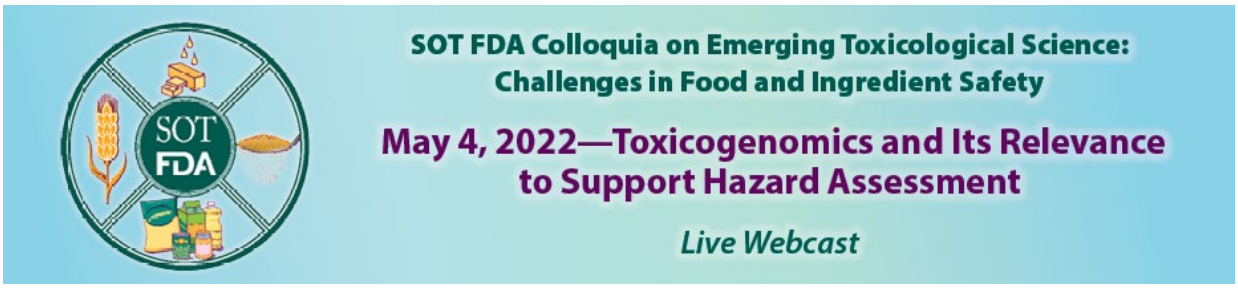

Real-Time Captioning

Note: This is not a transcript.

Schedule

9:00 AM–9:10 AM	Welcome from FDA Steven M. Musser, Deputy Director of Scientific Operations, CFSAN Welcome from SOT Michael Aschner, SOT President
9:10 AM–9:15 AM	Overview and Speaker Introductions Stephen Edwards, Colloquium Co-Chair, RTI International
9:15 AM–10:00 AM	Overview of Toxicogenomics and Its Contribution to Toxicological Science Carole L. Yauk, University of Ottawa/Health Canada, Ottawa, Canada
10:00 AM–10:40 AM	High-Throughput Transcriptomics for Chemical Bioactivity Screening and Tiered Hazard Evaluation Joshua A. Harrill, US EPA, Research Triangle Park, NC
10:40 AM–10:50 AM	Break
10:50 AM–11:30 AM	Toxicogenomics to Support Hazard Assessment Russell Thomas, US EPA, Research Triangle Park, NC
11:30 AM–12:10 PM	Application of Toxicogenomics in Toxicity Assessment of Chemicals in Food and Cosmetics Jorge Naciff, Procter & Gamble Company, Mason, OH
12:10 PM–1:00 PM	Roundtable Discussion Moderators: Co-Chairs All speakers

Welcome from FDA
Steven M. Musser, Deputy Director of Scientific Operations, CFSAN

OK. Good morning. Sorry for the technical difficulties this morning. Part of the world we live in, I guess, or are still adjusting to. I hope everyone's had a chance to join now or are joining shortly. Welcome to the Society of Toxicology's and the Food and Drug Administration's colloquia on Emerging Toxicologic Science Challenges in Food Safety and Ingredient Safety.

Today's topic is "Toxicogenomics and Its Relevance to Support Hazard Assessment." Next slide, please.

My name is Steve Musser. I'm the Deputy Director of Scientific Operations at the Center for Food Safety and Applied Nutrition at FDA. And I'm very pleased to be here doing the brief introduction and ground rules and background on the colloquia. Next.

So, this is a long-running partnership, this colloquium series, between the US Food and Drug Administration Center for Food Safety Applied Nutrition and the Society of Toxicology. We are very happy to have the Society of Toxicology as a partner in this colloquia series; it's been ongoing for eight years. We'll have two events this year; this is the first of those two. This is the 27th such symposium that we've had. And, you know, they've really been excellent series, every single one of them. And today's will be no different.

You know, we, our purpose here is to stimulate dialogue among the leading toxicology experts in future-oriented toxicological science related to all foods, assessment of foods, risk assessment, and food ingredients. And, and by that, I mean, you know, we really interested, at least on FDA's part, in knowing about these new approaches, but how are they validated, what are their what are their boundaries, what are what are the benefits, and what are the challenges of using them? How do we integrate them into modern approaches to risk assessment and chemical safety? So, from our perspective, these, these colloquia could simply not be more valuable.

It is not, however, this colloquia, a forum for soliciting regulatory advice or discussing food or food ingredient regulatory issues. It's important for me to say that because these colloquium series need to be confined to purely scientific discussions. Next slide, please.

Finally, before I turn it over to Dr. Aschner. The series are open to the public via webcast at no cost. We hope at some point to, again, have them available in person. But currently, they're available via webcast. It's a global audience, so people from all over the world. Industry, private citizens, government, and any others that I've left out are welcome and often attend. And finally, the recording and the slides are all going to be available at SOT's toxicology.org website.

Thank you very much for attending. I hope that you were all able to navigate the, the signing online and the web in our series, and I appreciate your patience in getting started. And with that, I'm going to turn it over to Dr. Aschner, and he can take it away. Thank you.

Welcome from SOT
Michael Aschner, SOT President

Thank you very much, Steve. And my name is Michael Aschner, and I want to welcome you on part of the Society of Toxicology. I am the incoming, or just started my presidency a couple days ago. So, I look forward to working with that all of you, and my regular job is at Albert Einstein College of Medicine. Can I have the next slide, please?

So, just a few words about the Society of Toxicology, its vision, and its mission. As stated here, we exist basically to create a healthier and safer world by advancing the science and increasing the impact of toxicology, and some of the strategic objectives that we have developed over the last few years are to promote trans-disciplinary science on cooperation

with other disciplines, expand global outreach, engagement, and collaboration, provide opportunities to capitalize on scientific advancements, to develop communication partnership with key influencers, and I believe one example of such interaction is what we're doing today, as well as the last bullet, which is to promote the use of sound science in decision making, which I'm sure we'll be hearing about today. Next slide, please.

So, this colloquium, as Steve mentioned, has been ongoing for, since 2014, I believe. And you can see in this slide, which was updated by Betty yesterday, and I want to thank her for putting together actually some of my slides, you can see that this is very well attended. I don't know how many people are from foreign countries today. I see that we have 70 participants, which speaks volumes on the importance of this colloquium. And you can see that between 2018 and 2022, we had 45 countries represented in these colloquia, so it definitely has incredible reach. And I'm sure it will continue to do so. Next slide please.

I won't read all the names here, but I want to thank everybody on the SOT FDA Committee. I know you've worked very hard to put this together; these are the members for this coming year, 2022–2023, I saw some of you online. So. thank you very much, Steven and many of the other members who are who are here today. I know Suzy has been also very active on this committee in the in the past few years and definitely promoted a lot of the issues when she served on Council.

And I just have a couple more slides, I believe, just to introduce a little bit the title, we can skip this one. And so, this is a scheme I guess that I got from, from a paper, I believe it's from Dr. Liu. And I think what is in front of you today is a very interesting topic, but I think it would be fair to say that there are some challenges as well. And this slide and the next one, basically, identify some of these challenges. I look forward to hearing about those. If we go from the bottom up of this pyramid. You can see that there are several key questions that are listed on the left. For example, what's the future of toxicogenomics study design? On the right hand side we have the options and the solutions. so do you use 2D cultures? Do you use 3D cultures? Do you use primary astrocytes if you're interested in the brain or do you use different cell lines?

If we go further up in this pyramid, and I'm not going to read everything, you can see that the question is, how should evolving technology in toxicogenomics be addressed? Do we use just mRNA, do we use the RNA-seq, do we do single cell RNA, and so forth going all the way up to, which of the toxicogenomic results is the one that's really the most relevant for the mechanistic information for predicting that toxicity and for helping in regulatory decisions? Next slide, please.

And I'm going to skip this one I'll go to the, to the last one, because I'm running over. So, I hope that the outcomes from this session that we're having today is to promote better understanding of the advances in this field. As you all know, this is moving very fast, and how the generated genomic data can enhance decisions in risk assessment. One issue that hopefully we'll discuss today is the appreciation for reproducibility. I think those of you, as I have, that have used different types of platforms, know that sometimes it's not easy to standardize and reproduce the results. So, how do you go about utilizing the data in the setting of regulatory decisions? And finally, I don't know if this will be discussed, but hopefully we can also start to better understand how machine learning, how artificial intelligence, and those kinds of factors will help us in predicting toxicity using data that are derived from a toxicogenomics.

And with this, I'll turn the platform to the moderators again. I want to welcome everybody on behalf of SOT, and I'll try to stay on till the end. And I look forward to an enlightening and productive discussion today. So, thank you very much.

Overview and Speaker Introductions
Stephen Edwards, Colloquium Co-Chair, RTI International

And thank you to Drs. Musser and Aschner for that great introduction. On behalf of Dr. Xiugong Gao and myself, it is our pleasure to welcome you to this colloquium on toxicogenomics and its relevance to support hazard assessment. As you've heard from Drs. Musser and Aschner, the field of toxicogenomics has matured to the point where the results from this technology should be useful for assessing the potential for chemical toxicity.

Today's colloquium will review the state of the science and explore existing applications for toxicogenomics for toxicity testing. We're then going to invite all of the speakers and attendees to participate in a roundtable discussion on the future of toxicogenomics, the remaining challenges and potential solutions that Dr. Aschner just mentioned, and additional uses of this technology moving forward. Our four speakers today are all experts in the area of toxicogenomics and have been heavily focused on the use of toxicogenomics to support hazard assessment.

Dr. Carole Yauk is our first speaker, she is professor of biology and holds the Canada Research Chair in genomics and environment at the University of Ottawa, where she focuses on the development and implementation of genomic tools for health risk assessment of environmental exposures and on improving regulatory assessment of heritable genetic effects. Before that she was the lead science for the genetics laboratory of the Environmental Health Science and Research Bureau at Health Canada for 18 years. Carole will give us an overview of the field of toxicogenomics and introduce the first of several case studies illustrating how these data can be used to support toxicity assessment.

Dr. Josh Harrill is a cellular and molecular toxicologist with the US Environmental Protection Agency's Center for Computational Toxicology and Exposure. He's a lead investigator for CCTE's high throughput transcriptomics and high throughput phenotypic profiling hazard screening programs that focus on the use of broad-based high content profiling assays for concentration response screening of environmental chemicals. Along with Carole and others Josh recently led an effort coordinated by the Organisation for Economic Co-operation and Development, or OECD, to define the criteria required for toxicogenomics data to be suitable to support decision making and this directly addresses some of the reproducibility issues that Dr. Aschner highlighted. His talk will focus on the state of the science for toxicogenomics and things to consider when using toxicogenomics data to support Hazard Assessment decisions.

Dr. Rusty Thomas is the director of the Center for Computational Toxicology and Exposure at the US EPA. His center has led the way for the use of new alternative methods including toxicogenomics for toxicity assessment. In addition, in 2013 he published one of the seminal papers demonstrating the utility of toxicogenomics to support risk assessment. He also serves as co-chair of the OECD extended advisory group on molecular screening and toxicogenomics. Rusty will tell us about how the US EPA is incorporating toxicogenomics approaches into their decision-making processes, and will discuss the international efforts led by the OECD to promote the use of toxicogenomics.

Dr. Jorge Naciff is a member of the Global Product Stewardship, Central Product Safety. organization in the Procter and Gamble Company, where he has been providing human safety support for multiple product categories, mostly in upstream technology and R&D. He has been doing research in the field of animal alternatives for safety testing for over 21 years, including the application of toxicogenomics for toxicity assessment. In today's talk, he will provide a series of case studies, highlighting the use of toxicogenomics to support toxicity assessment of chemicals in food and cosmetics.

Following our talks, we will invite all of our speakers to join a roundtable discussion, answer questions from the audience. We encourage you to enter the questions in the Q&A box at any time throughout the morning and we will cover as many questions as possible during the roundtable discussion. Without further ado, let me introduce Dr. Carole Yauk, who will provide us with an overview of toxicogenomics and its contribution to toxicological science. Thank you, Carole.

Overview of Toxicogenomics and Its Contribution to Toxicological Science
Carole L. Yauk, University of Ottawa/Health Canada, Ottawa, Canada

Thank you very much. Just gonna share my screen here. Can everybody see this now?

Stephen Edwards: Yes, we can.

Carole L. Yauk: Okay. I'm just going to try and move this. There we go. Okay, thank you very much. It's a real pleasure to be here and an honor to be invited to present to tell you a little bit about toxicogenomics and its contribution to applied toxicological sciences.

I have no conflicts of interest to disclose. Today what I'd like to do is review some basic principles and some recent advances, focusing on transcriptomic applications in toxicology. Ans we'll review knowledge gained from some recent case studies that we have been conducting in partnership with Health Canada in decision making. So, I'll provide some background, why we're doing this, and then dive into some details on toxicogenomics, how it works and basic principles, what we like about it, how we're using it, including development of biomarkers and dose response evaluations, and then some case studies. The first one on one chemical and then following that two case studies on small groups of chemicals. So, let's get started.

