



## **#273: Life by numbers: Systems biology and its approach to researching disease**

VOICEOVER

Welcome to Up Close, the research talk show from the University of Melbourne, Australia.

DYANI LEWIS

I'm Dyani Lewis. Thanks for joining us. Science is very often about chipping away at a problem one component at a time. By piecing together fragments of knowledge on how each component of a system operates in isolation, an understanding of how the entire system works can emerge. But systems biology - an emerging field of research in the life sciences - takes a different approach. It draws together and integrates large amounts of disparate data to take an encompassing approach to the study of life and its myriad processes. Instead of looking at one protein or unpicking a single biochemical pathway, systems biology embraces the complexity of an entire system to find patterns and networks in the data collected and, from these patterns, seeks to explain biological behaviour and function.

Today I'm joined in the Up Close studio by a systems biologist to talk about his field and how he goes about researching complex biological processes such as disease. Dr Michael Inouye is a senior research fellow and head of the Medical Systems Biology Lab within the Departments of Pathology, and Microbiology and Immunology at the University of Melbourne. Mike is also an honorary senior research fellow at the Institute of Molecular Medicine, Finland, at the University of Helsinki. Welcome to Up Close, Mike.

MICHAEL INOUYE

Thanks, Dyani.

DYANI LEWIS

Mike, you are a systems biologist as I said in the introduction. How does that differ from a garden variety biologist?

MICHAEL INOUYE

So following on from how you were speaking about biologists tend to study individual

components, really what systems biology is trying to do is trying to integrate across these different pathways and trying to understand broader behaviours using really as much data as we can. And it really is a data driven science. So really one of the main differences between us and what has come before is that systems biologists tend to be very well versed in mathematics, in statistics, in computational sciences, because the data that we're dealing with is tens of thousands or millions of variables on individual samples or cells or populations of cells. And really to be able to handle that data, you need to have a very strong basis in quantitative sciences in order to extract out the little nuggets of useful biological information in all of that data.

DYANI LEWIS

So if I were to visit you in your lab, what would I be looking at?

MICHAEL INOUE

So our lab is a dry lab. Effectively everyone is very much focused on their computers and programming in various computer languages and doing statistical analysis of data. Obviously we work on ways to also visualise a lot of this data in order to communicate to our own brains what we're seeing.

DYANI LEWIS

The reductionist approach of taking component by component has worked very well for science in past years and decades. Why do we need this systems biology approach?

MICHAEL INOUE

There are two main answers to that in my view, one of which is a technological explanation which is that with genomics and transcriptomics and proteomics and other sort of -omics technologies, you're able to measure tens of thousands, millions of variables for individual samples. Why not? If you could do that and actually be able to quantify DNA, RNA proteins, other sort of biomolecules, you should be able to understand more about the system. And really systems biology is very much needed in order to understand the variation within those components and between them. The second explanation really is that when it comes to drugs and trying to increase the efficacy of the compounds that are going into clinical trials, current biology, current molecular biology and the disease models that we have are failing quite spectacularly where we have greater than 90 per cent of the molecules that go into clinical trials are failing for efficacy and safety issues especially. That means that we have to really improve our pre-clinical models of disease. And systems biology is one very promising way in which to do that.

DYANI LEWIS

These drugs are really getting to the trial process without properly understanding all of the interactions at play. Is that what's going on?

MICHAEL INOUE

I think that's a fair assessment. Really we haven't had the ability to really understand all of the variables in play until very, very recently.

DYANI LEWIS

Now you mentioned different -omics technologies. Can we just go through some of those? So you've got genomics which is looking at the genes in a body. What other -omics are there?

MICHAEL INOUYE

So genomics specifically is really looking at DNA variation, so if you have single point mutations in the genome, if you have copy number variation which is duplications or deletions of genes. Really what's special about genomics is that DNA is really one of the only, sort of, molecular classes that we can be sure has a causal effect on traits and diseases that come about in an organism, because at least the genome of the organism is pretty much set at conception. And that genome obviously, though, there's other ways in which it expresses itself. That starts getting into transcriptomics, as I mentioned, where you're starting to look at the RNA that's being transcribed off the genome.

DYANI LEWIS

So it's basically looking at which genes are switched on and off in given tissues at given times.

MICHAEL INOUYE

Switched on and off and quantitatively if we can possibly, yeah, and getting an idea of the time axis with the organism as well.

DYANI LEWIS

Then you've also got presumably what those genes are producing, which is where metabolomics comes into it?

MICHAEL INOUYE

So really what those genes are producing is - in terms of RNA and then into proteomics and trying to get an understanding of protein composition of the cells that you're working with. Really metabolomics is another really exciting area that's coming into play more and more. It's an area that I've worked in a bit recently that's trying to get a better idea of how the genome, the transcriptome, the proteome are interacting with the extracellular metabolites.

