether they're weighing in on unit economics or the marketplace aspects or machine learning or cloud infrastructure or whatnot. The other general partners have a huge set of domain expertise and other experience to contribute, just as they do with all our other companies.

And it's not all technical expertise, either. It's taking advantage of their — and our operating teams' — core competencies in hiring, sales, marketing, SaaS (which is very different from consumer applications), and so on. It's about building enterprise sales, pricing strategies, etc.

That's the most exciting thing here, business-wise. Because the reality is that these bio startups look exactly like software companies, especially after they achieve product-market fit and gain their first customers. When I said these companies are more like software companies, I meant it — not just in their tech core but in how to build and scale them.