We all know conventional toxicology has served us incredibly well for many decades, but it does have its limitations. For example, it's primarily relying on the analysis of apical endpoints or diseases associated endpoints, and these tests are low throughput and take a long time. This requires high-dose exposures and low-dose extrapolation and animal to human extrapolation, which is challenged by the fact that many of these tests don't provide mechanistic information. And finally, they use large numbers of animals, which is costly and by today's standard increasingly considered unethical. And as a result, we've had many proposals for a major change in toxicity testing. Some of the important reports shown up at the right and incredible momentum around the world toward this objective. So, we want modern testing to be more human relevant faster, cheaper, cover more biology, and be more ethical, mechanistic, predictive, and protective.

And in order to do this, we're harnessing all kinds of new data sources and today we're focusing on toxicogenomics or TGX. This could be any of the ‘omes, many of them shown over here on the left, and what these ‘omes do is open up this black box which is the mechanistic response to the chemicals and analysis of these ‘omes requires close collaboration with data scientists and bioinformaticians because they produce large, complex data sets that are challenging to interpret and so data science, bioinformatics, and biostatistics is a key component of these analyses. And today, we're going to focus on transcriptomics.

We like transcriptomics because it provides very early mechanistic biomarkers; we can observe gene expression changes that are disease predictive at very early time points, well before visible changes like tumors, for example in this in this diagram, occur. These transcriptomic changes I'll argue today are occurring at doses that are similar to the doses that caused the disease associated changes. And because it's a global gene expression profiling approach, it produces rich mechanistic information that can help us to understand potential human relevance and mode of action.

We like it because these technologies have been around for a very long time now—beginning in the 80s with PCR through the 90s when QPCR was invented and DNA microarrays were invented. The major uptake of DNA microarrays in the 2000s and the invention of RNA sequencing, and by 2010, the widespread use of microarrays but also opening up the possibilities of RNA sequencing to the world through technologies like Illumina Hi-Seq, which really increased efficiencies and brought down costs, and more recently, now we're moving into the era of single cell RNA sequencing, spatial transcriptomics, and multi omics approaches.

So, we like it because these are established technologies, and they also produce reproducible data, and I'm going to show you some of that today. The technologies themselves in comparing the same samples yield, with the same pipelines, virtually identical results, but biology can be messy. So, here we have some of our mature technologies: microarrays, quantitative real time PCR, and RNA sequencing. And we also like that this technology is actually continuing to evolve and getting faster and cheaper. So, again, enabling more and more labs to be able to do it. One of the technologies that I'll talk about today is TempO-Seq, which is a targeted sequencing approach. The advantages of this is you can work on RNA, but in cells and culture, you don't need to, you can work directly on cell lysates. Because it's a targeted approach, we don't need to use as many reads in a sequencing reaction in order to cover the transcripts that were interested in, and this also makes the bioinformatics more straightforward and the technologies available today, enable barcoding of samples, therefore we can multiplex hundreds of samples in a single sequencing run.

How it works in a cell culture model would be we would lyse the cells, and the Tempo-Seq libraries have probes for every transcript of interest shown here. There's two probes for every transcript, and they kneel right next to each other here, and they have primers built into them. When these probes kneel, you can then ligate them in a ligating reaction, and that enables you to have a molecule that you can actually amplify to build a sequencing library. So, you're actually sequencing probes. The disadvantages: it's fantastic because it's really fast and it's much more cost effective. The disadvantages are that because it's targeted it only captures a portion of the genome, and it's still expensive so we do need to make some strides towards making this even more cost effective.

We like transcriptomics because there's published standards such as the minimal information about a microarray experiment. The US Food and Drug Administration had all kinds of

excellent consortia like the MAQC and the SEQC providing some ideas of best practices, and publications require making your data openly available and there's repositories like the Gene Expression Omnibus and the Chemical Effects and Biological Systems Database that allow everybody to access the data. More recently there's been strides towards increasing the transparency of omics work, and Dr. Harrill will talk about this in his subsequent presentation, but the OECD's Omics Reporting Framework is enabling sharing of both the data, as well as the analytical approaches to use to increase transparency and reproducibility. And finally, efforts from for example the European Consortium, or European Centre for the Ecotoxicology and Toxicology of Chemicals. This project led by Florian Caiment is developing pipelines like the Regulatory Omics Data Analysis Framework for increasing acceptance of harmonized practices in omics data analysis.

So we've come a long way. And there's great enthusiasm for many applications of toxicogenomics and decision making from mode of action analysis, through tiered testing and prioritization, potency assessment, informing grouping and read across, and deriving even a point of departure for use in risk assessment. Along with this enthusiasm of course are the concerns, some of which were raised at the start, you know, can we trust these new tools has there been enough validation, or what do we understand about sensitivity, specificity, reproducibility, and accuracy? If we switch, would we miss toxicological effects? Have we covered enough biology, for example? The gene expression changes that we are measuring are not adverse phenotypic changes. Therefore, would we be basing decisions on adaptive versus adverse effects? And since gene expression changes, as I stated, is one of the first cellular responses to a toxic chemical, will the doses at which we see these responses be extremely low, and therefore not these doses from not feasible for risk management. And then finally, really practically, how do we do it, and will it give us comparable results to what we're used to seeing? So, I hope in my upcoming slides I'll address some of these questions.

One thing that we've been focusing on, my group, is how to efficiently analyze and interpret toxicogenomic data. We want to do this in a in a manner that's not only efficient but it's also not subjective, so toxicogenomics produces these large data sets of exposed versus control cells, we get hundreds or thousands of genes that are differentially expressed. And over the years, a number of really nice bioinformatics tools and applications have evolved. And so, we can do these complex analyses, but it's been difficult to figure out how best to use that information in risk assessment. So, we've been working toward this more simplified vision for the use of transcript omics in toxicological evaluation and I split it into two boxes. First we take these large gene lists, and what we'd love to do is extract predictive biomarkers and predictive pathways that we can align against adverse outcome pathways in order to determine mode of action and predict potential hazards, so this is box one, and in parallel, we want to do the dose response modeling of the responsive genes, and we want to be able to identify then the dose at which these gene sets are perturbed and use that information in risk assessment. And the next slides I'll go over box one and box two to show you how we're taking steps towards developing these processes.

So, the first is biomarker development. How do we do this? In biomarker development cells or animals are exposed to chemicals that we know target specific toxicological endpoints or modes of action; we call these training set compounds. So, you expose the cells or animals, and do transcriptomic profiling on it. And this data set is subject to machine learning. And through this process of machine learning genes that predict the responses are extracted. In this example we have chemicals on the x axis genes on the y axis and this is a training set of 28 chemicals that were split into two categories: they either cause DNA damage so they're DNA damage inducing, DDI, or non-DNA-damage-inducing and you can see from these red

and blue, red and green patterns, indicating the up regulation and down regulation of gene clear, difference in the, the transcription profile of of the genes between the DDI and the non-DDI agents. So, this is biomarker extraction. And then we then validate the biomarker using chemicals outside that training set. So now we're confirming that the biomarker is working. In this example we have a dose response for two genotoxic agents, showing patterns that match the CDI.

So, this is actually the TGx-DDI biomarker, which is a project ongoing within the HESI Emerging Systems Toxicology and Assessment of Risk Committee, and really excited to say that just two weeks ago, we received notice from the FDA, that our qualification plan for this biomarker was approved, and we're setting off now on our qualification exercise. For biomarker validation for TGx-DDI, just a snapshot of some of the stuff that we've done, we've tested over 100 agents now with the TGx-DDI biomarker. Here I'm showing you two classes of agents: class one is direct acting DDI agents in class four is non-DDI agents. You can see how these perform in a conventional test, the chromosome damage assay and AMES test, and this bar indicates how TGx-DDI classifies these chemicals. So, the, the TGx-DDI biomarker shows high sensitivity and specificity. It's been tested across all kinds of gene expression platforms and performs well in a couple of different cell types.

We want to align these transcriptomic biomarkers against as I said adverse outcome pathways, an adverse outcome pathway is a conceptual framework to describe existing knowledge, linking molecular initiating events or MIEs, through a series of key events to an adverse outcome. It's really an excellent framework to organize information from different experiments and test systems, and it's based on biological plausibility or statistical inference. OECD has an AOPWiki and an AOP program. And this is where you can find many of the adverse outcome pathways that are both finished being developed and endorsed by the OECD as well as in development. At the end what you have, after you've developed an AOP, is something that you can use as a hypothesis for testing chemicals and whether they operate through a particular mode of action. You can also use this for predictive toxicology and for test paradigm development, as well as interpreting new, new approach methodologies, data from new approach methodologies and how it fits into the prediction of adverse outcomes.

The second box was dose response modeling, so here's a dose response, we're familiar with NOEL and LOEL, and of course, with benchmark dose modeling which is what we're using now for transcriptomic benchmark dose model for transcriptomic dose response assessment. So, with benchmark dose modeling we're applying mathematical models to identify a defined response above background and this has been used for decades and toxicology, so it's really nothing new. But what is new is the ability to do this in a highly parallel form across all the genes in the genome. And so for that we're using BMD Express, which was initially developed by Rusty Thomas, when he was up a Hamner Institute, and is now managed by Scott Auerbach at the NIEHS. So, with BMD Express multiple mathematical models are applied to model the dose response curves, so that we can get gene pathway, or biomarker benchmark doses, so I'll flip back and forth between benchmark dose, when I'm talking about *in vivo* data and benchmark concentration or BMC when I'm talking about *in vitro* data. And for this, then we want to think hard about how from this information we're going to derive a transcript on my point of departure.

So, putting one and two together. Over here we have an AOP that we've developed, where we have DNA damage as a molecular initiating event, which if inadequately repaired with lead to strand breaks, and then chromosomal aberrations or mutations. So, we're interested in developing, for example, an integrated test strategy to do high throughput screening of

chemicals to determine if they operate through this pathway. So, we can apply first the TGx-DDI biomarker, in this study we applied the biomarker to TK six cells exposed to 10 different chemicals that were either genotoxic or nongenotoxic. We then derived the gene benchmark concentrations or BMCs as a point of departure for the TGx-DDI biomarker *in vitro* and compared that to the high throughput COMET chip data for DNA damage, actually this was an HepaRG cells, sorry about that.

So, we can directly compare the molecular initiating event and one of the key events in this adverse outcome pathway. And when we do that, we have COMET assay benchmark concentration on the y axis and TGx-DDI benchmark concentration on the x axis, we see an excellent concordance supporting that these chemicals which are actually prototype genotoxic agents that I'm showing here, operate through this adverse outcome pathway, we see these black lines indicate that these assays give benchmark concentrations within tenfold of each other. So, these kinds of integrated test strategies are being integrated into Health Canada's GeneTox21 program led by Dr. Paul White.

We're not alone in being excited, of course, about transcriptomic benchmark dose modeling and deriving points of departure. And these are some of the seminal studies that provide the foundation for why we're doing this kind of work. So, back in 2013 Rusty's really important point paper showing that the lowest transcriptomic BMD from the most sensitive pathway was predictive of the cancer point of departure and also noncancer endpoints. Again, all of these transcriptomic points of departure were within tenfold of the atypical points of departure, across these different tissues in rats and mice.

And it's this paper here more recently, by Gwinn et al. with Mike DiVito and Scott Auerbach and others looking to challenge the system more, so if we only look at transcriptomic points of departure from male rat liver and kidney in five- day studies, considering hepatotoxic chemicals, chemicals that are toxic through to other tissues, as well as negative, you can see that in this plot actually is showing the comparison against any species, both sexes, and any organ, that even just looking at liver and kidney gets you pretty close to the apical point of departure for these kinds of effects.

We’re not the only ones working on transcriptomic biomarkers as well, so this is work by Rooney at all from Chris Corton’s lab, again using the AOP framework. In this case, considering molecular initiating events that are known to be associated with rat liver cancer, and the hypothesis is that transcriptome biomarkers that can measure these molecular initiating events, and the downstream KEs in liver cancer AOPs can be used for the early identification, as well as their associated doses that are likely to be tumorigenic in the liver to your bioassay. And so, he developed a series of biomarkers, and the biomarker accuracy in these experiments was between 91 and 98%. These studies served as one of the starting points and the basis for the Emerging Systems Toxicology and the Assessment of Risks Carcinogenomics Working Group. The objective of this group is to drive the acceptance of a weight of evidence approach using genomic biomarkers for application to *in vivo* shorter-term rat studies to inform liver tumorigenic risk in the rat to your cancer assay.

So, transcriptomic biomarkers that can be used as a rationale basically to potentially waive the need for the two-year cancer bioassay. Another exciting development in this project is the use of air corrected sequencing to measure clonal expansion of cancer driver gene mutations as another biomarker of potential tumorigenic risk so the project leads on this are Dr. Keith Tanis at Merck, and Dr. Chris Corton from the US EPA.