DYANI LEWIS

How much of this data is publicly available? Many people would be familiar with the Human Genome Project. Are other large datasets like this available to you?

MICHAEL INOUYE

Yes, there's a very strong vein of data sharing in the systems biology community, which largely has come out of the Human Genome Project. So there are large datasets available for really anyone with an interest in this area to start playing with. It's really a boon for researchers such as myself, because it allows us to really try and understand other systems and look and see things in our own data and try and find those similar relationships in other public data. So it's a real advantage for the

community.

DYANI LEWIS

Mike, one of the newer fields of big data in biology is collecting data about a person's microbiome or all of the genes in all of the organisms that live on or in us. Is that something that can also be integrated into this systems biology approach as another layer of complexity?

MICHAEL INOUYE

It certainly is, Dyani. The thing about microbiomics is that we've come to appreciate that really we share our bodies with millions and millions, trillions of microbes as well as with our own human cells. They are very intimately involved in a lot of the biological processes that we would consider make us human, so getting a better handle on what they're involved in, you know, bacterial communities that live in our skin, in our noses, on our feet, et cetera, that - understanding, so how they affect disease and how they affect how our body functions is certainly an area that many people are interested in mixing in with these other -omics.

DYANI LEWIS

I'm Dyani Lewis and you're listening to Up Close. I'm speaking with systems biologist, Mike Inouye, about his research into the molecular networks within us. Mike, let's have a look at some of the specific studies that you've been working on. One has been looking at cardiovascular disease. Why is this particular problem amenable to a systems biology approach?

MICHAEL INOUYE

So when it comes to cardiovascular disease, it's something that we very much care about, because it's really the leading killer in a lot of Western countries and getting a better handle on the pre-clinical models of disease will allow us to get some major runs on the board when it comes to designing therapeutics and other sort of intervention strategies. Really what makes systems biology such a good partner for CVD is that it is a very multifactorial disease. It's very dependent on genetic factors, on environmental factors in terms of molecules also circulating in the circulatory system. That also means that it's a very accessible thing to work with, because you can sample large numbers of people and get the profiles that you need very easily.

DYANI LEWIS

So how did you start going about looking into this area of cardiovascular disease?

MICHAEL INOUYE

So what we were trying to do was link variation within the human genome with variation in metabolomic profiles in circulation to understand gene expression as an intermediary between the genetic variants and the metabolites, trying to get an understanding of the biomarker metabolites in circulation that seem to predict whether some individuals get cardiovascular disease and others don't. So we looked at things like these so-called good cholesterol and bad cholesterol, high-density lipoprotein, low-density lipoprotein, and looking at subclasses thereof. So these are

molecules that are quite complex. Their composition is made up of all sorts of different biomolecules. And with metabolomics, we can actually look at not just the whole thing, but we can start looking at smaller and smaller particles and trying to get a better understanding of the composition.

What we did was that because cardiovascular disease has a strong genetic basis, we basically - through my collaborations with Finland, we looked at about 6600 people from around Finland and looked at little point mutations in their genome and tried to correlate those with these circulating metabolites. And what we can do is we can initially link the SNIPs with metabolites that they seem to be correlated.

DYANI LEWIS

These SNIPs, they're single-nucleotide polymorphisms, so just point mutations in the genome.

MICHAEL INOUYE

Yeah, and then we can say, okay, well, if we actually look at the genes which are close to these SNIPs in the human genome, how are they then associated with the metabolites? By linking these various molecular classes together, we can get an understanding of how the system is operating and then we can progress to look at how it might be operating in terms of atherosclerosis. We ended up finding that there were quite strong genetic associations there. These circulating metabolites, which seem to be associated with cardiovascular disease, seemed to be driven by genetic variants.

DYANI LEWIS

So did you need to choose people who had cardiovascular disease to be part of this study?

MICHAEL INOUYE

Not initially. Cardiovascular disease came out of a lot of the genetics of metabolism that we were looking at. So some of these new genes that we found seemed to be previously involved in atherosclerosis which is a major part of cardiovascular disease. What we ended up doing is once we had identified those SNIPs which were associated with these metabolic networks in circulation is that we began to pull a lot of public datasets and private datasets. As we talked about before, it was very useful to have these available. We started looking at how these SNIPs influence transcription of genes that were physically close to them. We then tried to link those genes to the metabolites that the SNIPs seemed to be associated with and then in turn look at the gene expression in cardiovascular disease and in healthy controls.

DYANI LEWIS

So the genomic data that you had you said was the SNIPs or single-nucleotide polymorphisms. What are they and why are they useful?

