Okay. So our next speaker is Ihor Lemischka. I'm not going to try to pronounce it. Sorry.

IHOR LEMISCHKA: Oh, actually, I should thank you for not trying to pronounce it.

CHRIS CHIN: Oh, okay. Okay. There you go.

IHOR LEMISCHKA: Okay. First of all, a disclaimer. I'm neither a bioengineer, nor do I do anything beyond the most rudimentary types of imaging.

But I hope that you'll see that some of the problems that we're grappling with in our systems -- namely, pluripotent stem cells -- are very similar to the ones that I've been hearing about this morning, and to name two, one is the challenge of integrating multiple different types of datasets originating from very different experimental platforms.

And the second problem being how do you actually view the dynamics of regulatory networks as they process biological information?

The title here is my generic title, which essentially allows me to talk about anything I want, but, in reality, there will be two vignettes about

pluripotency, and, if I have time, there's going to be an et cetera as well.

So this is an illustration, a snapshot of the complexity of the mouse embryonic stem cell regulatory network as of about two years ago. This is hopelessly outdated, although clearly complex.

And these are all transcription factors here. And I'll point out that all of these edges and interactions and signs are based on experiments. So this is all based on experimental, empirical data and no computational stuff as yet.

Now, the problem with these things, as complex as they are, is that network diagrams are static portraits of essentially a dynamical system, and, as such, they provide only a wiring diagram of a regulatory network. And, by definition almost, because of the way the information is gathered, the additional complexity is that this network that I showed you in the previous slide is actually an average network that exists in a population of stem cells.

Now, to compound the problem here, most network diagrams incorporate only one or two molecular layers or levels of information, and, usually, these are

limited to transcriptome profiling or epigenome profiling, things that are now more or less routine in the Omix world.

So what we really want to do is to develop ways of addressing what actually happens in real time as a stem cell begins to roll down this Waddington landscape, and, in the end, how the regulatory network is processing biological information -- a series of inputs, say -- and how that information processing can be correlated with a defined, measurable change in cell fate.

So how to measure network dynamics. And, in essence, what we'd like to be able to do is to analyze a sulfate decision at multiple molecular levels as this decision is occurring in real time, and, in essence, convert regulatory snapshots into regulatory movies.

So the system that we use is illustrated here, and, basically, it's an RNAI or SHRNA-based system where, on a single vector, we remove an endogenous gene product -- in this case, the homeodomain transcription factor NANOG -- and we essentially have a doxycycline-controllable cassette downstream of that where we -- this cassette encodes in a doxycycline-dependent manner an SHRNA immune version of the same gene

product, and, in this case, NANOG. We've done this for many different things now. It's a very robust genetic complementation platform.

Now, these cells, then, depend crucially on doxycycline for maintaining their self-renewing pluripotent state. The system is very tight. They're fully pluripotent in the sense that they're able to give rise to germline kimers after culturing them in docks and following implantation into blastocysts.

So what we can do, then, using this system is culture the cells in doxycycline, remove doxycycline, shut down the given gene product -- in this case, NANOG -- and then, in the same experiment, over time, profile global changes in the epigenome, so various histone modifications, DNA methylation, elongating POL2 localization to measure active transcription.

We can also profile changes -- dynamic changes in MRNA levels, and, together with Tony Whitten's [phonetic] group in Manchester in the UK, we have been able to profile the entire nuclear proteum by mass spec.

And, finally, not to be not complete -- for completion's sake, but also for a good reason that

I'll show you in a minute, we've also been profiling microarnase [phonetic] in the same experiment.

So what we're trying to do here is introduce a very defined perturbation [unintel.], simply removing one gene product. That removal, that event triggers a cascade of biological and molecular changes, and we're trying to profile these changes ranging from changes in the epigenetic landscape all the way to the proteum here, nuclear proteum. And I won't have time to show you, but we've also now been incorporating global changes in the phosphel [phonetic] proteum as a function of time after removing one thing.