That's my background, and now I'm going to dive into some case studies, reviewing how we're using this information in different toxicological applications. The first case study is on tiered testing for human health risk assessment of hexabromocyclododecane, or HBCD, shown here. The objectives of this case study were to gain experience in applying a tiered testing paradigm to explore consistency across tiers, and to evaluate potential use in risk assessment.

This is our very simplified tiered testing paradigm, tier one is high throughput screening data, tier two is *in vivo* transcriptomic data, and tier three are conventional tests.

These are the methodology for the paper that I'll go through and very high level. Tier one, as I said was high throughput screening, that specifically was ToxCast and Tox21 data, this tier was analyzed by Marjorie Moreau and Andy Nong at Health Canada. And the ToxCast assays for HBCD the summary up here shows they identified the site of toxicity limit, and a number of active assays below that site of toxicity threshold, the varying degrees of activity, they calculated the AC50s of these and they did pathway analyses, and they determined that nuclear receptor activity was a major, was very enriched within these site of toxicities, active ToxCast assays.

In tier two Reza Farmahin from my lab us did RNA sequencing, conventional RNA sequencing, poly-A RNA sequencing on rat livers from male and female Fisher F344 \ rats, these rats were exposed for 28 days to HBCD, and then we subjected them to some of the analyses that I showed on my previous slides, including benchmark dose analysis and pathway analyses.

Tier three was done by Anne Marie Gannon and Ivan Curran at Health Canada. And these were standard rats subchronic studies with all the conventional endpoints built in.

I'm going to jump right into the transcriptomic data. I told you a little bit about very high level about what Marjorie and Andy found with ToxCast and the *in vivo* transcriptomic study first of all confirmed many of these effects that they were occurring *in vivo*. However, it went a little bit further provided us a little bit more information for example we saw some differences between the sexes. These are the number of differentially expressed genes in males versus females so you can see there's this overlap here. When we break that down a little bit further and look at upregulated and downregulated genes, we can see the reason there's more genes in the males is because they have a greater response at the highest dose of the exposure. However, as has been shown by many, many others when we begin to look at the pathway level instead of the individual gene level, we see a much higher concordance between the males and the females, and we see that HBCD causes changes in a variety of pathways, such as xenobiotic metabolism, oxidative stress, hormone imbalance, fibrotic activity, etc.

Both the tier one and tier two assays were explored in terms of determining whether this was HBCD was causing its effects through CAR and PXR so this was a big finding within tier one and so we wanted to look to see if that could be seen within the transcriptomic data as well *in vivo*, and using publicly available biomarkers, here you can see a number of biomarkers on the, on the y axis, we found that at every single dose, we had this is log the minus log p value on the x axis so the larger the bar, the more statistically significant. We found enrichment of a CAR PXR signature in both males and females, supporting what we saw in tier one.

So, putting everything together then here are the tiers on the left, and this is our sort of loose adverse outcome pathway for HBCD. In white, are boxes that we observed both in tier one and tier two, in gray are things that we inferred from tier two and tier three, and in black are things that we measured only in tier three so HBCD appears to activate nuclear receptors in particular

caCARr PXR, leading to the induction of liver enzymes, and this downstream cascading effects. And in fact, the black boxes here, the conventional tests, were predicted from these upstream tears as well.

Moving on then to the BMD analysis, what we are showing in this plot is BMD on the x axis and the number of genes from the transcriptome that have these BMDs. So, we have males on the left and females on the right. I was really amazed when I first saw this plot, about the almost completely perfect overlap of this beautiful bimodal distribution and this is very typical of benchmark dose analyses to see these by bimodal distributions, with some genes, peaking with a low benchmark dose and others peaking with a high benchmark dose. So, here you see a with a benchmark dose of 1000 milligrams per kilogram diet we have about 20 genes in the males, and about 24 or so genes in the females very consistent between males and females.

Reza, you know, one of the big questions is what, how do we derive the transcriptomic points of departure there's many ways to do this and this is one area of active research. Reza came up with a number of different ways that we could identify points of departure. I'm showing our here, the way most typically used is the lowest overall pathway, so this would be a gene set that reflects a molecular pathway or a biological process, and it requires at least 5% of the genes in that gene set to be to fit a benchmark dose analysis, a benchmark dose model, and a minimum of three genes. So, you can see again great consistency between the males and the females. And you also see that this lowest overall pathway is highly consistent between the males and the females as well and is the lowest, and in fact of interest, the lowest atypical effect was on the thyroid gland, and the lowest pathway was also associated with thyroid hormone metabolism.

Andy and Marjorie then then converted, did *in vitro* to *in vivo* extrapolation to derive administered equivalent doses to allow us to directly compare the tiers to each other. And so here you see, each of the AC50s converted to an administered dose as well as the tier two and tier three *in vivo* transcriptomics equivalent doses. And so, you can see for example that tier one as we expected is the most conservative, has the lowest ratio between the lowest active assay, the smallest ratio between the lowest active assay and the estimated human exposure, whereas tiers two and tier three as we expected are really highly overlapping. So, there's great consistency here in the points of departure.

To conclude, then from this, we have hopefully demonstrated to the transcriptomics and high throughput screening predicted similar mechanisms and outcomes. And these were aligned with changes in the conventional toxicology endpoints. The liver transcriptomic profiles were consistent with potential effects in the thyroid, a primary target organ of HBCD exposure. And our tPODs, so transcriptomic points of departure, were consistent with the apical endpoint points of departure. That bioactivity exposure that I showed indicated that tier one was the most conservative and tiers two were highly similar and overlapping. So, the question for you to think about then is, could we have actually skipped tier three, would you have felt comfortable not doing tier three?

Case study two: now we're moving into big groups of chemicals. This case studies focused on per- and polyfluoroalkylated substances or PFAS which are ubiquitous in the environment, I'm sure you've read all about them in the news. Because of their widespread use, their persistence, and high mobility. And there's a really strong body of knowledge on some PFAS like PFOS and PFOA. However, for a very, very, very large number, there's little to no information on them and so here's some of the products that they're found in.

So, our new challenge, then, is how do we begin to address a chemical group that has over 4000 chemicals in it, depending on how you define what a PFAS is? So, we clearly needed to move away from conventional RNA sequencing or microarrays to something with greater efficiency.

And for that, we used the TempO-Seq platform that I already described in detail on one of my earlier slides. So, the objectives of this case study where to use high throughput transcriptomics to acquire hazard mode of action and potency information on PFAS for risk assessment, and specifically to inform acceptable concentrations of PFAS in drinking water, and for cleanup of contaminated sites. We are very excited about this project because it was one of the first times that regulatory partners approached us to address, specifically, a problem that they had. And so, we're working in close collaboration with our regulatory partners, and I would, I would highlight Ivy Moffat and Luigi Larusso here. And this work is done with my partner in crime Ella Atlas, who is the expert on all the *in vitro* models in our group, and on metabolic disorders and other things associated with this project. The work that I'm showing was the work of fantastic people in the lab like Andrea Rowan-Carroll and Anthony Reardon and others as well.

In this example, we are working, now we've moved *in vitro*, so we're using a human liver spheroid model and that's because we focused on liver because hepatocytes hypertrophy was a critical epidemiological endpoint in the risk assessments of PFAS and PFOA. So, we obtain these livers steroids from a company inSphero there. These are metabolically competent groups, spheroids, that are from pools of 10 human donors. You can see a little picture here, they're about 2000 cells apiece.

And so, the challenge, the methods here, were that we expose the liver spheroids to 10 concentrations of PFAS and collected samples at four time points, one, four, 10, and 14 days of exposure, the media were changed every three days and so they were exposed to new chemicals, and site of toxicity was monitored, as we went and over least cytotoxic concentrations removed.

We first started looking at four prototype PFAS shown here. A short chain four carbon PFAS called PFBS, our two most well studied PFAS, PFOA and PFOS, with eight carbons each, and a long chain PFAS, PFDS. So, these were used as the prototypes to begin in establishing our experimental approach, where we then selected six concentrations in two time points to move forward with a much bigger list of PFAS to, to have to cover 23 PFAS in total in this particular experiment.

The first thing that we did was we looked at differentially expressed genes, to evaluate whether there were toxicodynamic similarities between the PFAS. Here, you'll see PFOA, PFBS, PFOA, and PFDS, and we're looking at genes downstream of specific upstream regulators. For example, here, you see PPAR alpha, and it's orange indicating that all of the PFAS activated PPAR alpha. And then you see inhibition here of some of these other upstream regulators .

What's striking to us, is it's great to see as we expected that PPAR alpha is activated across all the PFAS, but we see this very similar, highly similar profile from PFOS and PFBS, and we see that as well I'm sorry about the resolution on this slide, we see that as well over here these two profiles that look highly similar, this is PFOS and PFDS. So, people at PFOA and PFDS look a little bit different. But there's a high degree of toxicodynamic similarities across these compounds. In particular, the long chain sulfinates. The pathways that we saw activated, as I

mentioned were associated with PPAR alpha and down regulation of cholesterol biosynthesis as well as effects occurring potentially on liquid homeostasis pathways.

And so, what I'd like to spend a little bit more time on though is the benchmark concentration analysis, which was a really exciting part of this, of this project. What I'm showing you here are these spots represent all the median gene BMC from all the transcripts that fit models for each of these chemicals, and they're shown from the least potent chemical to the most potent chemical, and these are the confidence intervals. So, PFOS has the lowest BMC the x axis is down here, and it's significantly more potent than PFDS, looking at the median gene BMC, which is more potent than PFOA and PFBS shown here. But this is on day one of the exposure.

And it's really interesting to see what happens over time as we continue the exposure, with the median benchmark concentration coming down to be equipotent for PFDS and PFOS, with PFBS staying up here, and subsequently by the time we get to 14th day of exposure, we see PFOA point of departure coming down here as well or the benchmark concentration coming down here as well. With PFBS kind of wiggling around up here and always being less potent than the other chemicals. So, this is looking at the median gene BMC and these reflect hundreds of genes.

Another way to look at this, is to look at accumulation plots. How do you read an accumulation plot? Accumulation is on the y axis and then the BMC is on the x axis, and each spot represents a gene shown from the lowest BMC the gene with the lowest BMC to the gene with the highest BMC. So, for example this purple line is PFOS, and we can see there were about 220 genes fitting models. And these, you can see the distribution of where their BMCs were. So, an interesting thing to look at here is then how left shifted are these curves, you can see why PFOS's BMC was lower than the rest and the left shiftedness would reflect increasing potency. And so again, we can look at these curves over time and how they change.

So, median gene BMC is one way, looking at the curve shift. Another thing that we really liked was a paper coming out of Steve Ferguson's lab shown here, where Steve's group looked at, if they did BMC modeling, could they differentiate from between liver toxic and non-liver toxic drugs in HepaRG cells, and what they discovered was that if they set a threshold of requiring at least 105 genes fitting BMC models, then they could differentiate those chemicals that were liver toxic from those that were non-liver toxic. So, this is a different cell model, but we thought that was kind of a neat idea, so here I've drawn a line, and these are horizontal it's an optical illusion that they don't look horizontal across the 100 gene line, and we can again rank potency and this way we're resetting sort of a more meaningful threshold, must have a certain minimum level of biological activity, and we rank the potency at that, at that stage, but we can also move this bar up and down. And, you know, set it wherever we want. But when we do that, we see very little change in the way we've ranked the potency is between these different PFAS. So, this is again about potency ranking but if we want to do something like this as a point of departure, we'd need something more conservative than a medium gene BMC, or, or, you know, the 100th gene, we'd want to go much more conservative.

And so, discussions with our regulatory partners on this project led us to, to think about two points of departure, the first being the lowest pathway, because as I mentioned, this is one that's widely used by many people. And the second being the fifth percentile gene to align against how that's typically, this is typically an endpoint use the fifth percentile AC50 in analysis of ToxCast data. And so, Andy, again, did the IVIVE conversions to administer an equivalent dose, so that we could compare our *in vitro* to our *in vivo*. So, these green bars on

this plot represent the points of departure of conventional animal tests used in the risk assessments. This box represents the median gene BMCs and their interquartile ranges, and here's where our transcriptomic points of departure fall, and you can see that even though we're in a human liver spheroid model in vitro, we're actually when you convert that almost completely aligned with the *in vivo* point of departure. That was for PFOS. Here's for PFOA where you see this, the declining again median gene BMCs, we see a little bit more stability when we're just looking at the, the lowest pathway in the fifth percentile gene. And again, we're very close to the conventional test endpoint.

Comparing that then to the human exposure, which is shown in the dashed red line here, and looking at this paper by Perez et al. in 2013, where they saw about 0.01 micro molar of PFOA in human livers and 0.08 micro molar PFOS, our lowest pathways were about an order of magnitude away from that so 0,6 and 0.8, suggesting the changes in gene expression actually would be expected at current population exposure levels.