MICHAEL INOUYE

SNIPs are mutations in the human genome that have occurred, sort of, somewhere way back in history and just through inheritance have increased in frequency. If you

assay any particular person off the street, you will actually find that there are many, many point mutations throughout their genome that differ from other people in the population. This is where we get that variation in the human genome that we're able then to start looking at how that's associated with the expression of different genes and the levels of proteins and how it might be involved in complex traits and diseases.

DYANI LEWIS

So these aren't necessarily mutations within important genes that are causing some kind of disease or anything. They're just changes that are fairly neutral.

MICHAEL INOUYE

No, they can be both. They can be both within genes and having a very strong association with disease or they can be relatively benign and not have much of a function. We're really just trying to - at this point, getting an idea about which ones are associated with disease and which ones seem to be relatively benign.

DYANI LEWIS

This genomic data you then link to what was going on in the blood of the people that were in the study?

MICHAEL INOUYE

Yeah, using the skills of computation and statistics, we're able to handle millions of these SNIPs that have been assayed across thousands of individuals. We can actually run statistical tests to test whether they are associated with metabolite levels that are in circulation.

DYANI LEWIS

Once you've got the associations between which SNIPs are important for which metabolites going up or down in different people, how does that help you identify the important or biologically relevant genes that are involved in the process?

MICHAEL INOUYE

So with these SNIPs, we can actually look at what genes are close to them and through which their effects might be flowing through and understanding how that's then modulating metabolite levels. So using the human genome and the annotation that a lot of people have been doing on that, we can actually just go online to a website and say, okay, well, where's the SNIP in the human genome and where are the closest genes? We can then look at those particular genes in other datasets and other tissues and try and understand how this SNIP is affecting gene expression levels and how then those might be affecting the metabolite levels which the SNIP itself was associated with.

DYANI LEWIS

Mike, humans are very good at detecting patterns in the world around us, so do you need a computer program to look for these types of networks and associations or can you just pick them out by looking at the data?

MICHAEL INOUYE

We definitely need computers. This is one of the real strengths of systems biology in that the data is just too much for the human eye. You have to deal with literally millions of variables for each sample and try and do mathematics on that. And it's impossible with a paper and pen or in Excel, so we actually have to use computer programs to manage the data and to run statistical tests on.

DYANI LEWIS

The types of networks that are involved, are they all things that you could have, I guess, predicted genes that would have been expressed together or metabolites that would have gone up and down together or are there some surprises in the data?

MICHAEL INOUYE

It's really a mixture of both. So we know a bit about what we're looking for. We can look and see that a lot of the bad cholesterol and subclasses thereof cluster together and they tend to move together. A lot of the good cholesterol cluster together and move together. But then we're also surprised by a lot of things. For example, in one of the studies, we found that the good cholesterol, high-density lipoprotein, actually isn't an entirely coherent thing, that there are subclasses of HDL which move in totally opposite directions to the larger subclasses. They seem to operate more like the bad cholesterol. So it's using these approaches like metabolomics and actually looking at as wide as possible of the metabolic system that you end up seeing things that both validate what we know and new things which really surprise us.

DYANI LEWIS

So when you cast that net of what metabolites to look at, I guess - in a way - that's predetermining what sort of things might be involved. Like how broadly do you cast that net in the first instance of what metabolites within the blood sample you might test?

MICHAEL INOUYE

Really that's determined by the technology and what we're able to glean from - in this case - a nuclear magnetic resonance machine. There are obviously other technologies which we can start mixing in and trying to get a much wider view of systemic metabolism. Technology determines how much of various molecular classes we can quantify, but we also have to be careful about technical artefacts getting into the data. We have to use very rigorous what we call normalisation strategies, filtering strategies to accommodate that so that we can be confident that we can cast as wide a net as possible and not let non-biological things into our studies.

DYANI LEWIS

This metabolomic data contrasts a little bit with the genomic or transcriptomic data where you're really measuring the whole genome.

MICHAEL INOUYE

Yes.

DYANI LEWIS

By combining the metabolite data with the SNIP data that you have, you've been able to pinpoint a couple of genes that are involved in cardiovascular disease. How do you go about confirming those associations?

MICHAEL INOUE

So in terms of validating the role that these genes might be playing in metabolism and atherosclerosis, we take these genes that we identified from the SNIPs to metabolite associations and we start to look at how the expression of these genes is associated with metabolites in independent datasets. And through that, we can be more confident that - outside of the original data that these genes were found in - that they are involved in metabolic processes that the SNIPs themselves would have suggested that they were involved in. Using other independent datasets - so you can see the idea of independence that really makes us more and more confident that what we're seeing is real, that we can then go and look at these genes in liver tissue in mice and find associations with atherosclerotic lesion area.