So this raises a bunch of challenges, and this just -- I won't have time to go into this, but, together with Avi Mayan's [phonetic] group at Mt. Sinai, we've been able to develop these interactive movie-like, two-dimensional, heat-map views. And if you go on line, one can click this and you can actually watch clusters evolve over time in the Y axis here and across molecular levels on the X axis -- in this case, H3K4 trimethylation, DNA methylation, RNA and, of course, you can integrate the protein-level changes here as well.

So this gives us -- In the background here, all these pixels are live, so you can click on them and retrieve as much prior knowledge information as we've put in there.

But some of the things that we have are things that you can ask -- Which of these RNAs, which of the proteins encoded by these RNAs have been documented to interact physically with each other from prior studies? So there's a wealth of information here, and I think quite a bit more sophisticated than this snapshot traditional view here.

Now, just to give you an example, a teaser, of one of the things we've been able to use this for to make biological predictions, which any database is only as good as the predictions that it makes.

So we noticed, in our previous studies, that -- and current studies -- that when you compare changes in nuclear proteins to changes in the MRNAs that encode these proteins, quite frequently, over 40 percent of the time, we don't see a direct correlation between RNAs going up and proteins going up. In fact, in 40 percent or greater of the time, there appears to be no correlation or even an anti-correlation.

So we went ahead and tracked, then, attempted to merge the changes in micro RNAs occurring over time to changes in their predicted targets, MRNA targets that occur over time, and, currently, we've made a lot of predictions here, and we're currently testing them.

So this is, I think, one value of this type of an approach, and, of course, there's many others that I don't have time to get into.

So this is kind of a summary of what I've told you so far. We do a very defined perturbation on day zero. This doesn't just work for transcription factors.

We've done epigenetic modifiers, kinases, phosphatases, RNA binding proteins, and you can track then what happens over time across these multiple levels. And, of course, then you can measure what the outcome, the cell-fate outcome is using your range of cell-fate markers or germ-layer specific differentiation markers.

So the idea is to correlate a specific perturbation. The black box here, which is the information processing part, the machinery of this with a defined output after the process is done.

Now, one thing that we've also been doing with Ari's group is to construct a large prior-knowledge database that incorporates essentially everything that has been done in the pluripotency field and basically integrating epigenetic chip, chip, chip seek, protein, protein interaction data, loss-and-gain-of-function experiments, followed by temporal profiling of changes, phosphoproteomics, proteomics, functional RNAI hits, histone modifications, array datasets and micro RNA and predicted targets.

And this is called ESCAPE. I'm convinced that the person, Hulay Zoo [phonetic], who did this database, along with these people, basically thought that ESCAPE would be a cool acronym and then retrofitted it to words. Embryonic Stem Cells Atlas of Pluripotency Evidence doesn't exactly roll off the tip of your tongue, but there you have it.

Now, ESCAPE has been very useful, because what ESCAPE has built into it are functions to predict the regulatory network wiring diagram or architecture. So you can do this based on temporal profiling of gene expression changes and epigenetic chip, chip, chip, seek experiments.

So what we've done then is to construct such a network that I'll show you in a moment and then basically trained the network, based on an independent dataset, where we profiled a series of markers and individual embryonic stem cells grown under two different conditions, and then, basically, made some predictions here and tested them.

And this is the kind of network that you get. This is a 30-node network that has not only the core pluripotency gene products, but also outreach to early commitment markers defining the three different germ layers as well as trophectoderm.

Now, for each of these nodes, we've also, for various reasons that, again, I don't have time to go into, we've used the bullion [phonetic] approximation here. We've been able to extract out for each node a bullion transition function, given the inputs to that node that regulate the ultimate expression of that node.

And we've done this for every one of the nodes. So, now, we can make predictions, run dynamical simulations and ask what happens if we do an in silico RNAI, either single RNAIs or combinatorial RNAIs, run the dynamical simulation, wait until a new attractor or new steady state is achieved or is obtained. This

gives you new values of zeros and ones for all the remaining nodes in the network.