And then of course you're interested in all the rest, and I don't have time to get into all the details of this but one of the things we did was we start separated the carboxylates from the sulfinates, and here I'm ranking them instead of in potency order in order from the shortest chain length to the longest chain length, and as expected, when it came to benchmark concentrations and points of departure, we saw this correlation between chain length and increasing potency for both the carboxylate and the sulfinates, and we could observe that as well in the gene accumulation plot so our regulatory partners now are using this information to help them in grouping these chemicals for Read Across and informing some of the data poor chemicals.

To wrap up that case study then, these human liver spheroids exhibited robust transcriptional responses and what I didn't show were some phenotypic changes following POAS exposure. And so, there is toxicity from these PFAS to human liver spheroids. Some of these PFAS demonstrating increasing potency over time, this was, these were nominal concentrations,

so we're doing some chemistry on this now, but indicates the importance of considering that toxicokinetics of the chemical and the importance of time series analysis. PFAS toxicity, as expected, increased with chain length. But we hope overall that you can see the utility of the transcriptomic BMC modeling for the potency comparisons. So, there's all kinds of additional work that we're doing on this now, including mode of action analysis, chemical analyses, thinking about phenotypic profiling and predictions of health effects in humans, and we're now looking at PFAS in mixtures.

And then just really quickly, our final Case Study and this one is unpublished, I stole the slides from a fantastic PhD student named Geronimo Parodi-Matteo, this one again in collaboration with Ella Atlas and our regulatory partner is Tara Barton-Maclaren and her group from the Existing Substances Risk Assessment Bureau at Health Canada and working in collaboration with awesome biomathematicians Matt Meier and Andrew William. In this example, we're studying bisphenol A replacements. And so, we're working in MCF7 cells which express the estrogen receptor and it's a little bit more hypothesis based. So MCF7 cells are exposed to nine concentrations of sixteen different bisphenols as well as BPA. And we are applying TempO-Seq and our standard pipelines to look at transcriptomic benchmark concentrations and similarities and profiles.

The first thing that we did of course was we went to existing biomarkers and Chris Corton from the EPA has an actually an excellent ER alpha biomarker that's highly predictive. ER activation biomarker in MCF7 cells, and so we applied that here in this column to determine which of the bisphenol A replacements activated this biomarker, you can see that was done here. And then we derive benchmark concentrations and points of departure for all of these replacement chemicals, this big red box indicates that we could identify some chemicals that were inactive in MCF7 cells. They didn't perturb enough of the transcriptome to allow us to identify point of departure. So, people often ask well what about negative chemicals? Here I'm suggesting that we can identify inactive chemicals.

So, I talked about the 100th gene in this case we wanted something a bit more conservative, so we went with the 25th gene, and we're comparing that to the lowest median pathway and to ER-specific points of departure as well. And that's shown over here ranked from the most potent BPA to the least potent and the inactive chemicals. So, you see, the, the sensitivity of choosing these kind of pathway- or hazard-agnostic points of departure, showing tons of activity below where estrogen receptor appears to be activated and really interestingly, the beautiful concordance between the 25th gene BMC and the lowest pathway BMC.

We also did gene set enrichment analysis and to look at toxicodynamic similarities. And here, just like in the PFAS case study that I've shown you, you can look across these chemicals at perturbed upstream regulators or canonical pathways, and you can see a high degree of concordance in terms of activation and inhibition of these pathways and Gene Sets providing a strong rationale for putting these together in groups for analysis.

Alright so that brings me to the end of my presentation, the overall conclusions or the lessons learned. I hope you can see that we are building momentum here in demonstrating that we can use this kind of information to identify specific, this kind of information being gene expression profiling, to identify specific hazards and to define modes of action, and how adverse outcome pathways are helping in the interpretation of these transcriptome analyses. I hope that I've shown you how BMD modeling can be used to compare potency of chemicals, as well as made you think about how we might use this to derive a transcriptomic point of departure. And I hope that I've convinced you that the doses that which we're seeing transcriptional changes are actually consistent with those causing adverse pathologies and not, you know, not a lot more sensitive.

The case studies are really helping us in my opinion to build confidence in the application of genomics in regulatory evaluations and really help to define context of use and we're really appreciating the close collaboration that we are having with our partners who actually are conducting the risk assessments and helping us to identify their challenges and helping us to determine what's feasible. So, we've seen we're seeing growing use in our applications across different regulatory agencies, and we think these case studies are required to continue, especially focusing on demonstrating applicability across broad chemistries in areas of biology, that will be really critical to mainstream integration.

So, we need to continue to develop more biomarkers, we need to continue with our validation exercises and refining of course and validating our approaches. There's a lot of emphasis from our partners in establishing best or what we might call acceptable practices including moving towards international harmonization of some of these approaches, developing potentially guidelines in the future. And also, we believe that we should begin to, we should also continue to do these case studies to build confidence.

Just like to finish by acknowledging all the important people I hope I mentioned most, most of them as I went, but also a shoutout, not only to my Health Canada partners, but to people around the world who have been exceptional collaborators, especially Scott and Steve at the NIEHS and Rusty, Josh, Logan, and Chris at the EPA and my HESI partner the executive director Syril Pettit, the TGx-DDI Working Group and the Transcriptomic Point of Departure Working Group, and thank you for the funding to the funding agencies who have supported this work over the years and to you for your attention.

Thank you.

Stephen Edwards: Thank you so much, Carole, it's just amazing the amount of work that you and your collaborators have done in the past few years, since we got a late start due to out technical difficulties, we're going to hold all questions until the panel discussion, but I do encourage everyone, you should have a Q&A icon at the bottom of your zoom screen. So, please click on that icon and submit those, those questions; clarifying questions, Carole may be able to deal with during the other talks and then we'll have our more detailed discussion at the end.

So, with that, I would like to introduce our next speaker, Dr. Josh Harrill, who will talk about high throughput transcriptomics for chemical bioactivity screening and tiered hazard evaluation. Thank you, Josh.

High-Throughput Transcriptomics for Chemical Bioactivity Screening and Tiered Hazard Evaluation
Joshua A. Harrill, US EPA, Research Triangle Park, NC

Yes, thank you for the introduction, Steve. Can everyone hear me okay?

Stephen Edwards: Yes, we can.

Joshua A. Harrill: Okay, and the slides are displayed properly?

Stephen Edwards: Correct.

Joshua A. Harrill: Yeah. Well, thank you, thank you for the coordinators for inviting me to come in and speak in the session day, this is some, some very exciting topics about some very exciting science, and I'm here today to share what we've been doing at the US EPA in terms of high throughput transcriptomic screening for chemical bioactivity, and how that fits into tiered hazard evaluation. A conflict of interest statement: the views expressed in this presentation my own, they are not the US Environmental Protection Agency's, nor do mention of trade names or products represent endorsement for use and I declare no conflict of interest, regarding the concepts in this talk.

So, as far as what I'm going to be talking about today, my objectives are to provide a broad overview of next generation blueprint of computational toxicology at the US EPA, with a particular emphasis on the role of transcriptomics in that strategy. I'm going to touch on a variety of different topics related high throughput transcriptomics in chemical screening, including a targeted RNA-Seq technology, some novel bioinformatics workflows and open-source tools that we've developed or adapted for analyzing these data. I'll touch base on

transcriptomics reference materials and how we're using those to inform our science, and also at the end, speak more about this international effort to develop an omics reporting framework.

So, I'll start here with this diagram over on the left side of the slide, this is taken from what we call the Next Generation Blueprint of Computational Toxicology at the US EPA. And what this diagram shows is a variety of different research topic areas that are present in the CCT's portfolio. We have different research efforts, focused on chemical characterization, toxicogenomics and *in vitro* disposition, exposure assessments, and hazard evaluation which, which is where a majority of my work comes into play. We pursue each of these topic areas to the combination of computational modeling and high throughput screening and have several other areas of focus related to outreach and training, establishing competency models and characterizing the uncertainty and variability in these tools, and also developing software and IT tools within deliver information to the public and the nation state. And so, I'll start here.

Stephen Edwards: Josh, your, your audio is kind of fading in and out. I don't know if it's possible to get closer to the mic or anything.

Joshua A. Harrill: Give me, give me one second here. Um, I've had troubles with my microphone in the past. Bear with me, one second. is that any better, Steve?

Stephen Edwards: I think so; it's kind of fading in and out so I'll let you know if it starts fading again.

Joshua A. Harrill: That might be the best that I can do, I apologize. I'm sorry. So, anyway we'll focus on this, this area over here of hazard evaluation, which is where majority of my work takes place. So, traditionally when you think of hazard evaluation and CCT, what you think of is ToxCast and the ToxCast program uses targeted high throughput screening assays, to expose living cells or isolated proteins chemicals and assess the bioactivity and potential toxic effects of those chemicals. So, this diagram from Richard et al. in 2016, kind of summarizes the ToxCast program in total. And basically, what it was is, is there were several different phases of ToxCast testing, they used to use mostly targeted assays that measured the effects of chemical X on target or biological process Y using these more traditional high throughput screening assays. And while the success approach provided a rich collection of biological data from people can infer the bioactivity of chemicals, it also provided an incomplete coverage of human biological space.

So, our new strategy for hazard evaluation is to improve the efficiency and increase the biological coverage of our screening efforts by using what we call non targeted profiling assays. And these assays are designed to cast the broadest net possible for capturing and potential molecular and phenotypic responses in human cells to chemical exposures.

This is another diagram for the next generation blueprint of computational toxicology at EPA just describing where these types of the profiling assays fit in a, in a NAMs-based tiered hazard evaluation approach. And so, the high throughput profiling assays exists within tier one of the strategy, which I'm showing in the red circle. Basically, it's our, it's our first pass at actually testing chemicals and spending some time on the biological activity of those chemicals. So, in order to be a high throughput profiling assay, we consider those to you must meet several criteria for those. One is that the technology must be able to yield bioactivity profiles that can be used for potency estimation, mechanistic prediction, and evaluations of chemical similarity, that the assay be compatible with multiple human derived culture models, that are capable of being run in concentration response screening load, and it helps to they are

cost effective. And to date, EPA has identified and implemented to such scientific profiling assays. The one I'm talking about here today, high throughput transcriptomics, and the high throughput phenotypic profile in the cell painting assay, which I won't talk about today.

As far as how this tier one high throughput profiling data is used. Really this diagram here is focused on point of departure estimation. And so, you can follow one of two different paths. If the information from the tier one assay, you don't have any previous information on the biological target that a chemical might hit, then you can actually take this tier one data and derive a molecular point of departure, based on the perturbation of any particular gene or pathway or phenotype, that the chemical maximally perturb, and you follow this right hand path to define your point of departure. If you do know something about the molecular target that chemical effects, or you can infer that information from the tier one data, then you can go down the left side of this diagram, and actually confirm your biological activity in orthogonal confirmation assays, and then start to feed that information into either existing or de novo AOP constructs as Carole described to come up with more mechanistic or adverse outcome based points of departure.

As far as the technology that we're using in our high throughput transcriptomic screening program it is the TempO-seq technology that Carole described in her presentation. I won't dive into the assay chemistry because she already did explain that in sufficient detail, but if anybody wants to know more about this this targeted based transcriptic assay, I refer you to this manuscript by Yeakley et al. in 2017.

So, the version of the TempO-seq assay that we're using is a whole transcriptome assay, it's capable of measuring the expression of greater than 20,000 transcripts in every sample. It requires only picogram amounts of pulling on a per sample and important for our purposes for high throughput screening, it's compatible with both purified RNA samples or cell lysates, and so cell lysates is the predominant sample type that we operate. And as explained earlier these lysates are barcoded according to sample identity, such as the position on a multi well plate, and combined in the library through sequencing using standard instrumentation, such as a heist. I's more efficient to actually sequence the ligation products that are part of the TempO-seq assay, as compared to traditional RNA-seq, because really at all your sequencing are these 50 base pair reads from the ligated probes.

So, we've, we've used this technology to actually build a high throughput transcriptomics platform for screening and environmental chemicals and we published that paper in 2021 in Toxicological Sciences, which you can see here. And really what I'm going to be showing throughout the rest of the talk are results from this particular paper, which we call our paper, in which we screened 44 toxic chemicals an eight-point concentration response in the MCF7 cell model. And also, in subsequent studies, really expanded the number of chemicals that we were testing in the, in the scope of our screening efforts in what we call our NCF7 screen and tested the more than 1700 different chemicals, using this eight-point concentration response screening paradigm. And the two-year assay technologies that we used to this, one cell transcriptome Tempo-Seq, and then an imaging-based cell liability assay to give us an idea of what concentrations may be acutely cytotoxic in our test system.

So, what I'm showing here is the experimental design we used in both the MCF7 palate and the MCF7 screen. So, what's shown over here on the right side is a plate map of our best place. And once you can see is that within these those plates, we have 44 chemicals are ready to read a point solution series that you different shades of green in the top and bottom parts of

this plate. We also have a number of vehicle control wells which is the brown DMSO being a vehicle control, and a trio of non-treatment controls within this plate as well.