We can also go and look at these genes in human arterial tissue and look at arterial tissue that is relatively healthy and those which have atherosclerotic plaques and look at the expression of these genes and how it's different between those two, between disease and health. We actually find that indeed when we look at these genes in humans in arterial tissue that they are four to six times increased in expression in the atherosclerotic plaques, so this gives us quite a bit of confidence that the genetic associations that we originally saw with metabolites that were themselves associated with cardiovascular disease do indeed indicate genes that are involved in that process.

DYANI LEWIS

You're listening to Up Close. I'm Dyani Lewis and my guest today is Mike Inouye, a systems biologist who has been investigating the complexities of cardiovascular disease. Mike, you chose to work on a population of Finnish people, specifically from Helsinki in the first instance. Why is the Finnish population particularly good for this kind of study?

MICHAEL INOUE

So the Finnish population is an especially attractive population to work with, because they're very positive on medical research. They're very willing to enrol and participate in medical research studies. This is largely because they have a very unique genetic profile, especially around Europe, and that predisposes them to a lot of diseases which are at much higher frequencies in their population than elsewhere. So the people know this and they want to know why. They're very open to the explanations that science can provide them. That means that a lot of data ends up being collected on them. Finns tend to also be very technologically oriented, so they were very willing to embrace very early -omics technologies and start to profile the genomes of people in the population and their transcriptomes and metabolomes. All this data together - because they also have electronic hospitalisation records, we're now able

to start linking those together with all these various molecular profiles and genomic profiles and begin to look at the relationships between all of this data and disease.

DYANI LEWIS

Is it just the population interest in these kind of technologies or is it also something about the genetics?

MICHAEL INOUYE

The genetics as well is quite interesting, because you get a very unique SNIP profile - as we discussed earlier - amongst the Finnish people. You also get enrichments for low-frequency loss-of-function variants, so these genes which actually just by chance end up being inactivated in humans that apparently are walking around on the street healthy. You get profiles of these which are quite unique. That means that we are able to look at things that we're not able to in other populations.

DYANI LEWIS

If you take a community like here in Australia or I guess in many parts of America, they would be far more varied than, say, in Finland. Is that also a contributing factor to why the Finnish population is attractive?

MICHAEL INOUYE

It is. So when you have people with different population histories who enrol in the same study and you start looking at all that data together, that is something you have to be conscious of. There is what we call population structure confounding within the data. That tends to be also true with the Finns, but it's somewhat less when you start mixing in people who are originally from Italy, people who are originally from the UK, people who are originally from China, as we have here in - especially in Melbourne and Australia more widely.

DYANI LEWIS

So where does this research go to from here?

MICHAEL INOUYE

What we can do actually with some of these systems biology studies is we can continue looking - not just in mice but in humans - at genes which are knocked down or knocked out in people and look at how that is influencing biomarkers in circulation and at their probability of getting disease later on, so this represents yet another layer of replication for these studies that we've discussed today.

DYANI LEWIS

Based on the kind of networks and associations that you can pull out of the data that you're collecting, can we make predictions about factors that are associated with cardiovascular disease and perhaps into personalised medicine even?

MICHAEL INOUYE

So I wouldn't quite go that far at this point. What we're doing is we're trying to identify

therapeutic targets which are relevant at least to the population and the data that we're working with. The associations that we find, I mean, they are statistically predictive, but the size of their effects is still not large enough that we can start talking about things like personalised medicine and selecting different treatments for different genetic and molecular profiles of individuals. But it's something that we want to work towards in the future.

DYANI LEWIS

But it very much comes back to then identifying the right drugs to go on to clinical trials.

MICHAEL INOUYE

It definitely is. I think that's certainly something that we hope ends up coming out of this work and follow-up studies.

DYANI LEWIS

Mike Inouye, thank you for being my guest today on Up Close and discussing your work in systems biology.

MICHAEL INOUYE

Thanks, Dyani.

DYANI LEWIS

Dr Michael Inouye is a senior research fellow and head of the Medical Systems Biology Lab in the Departments of Pathology, and Microbiology and Immunology at the University of Melbourne. He is also an honorary senior research fellow at the Institute of Molecular Medicine, Finland, at the University of Helsinki. If you'd like more info or a transcript of this episode, head to the Up Close website. Up Close is a production of the University of Melbourne, Australia, created by Eric van Bommel and Kelvin Param. This episode was recorded on 23 October 2013. Producers were Eric van Bommel and myself, Dr Dyani Lewis. Audio engineering by Gavin Nebauer. Until next time, goodbye.

VOICEOVER

You've been listening to Up Close. We're also on Twitter and Facebook. For more info, visit [upclose.unimelb.edu.au](http://upclose.unimelb.edu.au). Copyright 2013, The University of Melbourne.

© The University of Melbourne, 2013. All Rights Reserved.

---

**Source URL:** <http://www.upclose.unimelb.edu.au/episode/273-life-numbers-systems-biology-and-its-approach-researching-disease>