And then we went and did the actual RNAI experiments in the lab and tested, by doing real PCI measurements, the expression levels of all the remaining nodes.

So these are the predictions here. These are the RNAI experiments removing ESRB, NANOG AK4 [phonetic] individually or in combination.

These are the actual measurements of the remaining nodes on the Y axis here, and this is the overlay which basically says, in yellow, means we get good experimental validation to the computational prediction. And this is about a 75- or 80-percent validation rate, which, for a network of 30 nodes or so, is, from what I'm told, is pretty darn good.

So we're building this network out further, of course, also doing a lot more experiments. And I'll just show you one experiment.

So here are two fluidyme-based single-cell profiling experiments where we profiled 96 individual ES cells across 96 markers, both pluripotency as well as commitment markers.

And, basically, these -- In red, of course, means the marker is expressed. In green it's not expressed. And we did this under two different conditions, either serum plus leukemia-inhibiting factor, LIF, or serum-free, two-I [phonetic] and also containing LIF.

Now, the important thing to remember is that, in both cases, the cells are fully pluripotent. So they're able to give rise to all three germ layers. They're able to contribute to the germ line, robust kimerism.

But, as you'll see in a minute, when we used these single-cell data, both from the serum conditions and serum-free conditions, to train the predicted network from -- extracted from the database, you basically get different versions of the network.

Here's serum plus LIF. Here's two-I plus LIF. And this is the comparison. And the remarkable thing is that this single regulatory network responsible for the pluripotent state of these cells can be extensively and even dramatically rewired depending on what the culture conditions are, either with serum, which is, of course, [unintel.] and also serum free, completely chemically defined medium.

So this is intriguing to us, and leads us to propose sort of a model like this. But, in fact, the Waddington landscape isn't a point, but it's actually a plain, and that the pluripotent state actually doesn't exist, but is actually an amalgam of many, many different alternate states that are constantly fluctuating when looked at in single cells.

And we're trying to address this now using more sophisticated technologies like putting multiple color markers into the same ES cells and try to see if we can then sort out cells that express, you know, twice as much NANOG as AK4 or various combinations and see if we can correlate these different values here or different states of the pluripotency network to different biological properties of these cells.

Okay. So that's vignette number one. I can see I'm not going to be able to have time to go into et cetera, but that's okay.

So some years ago, a couple of years ago, we revisited this idea of kinase phospo-regulatory signaling in ES cells that the field thought was all done and signed, sealed and delivered. And we simply asked are there other kinases or phosphatases that are important in maintaining the ES cell or pluripotent cell state?

And this is going to be a good example of, you know, a counter argument to what reviewer number three always says, that this is not hypothesis-based science, because, as you'll see, it actually leads to a completely new mechanism of pluripotency versus a commitment switch that nobody thought existed in ES cells.

But to make a long story short, a post doc, Dung Fang Lee [phonetic], and a student, G. Sue [phonetic], in the lab basically came up with a list of a little over 104 kinases, phosphatases, a few regulatory subunits that had properties of being enriched in undifferentiated ES cells or I PS cells.

They then went and knocked each one of these down using SH RNA, identified a set of five here that were important for pluripotency; i.e., you take them away, cells differentiate.

They built rescue clones, these genetic complementation clones, for each one, showed that pluripotency depends, as expected, on doxycycline looking at undifferentiated morphologies, and this is backed up by a lot of marker analyses as well.

I'm just going to focus on one of these. They're all interesting, and if somebody wants to ask me later I'll tell you why. But we focused on aurora kinase A, and aurora A is a member of three homologous [unintel.] kinases. They're mitotic kinases in somatic cells. They're linked to cancers. Knock out of aurora A leads to embryonic lethality at about the 16-cell stage.