In addition, on every dose plate that we use in the studies we include triplicates of three different reference chemicals, so genistein, which is the estrogen receptor agonist at 10 micromolar, sirolimus which is the agonist at 0.1 micro molar, and then trichostatin A which is an H deck inhibitor at one micromolar. And these reference treatments are part of our experimental design, because we use them to track assay performance of the TempO-Seq assay, including the response of the MCF7 cell that we have in the plate, so we want to make sure that our cells that we've actually used in the study are responding in the proper way to these different mechanistic reference chemicals.

So, this is our dose plate over on the left. And what we do is we use a specific type of technology called an acoustic dispenser to actually randomize all the different treatments across our assay, which is shown over here on the right side. So, our assay plate columns two through 24 container the MCF7 cells and our first column is actually empty and reserved for what we call reference samples. And so, these are another type of sample that we use to track assay performance in this high throughput screening paradigm, but these, these particular samples are designed to be independent of chemical treatments and the responsivity of the culture, so they are basically commercially available standards that you dispense into each and every plate of different types, we have reference RNAs in blue and reference cell lysates in yellow, and buffer blanks in purple. And so, this allows us to actually evaluate the performance of the TempO-seq assay independent of the performance of our actual culture. So, so two very different purposes for the reference treatments and reference them. And then the bottom, the bottom wells here on this plate are actually reserved for our sequencing vendor, which is by HubSpot.

So, kind of diverging away from our screening, kind of screening results for a moment I wanted to touch base a little bit more on the actual use the development. So, reference samples, again, to provide objective evaluation of the technical performance of the omics assay, not the biological response of an *in vitro* test system. We use reference treatments. These reference samples are processed in parallel with your test samples. They should be subject to the same manipulations and assay conditions as the actual test samples themselves, and they pass through all the same processes in terms of generating the transcriptomic libraries that you sequencing, and they should be implemented in a manner that facilitates monitoring of the consistency of your omics assay and results generated both within studies, across studies and potentially across laboratories over time. If there are there are other laboratories that are implementing similar things.

So, what we did in what I'm calling the early days of high throughput transcriptomics and you can see the date was really only a few years ago, 2017 to 2020, is that we identified two different pairs of reference samples for incorporation into our study. The first reference pair number one is a purified RNA pair, these are commercially available from Takara, they are a universal human reference on a, which is a, an RNA type that's derived from a combination of different tissue types within the human body, and also a Takara human brain reference RNA. So, an RNA type that's derived from different brain regions in human tissue. And these two different sample types have vastly different gene expression profiles, and we can leverage that, they'll help us track assay. Now, these Takara samples very comparable to the samples that were generated as part of the Microarray Quality Control Consortium. But the reason we went with the Takara products is instead of the ones that were actually from MAQC is that the

ones used in MAQC were actually [inaudible]. Not really optimal for evaluating performance of a transcriptomics assay that was performed in this cell lysate format that we were using.

In addition to be purified RNA pairs. We also developed a pair of bulk lysates here at EPA labs, which are basically MCF7 cells cultured in mass treated with either our vehicle control DMSO or TSA, lysed [inaudible] of the exact same sample pool in every single plate.

So, these are what we use early on in our transcriptomic screening efforts. But after we used them for a while we actually did receive a need to develop a more replenishable human derived transcriptomic reference samples that can be compatible with multiple assay technologies and available to consumers with both purified RNA and so cell lysate forms.

So, that's what we did, we partnered with BioSpyder it's actually engineer seven transcriptomic reference samples. So, these paired reference samples were prepared by combining the genetic material from different human derived cell lines focusing on the different condition. These different pairs were formulated to mimic the performance characteristics of the MAQC samples and to cheer for us is that we were able to prepare them as the purified RNAs and cell lysates, so the lysates were prepared first and then we actually did a mass RNA extraction on those same lysates to prepare the RNA samples.

And what we, what we can see when we actually look at the number of genes detected in the reference samples and the number of genes in common between like pairs reference sample A and B is that we did get similar numbers and protective genes in our engineered samples, as compared to both the MAQC and the Takara samples, and that the expression profiles for the lysate, shown here on the right side of the slide on the x axis, fold change versus the fold change in our purified RNAs was highly correlated, so we're actually seeing the correlation between the purified product in the lysate product, and a very large range and full change values, when you actually compare a sample A to a sample B. And this this range of full change values, but we actually use to evaluate assay.

So, I'll show an example of that on this slide, shown over here on the left-hand side, this is a heat map of the correlation of log2 fold change files from all the different plates that are present in a string. So, on both the x and the y axis of this chart, are dozens of plates of high throughput transcriptomics but we installed these [inaudible] each and every plate and then measured the correlation of the log2 fold changes across each plate.

And so, what you can see this visually inspecting this type of plot is that colors a little more of this yellow and white are very highly correlated and colors near this orange color are not correlated at all. And so, what you can see is that when we compare our human and universal human reference or a brain reference RNA, the correlation across plate are very, very high, except for a couple instances you can visually see which show up as [inaudible]. And so just doing the simple visualization of these log2 fold change correlations, allow us to flag plates, which may not be performing up to spec for what we expect the with our high throughput screening. We can take that information and actually plot it in a slightly different way over here on the right hand side. This is a visualization of a statistic called a D-statistic, which is mean correlation fold change to like sample players across all the plates in the study. And so, you can see the plates maybe underperforming will have lowered D-statistics in the lower ends of these distributions versus the majority of the plates up here in the studies, so this is just another way for us to use the samples to flag plates within our screens that may be underperforming and may need additional scrutiny in terms of determining whether they should actually be included in our study.

Now, switching gears from reference samples, I'll talk a little bit about the high throughput transcriptomic bioinformatics workflows that we built the handles. So, this is a diagram from the Harrill et al. paper that we published. And then what I'm showing up here is basically two levels of data management and data analysis. So, we have a raw data processing, is that we go through, we actually take the raw FASTQ reads from our high throughput transcriptomics and align them to a probe manifest, using the HISAT2 open-source tool and generate count matrix. Basically, count matrices, which are samples and columns and rows. And then from each of those. Each of these columns are great sample, we can calculate a number of different sample level quality control criteria as well, to help us read out underperforming samples, and I'll talk about that in the next slide.

All this information from the alignment strikes and fed into a Mongo database, and then we drawing, we for each individual chemical in our study, we actually draw the samples that are associated with different doses of those chemical and their corresponding vehicle controls and run a couple of type of concentration response analyses pipelines, in order to derive molecular point of departures through these chemicals, as described in the previous talk, so really our focus here is to make this entire pipeline reproducible and open source, all these tools are available in GitHub, to automate and efficiently execute these computationally intensive steps. And at the bottom half down here to focus on concentration response modeling and molecular point of departure.

So, in terms of our sample level QC metrics. These are also from the Harrill et al. paper in 2021. We basically came up with a collection of five different, five different sample level quality control metrics, to help us identify underperforming samples. The metric up here on top is related to our cell viability parallel track assay. So, that's the first thing we looked at and we reject any sample of our study where we're seeing a greater than 50% decrease in the viability of cells. So, that's a that's a pretty extreme threshold, but we do reject anything we see very cytotoxicity.

We then looked at both the number of mapped reads in every sample and the fraction of uniquely mapped reads in every sample and reject those that either have a very low total mapped read count, that's less than 10% of our target depth for a sample. So, for instance, typically will sync with a depth of 3 million, and we will reject any samples that come in below 300,000, so 10%, and more importantly we also tend to reject samples, where we see very poor mapping rates, so where the actual counts that we're seeing in the fast data outputs, do not match to any of the probes that are in the probe manifest that by department. And so, this this, you see a sample with low mapping rate, it usually means that samples poor quality.

So, in addition to those we've come up with a trio of different statistics based on the count distributions that are present in our samples. So, the first one here is called Ncov5. It's the number of probes that respond uniquely match read in each one of our samples. So basically, that you can think of this is basically a metric of the number of genes that are detected, and Nsig80 is the number of probes that are capturing the top 80% of signal in the sample, so you can consider this a metric our most highly expressive genes. And then we have this third one called the Gi coefficient, which is adapted from the field of economics an actually was designed to look at income inequality distributions in human populations, but we can apply it in the same way to look at inequalities in read count distribution across probes. And so, we use a principal notice the Tukey's Outer Fence principle to actually set up accept reject criteria based on the distribution of properties and accounts that we see in every single sample.

Just to visualize kind of what that looks like. And I apologize to the rendering quality of this figure. But anyway, the panels A through E represent each of our quality control criteria in terms of number of mapped in A, fraction of mapped reads in B, and then NCov5, Nsig80, and GiC in C, D, and E. And we plotted the distribution and the points of each of the sample types that are present in their study, their reference chemical, a test sample are being the so controls, so on so forth. And, for instance, any of the samples that actually fall below the horizontal lines on any of these plots on the case, above would be ones we would flag for some type of abnormal mapping rate or abnormal distributional property within the actual study and we would actually gate those out from our final analysis.

So, in total in this MCF7 screen, where the large screen where we tested more than 700 chemicals that involved more than 32,000 TempO-seq samples. We see that none of our lysis buffer blanks, which are just lysis buffer containing genetic material actually passed any of the QC criteria, so that's good, means their lysis wasn't contaminated in any way. And then a greater than 99% of the test samples within this 32,000 were acceptable based on these quality control criteria, and that's the plot down here in panel F. So, we retained bigger than 99% of each of the samples into the sample types that would present in our study. And so that's what we consider a fairly good pass rate for the quality of the data that's actually coming out in these large-scale TempO-seq studies.

Now focusing a little bit on the bottom half of our workflow. This is a portion of our informatics pipeline that's focused on the signature, it's focused on concentration response. So, we have built two different pipelines for actually accomplishing this. One of them is the is using the BMD Express pipeline that Carole mentioned in her talk. And using the signature aggregation approach [inaudible]. But what I'm going to talk about today is our, our this alternative pipeline here, which uses a tool called DEseq2 in a novel method that we've developed called signature concentration response modeling to actually also come up with the signature level points of departure.

And so, there's the signature level concentration response modeling approach actually brings together a few different open source tools that are present in the literature, both this DEseq2 tool which is a, a tool that was developed for processing RNA-seq data and generating moderated log2 fold changes for RNA-seq data, single sample gene set original analysis which is a, a specific type of analysis that provides a signature score, which I'll get a little bit more detail what this is here in a moment. And then tcplfit2, to which is a concentration response modeling package that we've developed here at EPA. So, taking together all these methods are intended to address coordinated changes in gene level expression in genes belonging to the same genes set or signaling factor.

So, a little bit more in our signature scoring approach. This basically shows how we take, we basically draw our, our chemical information from our queriable Mongo database. And so, for every individual chemical data, we'll take the normalized count data from chemical x, all those in chemical x, which are counted in the red triangle, its corresponding vehicle controls, estimate fold changes you using this DEseq2. We then bring into a put into play, a catalog of genes set signatures, with toxicological relevance, some of which are annotated molecular apartments. And we draw this from several different sources that are publicly available in terms of public gene set collections, including Bioplanet, the Connectivity Map database, a resource called DisGeNET, and the Molecular Signatures database.

So, so, we've compiled a signature catalogue of about 11,000 signatures, then we can then use the input in combination with the DEseq2 model and fold changes for computation of

signature stores, using this single sample GSCA approach. And so, again, what the single sample gene set enrichment analysis approach does is we actually store coordinated responses at each test chemical concentration, and what it's designed to do is to take, to take the information that you may see in coordinated, but not very large magnitude, gene expression changes, and basically convert them into an interpretable result is a signature score.

So, why are we looking focusing more on signature scores as compared to fold change? We can actually demonstrate that using our three different reference chemicals that we are expanding. So, what I'm showing here, I've got three panels genistein, which is our estrogen receptor agonist, sirolimus which is our [inaudible] agonist, and trichostatin A, so we scored these as either basically a week, medium, or strong stimulus that disrupt, that perturb gene expression of MCF7. And what you're looking at in terms of these histograms, is the correlation of either log2 fold changes across like samples that are on our different assay plates, which is the red distribution, versus the correlation across assay plates of signature scores, from the GSEA scoring algorithm.

And what you can see is that when we actually compare the correlation of log2 fold change versus the correlation signature stores, the signature stores consistently have a higher reproducibility or higher correlation, then the fold changes, especially for chemicals and calls weak transcriptomic responses. And so that's one of the reasons I would look into the single sample GSCA is kind of our bread and butter.

We can also share this signature scoring method actually identifies what we what we identify the expected biology for these chemicals. So, what we're showing here on this plot is our different molecular targets your estrogen receptor versus our different reference chemicals shown them the different colors. What you can see is, for instance, genistein which know the estrogen receptor agonist causes a higher enrichment in the signature score for estrogen receptor pathways versus pathways that are associated with the different molecular target that that genistein seem should not affect. And then you see that same, same pattern with sirolimus. We see an enrichment of the interplay three pathways and sirolimus. And then trichostatin A we see the same thing, so based on these results you can see we actually identified the expected biology, with the signature scoring approach, which gave us more confidence in the biology that we actually see what this would be associated with reasonable mechanism.

So, this slide demonstrates how we then take the SSGA results and actually calculate our BMCs or our molecular points of departure. So, this is an example for the chemical lovastatin queried against a hallmark pathway for cholesterol homeostasis, which is the, the biological activity this drug should affect. What you can see it as a function of dose, your signature scores actually go up. And we're able to calculate a benchmark concentration, based on the point at which the signature stores pass outside of the [inaudible] calculate. And so, if you take this information and look over on the right-hand side. This is an accumulation plot of signature scores, similar to the ones in saw in the previous talk, but you actually can see that each point is a signature benchmark dose where the confidence downwind. And you can actually rank those based on the from most potent at least potent and come up with a molecular point of departure based either on the most sensitive signature, or perhaps a statistic based on the distribution of [inaudible].

You can also take this information and collapse it across doses. To calculate this value which I call it SSAUC. So, this is basically a sign that scale the area under the curve metric that we

found the fairly useful for determining the mechanisms that chemical maybe for. So, mechanism over here on the left, concentration responsive over here on the right.

So, another example of how we put the signature scores in the play can be illustrated with these different estrogen receptor agonists and antagonists in this particular spot so what we're showing here are gene level beta for 4 estrogen receptor agonists, 3 estrogen receptor antagonists, and displaying the genes that are part as a full restraint signature, which are former strength being a strong estrogen receptor. So, basically what you can see in the data, why I wanted to show this, is it basically demonstrates that the gene level data [inaudible] is fairly noisy, but when we actually aggregate all this information at the signature level, we actually get the correct directionality of signature enrichment based on the estrogen receptor signature that we see here. So, basically expressing the full-strength signature down genes go down following antagonist treatment, and the strand signature up down genes go up following ER agonist treatment. And so, using the signature scoring approach and see not only is the correct biology identified as being effective, but the directionality of the response is consistent and expected as it should be.

This is a summary result from our large HTTr screening campaign, which included more than 1700 chemicals. And what we're showing here is a illustration of that, that bimodal phenomenon that Carole pointed out in her talk. And so, plot over here on the left, we actually take all the active signatures for each chemical, [inaudible] and the size of the bubble corresponds to the number of active signatures. And when you actually compare the fifth percentile of BMC the active signatures versus the median BMC the active signatures, you can see the population of red points, pop out, which are left shifted into the upper left-hand side of this plot. And basically, what that translates into our chemicals where you see this urban peak biological activity in the red histogram in the middle of this plot, versus a large peak of biological activity of higher concentration.

And what we saw we actually looked at this anecdotally, is that chemicals that display this type of response in MCF7 cells tend to be chemicals with non-pharmacological targets in this particular cell models so estrogens, retinoid acid receptors, of poly aromatic hydrocarbons, things like that, that actually cause this bimodal signal, where the contents of the signatures that are in the lower mode tend to be associated with molecular targets [inaudible]. In contrast, you can see another population of very chemicals in blue, where you don't really see this, this bimodality you see some bimodality really the bulk of the biological activity is left shifting. And those are those are just very potent chemicals that are causing all kinds of things to happen to the biology in the cell at the at the threshold for the bioactivity. So, these tended to include potent toxicants such as organic metallics dies, and other things that we know are fairly nasty.

We can take that information from our signature concentration response modeling and use that sign scaled area under the curve metric actually do clustering analysis as well. So, what this heat map shows is who's in the rows, signatures in the columns, and then the different styles are the sign scaled AUC in the value and here we're showing all the chemicals in a large screen that had a potency less than one micromolar. We actually clustered the chemicals based on these signatures, but you can see is chemicals that we know are estrogenic or glucocorticoid receptor agonists or have some type of known mechanism tend to cluster together. So, this is basically demonstrating that our signature concentration response modeling approach can help identified chemicals that are acting in similar or in the same mechanism.

We can visualize that data in a slightly different way as well. These are new map visualizations, we actually took the, the sign scaled area under the curve metrics and ran this new map of algorithms actually to visualize those data associations between chemicals in a, in a kind of a different visualizations case. And so, what you can see on this type of plot is that chemicals that the act more similarly to each other, tend to be clustered closer together. So, what you can see in part of this coordinate space of these little cluster of the chemicals that when you actually look down into the gene level responses and the signature level responses in these tend to be causing similar biological effects.

In this particular visualization, you see that the points on the, on the left-hand plot or colored by the potency in the fifth percentile BMC which is what we're changing the molecular point of departure, and you don't really see a large grading of effects from the lower right to the upper right that separates these, what really separates these two point clouds are the number of signatures that are actually in effect, which you can see on the right hand side of the plot line from purple to white.

And then you have this interesting cluster of points that tend to be way over here on the left side of this plot, so we're really interested to see, you know what these particular points were. And we investigated further, but actually overlaying information from an orthogonal assay on top of this heat map projection. So, what we're showing you over here on the left side of this plot our predictions from the ToxCast ER high throughput screening prediction model where chemicals that are predicted to be estrogen receptor arguments are shown in red, antagonists are shown green, and then the inactives are shown in blue. What you can see them indeed over here this this left shifting plot of points to be all estrogen receptor agonists that will be acting from the same molecular events in our, in our transcriptomics data.

And then we can also be the same type of prediction. And I must apologize there because actually got these two slides backwards the HTS prediction models actually on the left, and again you can see the red points are estrogen receptor agonists that are all clustered together here. And we can do the same on the left-hand side of the slide, actually using the signatures that are present in our signature catalog to do the same types of predictions.

So, in the interest of time, I'm going to move a little quicker through these last few slides, which really focused on the OECD Omics Reporting Framework. And so, this is we're setting aside the screening data for a moment. We'll talk more about reporting and reproducibility. So, this is a project that was conceptualized in 2018 through the OECD whose aim was to develop frameworks to the standardization of reporting omics data, generation, and analysis and to ensure that all the information required to understand and interpret and reproduce them this experiment and results are available in some type of reporting framework.

And so, the purpose of the project was to ensure, to basically develop a reporting framework for scientists to report how they've actually conducted an omics-based toxicity study and not to stipulate the methods of data analysis for the interpretation. That should be applied so we're not trying to define best practice here. We're rather trying to provide guidance on reporting of information and to foster transparency reproducibility.

And so, what this reporting framework actually looks like it has a modular structure, where each, each box in this chart represents a different reporting module. So, we have a study summary reporting module, a reporting module that describes the toxicology experiment itself [inaudible], and then different data acquisition and reporting modules for various

transcriptomics and metabolomics technology platforms, for example microarray, RNA-seq, mass spectrometry etc.

And then, downstream from that we have what are known as data analysis reporting modules. And these modules are different ways that people have actually looked at omics data in different contexts, so looking at differential molecules and differential expression, doing benchmark dose modeling, multivariate analysis and so forth. That actually provides reporting fields to state what they've done. As far as the computational steps needed to convert data that were required from an omics technology to an interpretable outcome.

And so we've harmonized this project was originally conceived as a transcriptomics reporting framework and metabolomics reporting framework, we've since harmonized those where we could, made these different modules that are shown in orange, amenable to any type of omics technology. So, transcriptomics, metabolomics, or hopefully soon proteomics. We've develop these reporting templates that provide reporting fields where somebody should report their data and develop the narrative guidance to descriptions and what [inaudible].

And then we're trialing these, these three points through a series of paired data submitter injuries and trials where the data provider selects a study, analyzes that study based on their, their, their practice they use, reports all the details into the reporting framework, and pass that on to the end user [inaudible]. And we've seen good concordance between data provider important data provider [inaudible]. See more details about this, they can look to this manuscript in Regulatory Toxicology and Pharmacology in 2020.

So, I think I've run out of time, so I'll just summarize what we talked through today. I've covered the innovations in high throughput transcriptomic screening, including open source tools and different experimental design considerations. Touched on the development of transcriptomics reference materials and the reporting framework we've been developing for toxicology.

And I'll just thank everybody, all the many, many people have been involved in all this work, and turn it back over to Steve. Thank you for your time.

Stephen Edwards: Thank you so much, Josh. Apologies to everyone for both our late start due to technical difficulties and some of our audio issues. So, we are going because of our late start, we're going to take a 10-minute break now, so it's 10:55 so will restart at 11:05, with Rusty Thomas. And we'll cover questions for Josh in the panel discussion. Thank you again, Josh.

Joshua A. Harrill: Thank you.

Break

Xiugong Gao: All right. The time is about 11:05. So, let's reconvene and then move on. So, hello, everyone. My name is Xiugong Gao and I am a research biologist with US FDA. I am Co-Chair along with Dr. Stephen Edwards.

Toxicogenomics to Support Hazard Assessment
Russell Thomas, US EPA, Research Triangle Park, NC

Xiugong Gao: I'd like to introduce our next speaker, Dr. Russell Thomas. For those who do not hear the speaker introduction, Dr. Thomas is the director of the Center for Computational Toxicology and Exposure to the US Environmental Protection Agency. So, today, Dr. Thomas is going to provide an overview of efforts made by the EPA over the last two decades in applying toxicogenomics to hazard assessment of chemicals. He is also going to share with us some current international efforts in the same area.

So, without further ado, Dr. Thomas, thank you and welcome.

Russell Thomas: Thank you. Good morning, or if you're not necessarily on the East Coast, maybe it's early, early morning or sometime in the afternoon or evening. Okay, so, can you hear me?

Xiugong Gao: Yes.

Russell Thomas: All right. Thank you, I appreciate the opportunity to talk to you today and share some of the experiences that we have at EPA. You heard a lot from Josh who really went deep on a lot of the more experimental and sort of data generation and analysis efforts. And I'm not going to be that technically oriented in my thoughts, if you're expecting that you probably going to be disappointed so my apologies for that upfront. But I'm going to try to touch on is a little bit more of a broader talk in terms of, let me share. There we go. I'm kind of having trouble a little bit sharing here. That is what I want. Alright. And I'll put it in presentation mode. Alright. Am I good? And I need to switch that, don't I, her?

Stephen Edwards: It's showing up full screen for me.

Russell Thomas: Oh, do I, am I doing good now?

Stephen Edwards: Now you're, now you're showing the two. You were good the first time.

Russell Thomas: Alright. How's that?

Stephen Edwards: Perfect.

Russell Thomas: Alright, cool. It always, depending on how many monitors you have and it's always a trial and error to try to get that right, so. All right, so let's get started. So, I do not have any conflicts of interest to declare but I always also have to give you my disclaimer that these views aren't necessarily the policies of EPA but are my views personally.

So, as Steve alluded to, my, I'm going to try to provide a higher-level overview of the application of toxicogenomics at EPA, give you a little bit of the history, and then go through that evolution, and then where we're at today, but also touch upon some of the more recent frameworks for applying toxicogenomics. And those frameworks aren't necessarily always specific for toxicogenomics but they at least from my perspective, are going to be important in continuing on the journey to apply toxicogenomics to regulatory decision making, so I'll touch upon some of those frameworks and Carole and Josh alluded to some of those, but I'll go into a little bit more on those frameworks and how they fit into sort of that journey in its application.

I'm going to go across a few more additional examples of ongoing EPA research to apply toxicogenomics in different regulatory contexts, and how that maybe will continue to build the evidence base that we continue to build for its application. And then touch upon a broad range of efforts, both within EPA as well as internationally in OECD on how to how we can build confidence in toxicogenomic methods for their eventual application, Josh touched upon a few of those efforts at OECD in the transcriptomic reporting framework and Carole did too. I'll go in a little bit, and add to that a little bit more broader across those, those efforts and touch upon a number of other additional complimentary efforts at OECD to build that evidence base. Alright so that's my, those are my objectives so.

Alright. So, I told you I, I go into a little bit more of the history, and some of its history, I will admit when I was preparing for this talk I either had forgotten some of this history or wasn't aware of a couple of the pieces in this history, and one of them that I had forgotten about when I until I started reviewing this was that EPA actually led the, was really, at least in my view, ahead of its time in certain ways, and that we had released an interim policy on genomics back almost 20 years ago in 2002. And in that policy if you can find it and read it, it's a bit of a challenge to dig that up on the web anymore. It's so old, but that policy expressed interest in using toxicogenomics data to enhance risk assessments and in a priority setting type decision context, right, but it would also consider toxicogenomics data on a case-by-case basis. In a weight of evidence type of approach, but it also conveyed at that time that toxicogenomics data at least in that setting and for those particular decision contexts that toxicogenomics data cannot alone be sufficient as a basis for decisions, but this, at least in my view was really an early stake in the ground for the agency's interest in toxicogenomics and what it can do to, and what it can contribute to these types of regulatory decisions.