Now, we also showed that if you look at databases of gene-expression profiles, aurora is very highly expressed in very primitive cells, so overlapping to a large extent with cells or embryo stages that also express the NANOG and AK4 and, to some extent, SOCKS [phonetic], too, pluripotency genes. And aurora is much less expressed in adult or somatic cell types.

Now, we showed also that aurora expression is down regulated following ES cell differentiation both in embryotic bodies and driven by retinog [phonetic] acid RNA level, also, by Western blots looking at the protein level.

Moreover, if you use the NANOG rescue, SOCKS 2 rescue and AK4 conditional clone, you're going to show that removing each of these transcription factors also results in dramatic down regulation of aurora.

Now, of course, we did what we always do. We built the rescue cells and showed that these are dependent on doxycycline for maintaining their pluripotency.

And then we could do experiments taking away or adding back dox and seeing what happens globally. And what we noticed, of course, is, first of all, that taking away aurora down regulates all canonical set of pluripotency factors, both at RNA and at the protein level relative to a control clone where there's no changes.

And, moreover, loss of aurora doesn't just trigger loss of pluripotency and differentiation, but it triggers mesodermal and ectodermal differentiation. So this is not random differentiation, as shown by profiling a whole range of markers at the RNA as well as at the protein level.

Now, when we did genome light expression analysis, we noticed that there might be a link to aurora signaling and the P53 signaling pathway. And we noticed this because we showed that, in fact, if you take away aurora by removal of dox from these cells, you get shooting up of P53 targets.

Moreover, if you inhibit aurora kinase with chemicals, you get the same thing, and, interestingly, the expression levels of P53 itself actually go down, although, in spite of the fact that P53's targets go up.

So this led to the hypothesis that perhaps aurora is controlling P53 and that's a requisite event for maintaining pluripotency.

Now, in spite of what reviewer number three says, the roles and mechanisms of P53 in regulating pluripotency are essentially almost completely [unintel.].

There was that set of five or six papers showing that you can improve IPS reprogramming by removal of P53, but the mechanisms for that have been largely unexplored. There's a couple of papers here, but, still, largely unknown.

So what we did then was to show, without going through a lot of data, we showed the P53 is actually phosphorylated on two Syrians by aurora, Syrian 212 and Syrian 312 in the mouse and that phosphorylation of Syrian 212 actually inactivates P53's transactivating transcriptional activity.

And, again, doing a lot of experiments here using phosphomimic mutants of P53 we showed that without this phosphorylation event, P53 actually sits on and up regulates massive programs of ectodermal and mesodermal differentiation.

So then we went and asked, Well, what happens if you remove aurora from mouse embryo fibroblast during Yamanaka reprogramming? And what you see is, if you remove aurora from MEFS, you actually -- you don't get five full decreases of reprogramming. You get zero reprogramming. So aurora is absolutely required for IPS reprogramming.

Now, over-expression of aurora, similarly, not so dramatically, but it improves IPS reprogramming efficiency. And all of this is done with only mild changes in the proliferative rate of the starting fibroblasts. So it's a mitotic kinase, but it seems to be acting in a non-cell-proliferation-dependent manner here. And we have other data that proved that more rigorously.

Now, what's interesting, also, is that if you remove aurora and, at the same time, remove P53 or ARF [phonetic], you completely rescue that requirement for

aurora, so you restore Yamanaka reprogramming to its normal level.

So this puts P53 immediately or closely downstream of aurora signaling both in maintenance of pluripotency in wild-type ES cells and also in the acquisition of pluripotency during Yamanaka reprogramming. And we showed further that it's the same Syrian 212 phosphorylation event that is required in both cases.

So this is the model here. Aurora sits upstream of P53, phosphorylates P53. This inactivates P53's transcriptional activity. Without aurora, P53 drives the expression of ectodermal and mesodermal gene expression differentiation programs, and this is in ES cells, but, also, we believe that the function of aurora is very similar during the reacquisition of pluripotency by Yamanaka reprogramming.

How much time do I have? I might be able to get to my et cetera.

MALE SPEAKER: [Unintel.].