Then in 2004, EPA published its first report on what those potential broader potential applications of genomics and chemical risk assessment could be. So, in addition to that, sort of stake in the ground in 2002 on that interim policy, it had more of a technical report describing its larger vision on what those potential applications could be and emphasized that those could be anywhere from prioritization, could be applied to monitoring type of applications. Reporting provisions, mode of action, as well as identifying sensitive populations and addressing the sticky problem that we continue to struggle with in addressing mixtures. Right.

However, at the same time it also laid out what the agency knew were a number of challenges in applying this newer technology and that would be linking that type of information to the regulatory endpoints and adverse effects that has really become the basis of our regulatory decision making, as well as interpreting that information in a, in a quantitative way for risk assessment. It also said that we needed to develop a more robust framework for how we can apply this data for regulatory acceptance and the age-old problem of training of risk assessors and managers to interpret this new data, right. So, if you're like me and you're looking back at this report, right you see many of those age-old problems that we are continuing to struggle with today, right, so 18 years ago we had these problems and, and I think we're still, although we're making progress I would argue, we're still struggling with some of those same issues today.

And that same year, though, on the agency used toxicogenomics information in its first cancer risk assessment and I think this is the first one. There could be another one, but this is the earliest one that I could find. And it used that toxicogenomics information in a mode of action weight of evidence approach for a particular pesticide mode of action analysis, and a risk assessment the time course toxicogenomics data was used, it was derived from rat olfactory mucosa. At that point, at a single dose and early gene expression changes, but there was a

time course data and early gene expression changes were evaluated in a mode of action context and interpreted to be consistent at least, and support the evaluation that this particular chemical had oxidative damage to DNA part of its mode of action, followed by self-proliferation, and that the late gene expression changes in that time course were interpreted to be consistent with tumor genetic potential. So, in this particular early support and use in its weight of evidence evaluation, toxicogenomics was just part of that puzzle and that overall building of the evidence. Right.

Then, in 2007, three years later, the agency released some interim guidance for microarray data submissions quality and analysis, to my knowledge that draft guidance was never finalized but it was sort of released as an interim sort of evaluation of, okay, how would the agency evaluate this type of data, how would we evaluate it for quality and how we expect it to be analyzed for the types of decisions that the EPA would be making with this? It also provided recommendations on performance-based evaluations for quality, what those recommendations would be, data analysis approaches, the agency data submissions and also data management practices as well. At that time, you know, we were expecting a, and genomic data to this day, you know, data management aspect of it is, and can be a particular challenge. You have a lot of it. And this sort of laid out what the data management practices would be within the agency and outside of the agency as well.

It issued a draft genomic data evaluation record template, DER. And also recommended the development of training modules and materials for risk assessors as well as cross-agency collaboration interactions between EPA and FDA, as well as other federal agencies, how do we as a whole federal government, deal with these types of data right and ensure that at least we're not inconsistent in how we're evaluating this type of data? But also encouraged case study applications, so that we can build confidence in in those types of decision making. Right. Again, a lot of these things are things that we're continuing to work on today, but this is an interesting read as well if you want to go back and look at the history of this.

In 2009, two years from that, EPA released its first case study. Remember that 2007 recommended a series of case studies and in building confidence and agency followed through and 2009 in applying this toxicogenomics data in human health risk assessment and particular on phthalate-related case studies. And in that case study it also outlined what it believed that the time can be a very systematic yet flexible approach to accommodate application of these toxicogenomics data in different health and risk assessment practices. Right. It focused primarily at that point in time in in using that data in a mode of action context as well as part of a larger weight of evidence, right, and that was consistent with some of the earlier recommendations and that the agency was making. It also provided some recommendations on best practices, and also highlighted the current limitations and many of those limitations were the same ones noted in, you know, seven years earlier, in those previous reports and, and in that interim guidance, that linkage to adverse effects the consistency and interpretation analysis methods, among others. Right.

So, there we are today. And I will say that, you know, we're 20 years later, and I'm going to sort of begin to then walk through some of the additional frameworks as well as confidence building approaches that we've added since that initial release and begin to build on that larger evidence base and foundation for how we, at least in EPA, are still pursuing its application of this technology and regulatory decision making.

One of those frameworks although this is not new. And Carole touched upon this, and Josh did, too, is this adverse outcome pathway framework, remember one of the limitations that

continues to be articulated, both in, in some of those earlier documents and articulated today by regulators is a linkage to the same adverse effect based, apical endpoints that we based many of our regulatory decisions on today. And can we continue to apply toxicogenomics data within a framework that allows that that connection? And this adverse outcome pathway framework, although it was introduced in 2010, it continues to mature as a framework and gains structure and confidence in how we can apply it in different decision making. And the bottom line and previous speakers touched upon this is that it's really an organizing framework and a way to organize that knowledge that existing knowledge, as well as the toxicogenomics-type data, and linking those molecular initiating events to a series of key events and ultimately to those adverse effects at the level of biological organization that's relevant to risk assessment. And I think this is really a part, a key part of the puzzle. In order to provide that interpretation of pathway and biological process-based gene expression changes in that proper regulatory content, at least for those chemicals that and we'll get into this a little bit later, that seem to act by that specific targeted mode of action context so that many of those as Josh pointed out those more selective acting chemicals that target a particular pathway or target a particular molecular target, okay, what would be the ultimate adverse effects, resulting from that specificity and that specific targeting of a pathway, or a protein or sector or an enzyme? Right. And that's I think the biggest and utility of this particular framework.

And the other thing that you know Carole alluded to was the review and acceptance of AOPs internationally, through the OECD and providing that confidence for its application, the endorse, maybe not acceptance but endorsement by OECD. And I think that goes a long way of saying okay this AOP is endorsed internationally for this particular application. Right. And that I think is the key fact and piece of that puzzle.

The second framework that touch upon was one that was published by a number of coauthors that we began to look at how we can use toxicogenomics data both *in vitro* and *in vivo* in a tiered toxicity testing framework. And then that built upon a number of data collection activities both the EPA before I was here, as well as at the Hamner Institutes and some of the *in vivo* studies that were done at the Hamner Institute's that that Carole pointed out, right, and how can you piece those together to get at that better faster, cheaper tox testing type of paradigm that we're continuing to pursue. Right. And in that that assembly of data and that tiered toxicity testing framework, you know the initial tiered testing framework include included a number of the *in vitro* assets for bioactivity, and then the, the analysis of that data to identify those that are selective acting as Josh pointed out and, and as I talked about in the preceding slide and those that are sort of bulls in a China shop. Those that don't appear to selectively target particular sector enzyme or biological target but those that that just wreak havoc at some particular concentration or, and the cells or dose and the tissue. Right. And that you can use some of these *in vitro* assays both to identify those molecular targets or to identify the dose at which that bull in the China shop chemicals began to perturb a broad range of patterns.

And then for those that appear to be either selective acting or non-selective acting, you can then evaluate in a risk-based fashion and follow through on potential targeted *in vivo* studies to confirm that mode of action or to identify what would be a point of departure in *in vivo* risk assessment context, and some of the studies that began to lay that foundation for this tier testing paradigm where those that initially showing this will be called tier one here in the middle. Let me get the laser pointer out. The tier one studies in the middle, looking at the protective in vitro screening points of departure and looking at when does bioactivity occurring in these particular cell types? When do you perturb any sort of pathways or biological processes in vitro? And in that particular context the majority, you translate that concentration

into administered dose, they tend to be protective terms of more traditional *in vivo* points of departure. Right.

And then those tier two studies where you're then taking those chemicals and looking at the concordance between *in vivo* pathway-based changes transcriptional, as measured by these transcriptomic technologies and looking for the concordance between the *in vivo* pathological effects. There's a much better correlation where when you're looking the same animal and using the same sort of pharmacokinetic-type extrapolations right that correlation becomes highly correlated and typically within a factor of three. Right. So, it's that building that tier testing paradigm of first looking at any bioactivity in vitro, doing the dose extrapolation saying, okay, I'm going to select a first tier protective points of departure, and then that second tier saying, okay, now that we have, we believe we have sufficient risk, we're going to narrow and refine that which a departure and that risk estimate through these *in vivo* transcriptome of studies to get an estimate of where we should be concerned From the transcriptomic pathway based-approach. Right. And this is going to be important in the later on in terms of its potential application, I think, in a chemical risk assessment type of context.

We also then after that particular framework released a subsequent framework of how you could potentially began to put these pieces of puzzle together in more in vitro, right, and building upon that tier one and that proceeding framework. And in those cases, really highlighting the potential application of toxicogenomics and potentially utility of toxicogenomics to evaluate in these in vitro screening paradigms that sort of got high coverage broad coverage high content type of data. So, essentially, you're attempting to cover through a series of cell types or 3D type of culture systems, the biological space that you may be concerned about in in a chemical risk assessment. And the good thing about toxicogenomics is you're not necessarily focusing on one particular pathway or one particular target, you're casting that broader net and looking for everything that could go wrong. Right.

And then beginning to at the same time, in the same way in this particular paradigm, looking for those that are selective acting and confirming that orthogonally as Josh pointed out through a different assay technology or a different cell based system, and then interpreting those pathway and target based perturbations in an AOP context, right, and then being able to walk through a risk assessment based on that specific mode of action, or for those that the bull in the China shop again taking that biological point of departure that bioactivity-based point of departure and developing a protective based number from a quantitative standpoint. Katie Paul Friedman showed this in a, in a more recent publication of how that can be done at least using a broader suite of ToxCast data and a broader suite of in vitro assays and showing that it was generally protective, more than 90% of the time if you apply this broad-based screening approach.

Some examples. So, those are the three frameworks that at least I'm going to fold into the talk later on that I believe begin to continue to advance, how we begin to think of a toxicogenomics data at EPA, and I'm going to go into a few examples of how we are both applying toxicogenomics within an EPA as well as some of the research to continue to build those. One of those examples is a more recent one where we've released, if you have noticed in last year, late last year, a new national PFAS testing strategy. And this is really to how the agency is intending to identify candidates of PFAS for additional testing, how does the agency through its test order type authority, identify which PFAS we should be testing right and should require testing of the companies that want to register those chemicals for commercial use in the US. Right.

And so, what we work through with our regulatory partners, was a way to develop initial PFAS-based structural categories, how do we route them in a consistent and overall holistic way right based on structure based categories, identify those PFAS categories with data gaps, right, and then either issue those test orders to fill those data gaps and eventually down the road, begin to refine those categories using in vitro mechanistic and toxicokinetic, as well as the *in vivo* test data.

If you've read through that strategy, you will note that we work through that, we focus predominantly at this stage on developing those categories, we began to group chemicals based on those structural similarities, breaking them apart, and then identifying using that structural based approach, what would be the most representative chemical of each category, most representative chemical of each category. We then pulled out and are nominating that for potential test order type of solicitation so they can do the more traditional test.

However, and downstream what we're also folding in is toxicogenomics data as well as other orthogonal data on refining those categories, maybe that structural category, you'd have identified the most representative structure in that category. Maybe that category although it's structurally similar has two or three different modes of action that are embedded in it. Right, that's not unusual, we know that from traditional category-based approaches, maybe parts of that category are active and specifically acting on a particular receptor or particular enzyme, we need to know that and this is where the toxicogenomics data and the other molecular and screening data can help inform, okay, maybe this particular part of that structural category is estrogen receptor active or PPAR are active while the rest of that category does not have that particular mechanism. Same with toxicokinetic. So, we're beginning to fold that in and layer it across that category-based structure-based approach to begin to make those testing level and regulatory based decisions.

The other example is the one that builds upon what Josh presented, where we were looking across three different cell types and Josh focused on the MCS7 cell type and, and we begin to be an added other different cell types of cells and the HepaRG cells that Carole talked about that. Testing about 1000 chemicals and concentration response, and then using that old transcriptome approach to both do the concentration response modeling for it when do we see any bioactivity but folding in and pulling into that signature-based approach that Josh was talking about. Which ones are those are selective acting in which ones are not? But if you look at those that generally are the bulls in the China shop, we can begin to draw that bioactivity based points of departure and begin to compare it to those *in vivo* adverse effect levels that we showed in that initial framework and seeing if this transcriptomic technology is similar to other high throughput screening technologies and deriving that more protective basements a departure. And it turns out that if you combine data from multiple different cell types that collectively evaluate a broader range of biological space that yes that we get that protective points of departure about 82% of the time and the most of the non-protective type points of departure word from neuroactive chemicals, so we need to be able to fill that biological space gaps that those mechanisms that we're not covering the neuro active chemicals, somehow and begin to plug in more neuro-based models into this testing paradigm so we can sort of fill that particular data gap, but this is again to build that foundation and that sort of tier one type application of an in vitro based toxicogenomics screening.