IHOR LEMISCHKA: Oh, yes, I've got 3-1/2 minutes. I have time for et cetera.

So this is a really new project, and this is attempting to program hematopoietic stem cells directly from fibroblasts using defined transcription factors.

And programming from one cell state to another has been shown as early as the late ‘80s from Hal Weintraub’s [phonetic] lab with [unintel.]. And, now, there’s been a spate of other examples showing direct conversion from fibroblast to different cell fates without going back through an uncommitted progenerate or stem cell.

But capturing, going from a fibroblast to a somatic stem cell that still can balance self-renewal and multi-lineage differentiation ability has been really difficult. And, in fact, in spite of 25 years of intense work by good labs, it’s been almost impossible to get any transplantable hematopoietic stem cells in vitro from a mouse or a human embryonic stem cell.

So this was a big challenge, and, basically, what we decided to do was to not maybe directly focus on converting fibroblast to hematopoietic stem cells, but rather to focus and try to kick start a developmental program that during normal mouse development and human

development generates hematopoietic stem cells, namely try to regenerate the hemogenic endothelial stage.

And we did that using a double trans-eunuch [phonetic] reporter cell line that basically has -- is very specific in terms of driving H2BGFP to only the very primitive stem pro-generative compartments of mouse bone marrow. And no need to go into the details of this, but if you look at prospectively isolated stem cells from bone marrow -- and this is work done by Terry Moore [phonetic] my colleague at Sinai -- you can show that this double transgene very specifically marks the early compartment.

So we used fibroblasts from this double transgenic and selected out a set of 18 transcription factors based on a variety of profiling studies, both from Terry's lab and from the available databases.

And then we stuck all 18 of these into the cells, in the embryo fibroblasts carrying the double transgenic, and basically screened for turning on for basically H2BGFP positive cells that would reflect the activation of this double -- this reporter system.

And what's really -- and we got that. And what was really interesting was we showed that we can get these

structures, and I like to call these two-dimensional vasculature, because they have endothelial markers as well, but you can see here budding off of what's going to be a non-adherent cell from this structure here, and that cell is now turning on the nuclear GFP.

So we did what everybody does in this business, basically progressively removed individual factors. And we now have a combination of four transcription factors that's most robust, three transcription factors is sufficient, but not as robust, and basically are now trying to work with this.

Now, these four factors are [unintel.] 2, EDV 6, GFI 1B and FOS [phonetic]. And we think that FOS, although it's by no means specific to hematopoietic stem progenerative cells, our hypothesis is that FOS is doing, in this context, what C-MIC [phonetic] might be doing in Yamanaka reprogramming.

And basically, we don't have function yet, but we have a lot of marker expressions showing endothelial markers as well as HSC markers in individual cells. And this is the rudimentary imaging I was talking about.

And, basically, just to cut to the chase, we've done RNA seek at different stages of this procedure here, starting with MEFs day 20 and finally CD45 KIP positive and KIP negative cells. And comparing to gene atlas, very gratifyingly to about 2,000 datasets, we see that the very top hits from our RNA seek analysis of the candidate HSCs most closely match bona fide HSC profiling studies from the database.

So we've very encouraged. We have a long way to go. This is all mouse. We have studies ongoing in the human system as well, and I'll just end there.

And these are the raft of people. Philippe back here. Terry, Avi Mayan, other people in the lab, as well as a raft of collaborators here. So thank you. [Applause].

CHRIS CHEN: [Unintel.]. Maybe I'll ask one. Yes, go ahead.

MALE SPEAKER: [Unintel.].

IHOR LEMISCHKA: Yes, that's obviously a good question. And we don't know what aurora does or doesn't do in adult stem cells. I mean, everything we've done here [unintel.] embryonic stem cells and

[unintel.]. So I can’t really answer that [unintel.]. I may be wrong. We certainly [unintel.].

CHRIS CHEN: Okay. Thank you very much.

IHOR LEMISCHKA: Thank you. [Applause].