The other potential data gap that we've had traditionally at EPA is yes that's fine we can test those data in vitro using chemicals that are nicely soluble in DMSO or other aqueous type vehicles, but what are you going to do about volatiles because a significant percentage of the chemicals up regulates are volatile organics or semi-volatile organics. And so, we're partnering

with a group in another laboratory here at EPA, and they're building a, they built already an in vitro exposure system that can expose air liquid interface cultures in concentration response format, right, and then we can do the same paradigm are we exposing cells and different cell types, in this case we've chosen a cell line as well as human primary bronchial epithelial cells in concentration response in those air liquid interface cultures, analyzing the whole transcriptome-based responses, and then doing that concentration response model, and they've been testing the traditional volatiles that EPA, that's on EPA's list, and comparing those transcriptomic-based pathway points of departure, with in this case the TOV time weighted average type values that are published for by ACGIH and in general, there's pretty good concordance between these transcriptomic-based points of departure, and those threshold limit values derived by the ACGIH, right, and showing that at least in this particular system those bioactivity-based to these were pretty concordant with where we would expect to set those, you know, points of departure from regular regulatory decision making type context.

The other application in terms of building those signatures for mode of action. And we're partnering with our Office of Pesticides to build those signatures for particular liver-based modes of action that they're concerned with that help them identify potential concerns from cancer-related mode of action. And this particular case, Chris Corton has been working closely with his counterparts in pesticides to both, identify what chemicals they want to study, to build that evidence for this cancer-related modes of action, treat those particular *in vivo* studies with the, the pesticides that they believe have established modes of action. Right, develop those particular signatures in particular for some of those PPAR, PXR, AHR, and other receptor-based modes of action in the rodent liver, develop those signatures, build those AOPs surrounding those particular modes of action and integrate that into a cancer-based mode of action framework, and also at the same time began to look at those signatures, as Josh pointed out, as a function of dose, at what dose he began to see that that cancer related mode of action, and what doses dropped below that cancer, nongenotoxic-type mode of action response signature. And you're doing this both the concentration, as well as qualitatively for that particular decision-making context. And more to come on this but this was presented at a Science Advisory Board meeting. I think it's now two years ago, right, and we're building on this and working further to develop this particular approach.

The other thing is that we're not just at EPA concerned about humans. We also have to think about other species in the environment. And so, we've been building a capacity to extend this type of transcriptomic-based screening to a number of other species and the people up at Duluth, Dan [inaudible] and others are beginning to expand that there was a challenge that was released by those NASA life challenges, the grand challenges that they released, there was one to develop a transcriptomic-based approaches for a broad range of species that the agency needs to regulate and we can close that challenge and awarded that to a particular company that can build this type of transcriptomic approach across multiple different species, but they've been pursuing this in the meantime before that grand challenge as well and exploring how this could work in in both the fathead minnows so fish related aquatic organisms as well as invertebrates and looking at that, can we apply this particular technology to sending those points of departure for eco-related species and so far they've been testing about a dozen chemicals they expect another 20 or so later this year, testing those in concentration response in a, in a 24 hour exposure, analyzing the whole genome transcriptomic and doing that same benchmark dose modeling.

You can see on the lower left-hand corner some comparisons with the, the apical-based points of departure and the transcriptomic-based points of departure in this case they're analyzing it on, as Carole pointed out, depending on if you analyze in a pathway basis or a gene by gene

basis, they're refining how they want to determine this particular basis, this particular data, but in general that bioactivity-based points of departure in fathead minnow provided a protected point of departure, based on the more traditional apical-based effect concentrations that have been used to set ecological risk assessment type values.

The other thing that they've noted is that, at least on the data they collected so far that these bioactivity-based point of departure in fathead minnow we're not necessarily protective of effective concentrations in invertebrate species like crustaceans. Right. And so, how this agency applies this across multiple different sentinel species is going to be really important.

The other research that's going on is the use of toxicogenomics and trying to evaluate the tradeoffs associated with timeliness, costs, and uncertainty, and this is a, an important component here in terms of, of how the agency doesn't just consider a new technology, but also evaluates that tradeoff and this framework was recently published in 2022 that provides that, quantitative evaluation, so if you get that data that toxicogenomics data sooner, right, it may be more or less uncertain, but does that timeliness of that information, provide you a better decision making overall when you consider the fact that many of those chemicals you being to expose to in the meantime while you make that decision; so many of these risk assessments take multiple years to complete, but if you had an earlier estimate 2, 3, 4 years earlier, what would be the overall public health and economic tradeoffs associated with that?

I will try to speed up here because I think I'm running short on time, but this is another foundational approach to not just build that scientific evidence, technical evidence base, but also the economic and public health arguments of saying okay what can toxicogenomics contribute to regulatory decision making when you consider at timeliness, that cost effectiveness and other information on top of that scientific validity and technical validity.

We're also trying to build scientific confidence in toxicogenomics methods and approaches in a work plan. There's different deliverables in that that are really important in how the agency sees moving forward. A couple of those that I'll highlight but only briefly is that we're trying to look at that in the context of the traditional methods so we are trying to evaluate baselines of the variability and relevance of the existing methods so that we can set the appropriate expectations for toxicogenomics and new methos, also developing a series of case studies that we talked about like Chris Corton's case study for evaluate its application from regulatory decision making, as well as develop an overall agency wide science confidence framework to evaluate the quality, reliability developments methods.

We're also working at, in partnership with the OECD, Josh talked about that reporting framework, we believe that the consistency in reporting these methods is one of the basic hurdles to getting its acceptance in, and for regulatory decision making, and that we want to extend those OECD efforts to a number of different decision contexts; that reporting essentially is the only the initial step to its broader application, and that we need to then begin to pursue development of other applications of reporting modules and case studies for grouping and read across or screen level risk assessment and mode of action and adverse effect type biomarkers so that's a work in progress..

We've also done some survey of members of OECD, regulators in particular, of how they're applying it. And in that survey, we asked a series of questions in terms of what are the challenges of using omics data and chemical risk assessment? We gave them multiple different answers and allowed them to rank order those. The first on the list for many of them, about 29% of the time was a lack of confidence in omics data. So, continuing that previous

theme that I don't know about the reliability, or the linkage to adverse endpoints that that I currently use. Number two was the lack of explanation of when omics data is submitted for evaluation so that consistency and that reporting framework is important. The lack of technical guidance, how do we evaluate this in a consistent way across our regulatory organization, and then that lack of confidence in omics data was also coming in at the back end as well. Right, and the lack of relevant examples.

So, the other question was, which of the following would increase your confidence in the technical aspects of dealing with omics data, and I think that won out for that was that linkage, that AOP linkage between, between, you know the transcriptional perturbations and that adverse effect and, and I won't go through some of the other ones as well in the interest of time. And then which the following could increase your confidence in aspects related to application, and number one on that list was developing of an internationally harmonized guidance document on its application in regulatory decision making. And so, all these are consistent themes that we've seen before.

I'll skip that slide, point out that we're trying to build that confidence in applying it in IATA, an integrated approach to testing and assessment framework. Right. And we continue to do that in what's called case studies in OECD. There have been two case studies that evaluated omics data that's been published and released so far. One of them was using *in vivo* transcriptomic profiles to support a mode of action similar to that 2004 risk assessment as well as other case studies since then, and then using that as the biological point of departure across multiple different cell types.

And then finally, in summary, I think, development and application of toxicogenomics for chemical safety decisions that EPA has been going on as I talked about for more than 20 years, right. I think some of these new frameworks and research tends to address many of those perceived and real gaps in applying the technology. We're still working on both at EPA as well as internationally to build scientific confidence, that effort’s ongoing. But yet, you know, it's been 20 years, right, and we can have that conversation that we've struggled with for over 20 years to continue to apply this technology and beyond just its application and overall weight of evidence and that's still remaining elusive. Although it's somewhat maybe a bit of a pessimistic kind of bullet but I think that we are at the point right now where the future is brighter. We are developing that evidence base in those frameworks where we can build that argument, that holistic argument that the time is now for beginning to apply it both in a quantitative and qualitative way to the types of regular decisions that we need to make.

And with that, some references, and I'd like to thank the people in EPA that as well as the people we collaborate with to continue to pursue some of the new toxicological approaches that the agency needs to make its decisions in the future. So, there you go. Thank you.

Xiugong Gao: Thank you so much, Rusty, for the wonderful presentation. So, I really appreciate you and your colleague at the EPA who has been leading the way in trying to incorporate toxicogenomics into the regulatory setting, so thank you very much. Okay. And since we are behind schedule, so we save all the questions for you answer during the roundtable session, and for the participants, please send in your questions through the Q and A, and we will try to ask them during the panel discussion.

Application of Toxicogenomics in Toxicity Assessment of Chemicals in Food and Cosmetics
Jorge Naciff, Procter & Gamble Company, Mason, OH

Xiugong Gao: So, with that, I think it's time for us to move on to our next and final presentation, which is going to be given by Dr. Jorge Naciff. So, Dr. Naciff is a member of the group of product stewardship of the Central Product Safety Organization in the Procter and Gamble Company. He's going to share with us today is work applying toxicogenomics in toxicity assessment of chemicals that could be found in many of the FDA-regulated products, including food, dietary supplements, and cosmetics. So, without further ado, Dr. Naciff, thank you and welcome.

Jorge Naciff: Yes, thank you very much for the kind invitation and for having me to share with you some of the data we have actually been collecting from the work using toxicogenomics in in [inaudible] fast, as fast as possible, with trying to convey the message that I want to convey so [inaudible] a little bit [inaudible]. It seems that somebody is trying to hop in, somebody's trying to control this for me.

Anyway, I have no conflict of interest to disclose, other than what I'm saying here is part of the last thinking on the use of the, of the toxicogenomics approach and not necessarily what we try to compete to the regulators thus far. What I will try to convince you or to convey to you as is, why do we need in vitro testing the first place. by now probably everyone agrees why we need to do that but nonetheless I will make some points on that. How do we test, and which case studies to test because using toxicogenomics as, as it has been already said by all the participants, it is not a done deal yet, not in vitro systems.

Nonetheless, we want to show you some of the work using the toxicogenomics transcriptomic profiling, as well as ligand affinity assays that are more targeted for a specific molecules and specific targets for specific molecules and specific targets for some of the chemicals we were looking at.

Stephen Edwards: Jorge, are you advancing the slides?

Jorge Naciff: Yes, I am.

Stephen Edwards: Could you share and share again because we're still seeing your intro slide?

Jorge Naciff: Okay. How about now?

Stephen Edwards: Okay, yes, I'm seeing an agenda; yep, perfect. Thank you.

Jorge Naciff: Okay, I don't know what happened but that was very strange. So, I'm going to give you an example of someone that person we have taking with mixtures something that already Rusty mentioned, right, how do we deal with that? An example I'm going to use this as a political so we have many, a few of them already have a reading, I will just mention one of them for the sake of time, and then how do we use this information with different types of data to get a really at the point of how, how do we use this information to the drive a reference dose, and therefore, compare concentrations with different actives, and a brief summary of the work.

Oh, now I don’t advance. Okay, now I advance it. So, the chemicals that I will include in the presentation, some food ingredients like caffeine, and theobromine, food additives such as preservatives and parabens, and supplement ingredients, those are the botanical extracts, as well as cosmetic ingredients. Some are actives, in some cases and some are preservatives. too, because we use them all.

Why do we really need to do these assessments in vitro? By now we are doing that *in vivo* data is reliable but not quite there. And on top of that, our industry has faced a challenge because we cannot use any more *in vivo* data to support our risk assessment process that is described in this slide, and you suddenly published already there so I won't go into the details how we do that, but we need to get to a certain point where we can actually take an assay that allows us to get the same type information we used to derive from *in vivo* studies using in vitro systems and get this information to derive reference dose and use that information to extrapolate to the *in vivo* situation for our frame of risk assessment process, so I won't go into the details of this particular approach, but it has been describing two case studies. Having already published by OECD, and you can take a look at that particular with the parabens and the caffeine case study that I will describe to you.

What we're trying to do to do here is to the define, as everyone already said, a dose- and time-dependent response analysis of transcriptome in human-derived cells to use for the assessment of ingredients of interest for the cosmetic. And also, also to use this information to better define the biological states that these activities are these chemicals that are active may be covering the human being. The idea is also to integrate these data with other streams of data in the form of mode of action identification, or better yet, Hazard Assessment and general hypothesis testing for specific activities for which a specific mode of action is not there or is not being characterized yet.

And then, at the same time we also use this information to better define group of chemicals or related chemicals, under the proposition of our assessment assays that's assessment, in which we have data for a few chemicals, but there are chemicals that look alike, based on physical, chemical properties metabolism, etc. the parameters. Do you think elicit is the same biological response, that's something that we need to rely on in vitro, and it’s a way to also use this transcriptomic profiling to assess that and also to get an insight on new materials to define mode of action for this class of materials are the criteria, materials, and at the same time also use this information to define what is the outcome that we can elicit when we get this chemical in the system. And how do we use this information to define where a dose is not able to elicit this response, and which critical events are listed that higher doses, when the system is exposed to these materials?

So, the transcriptional profiling methods that we have used have been various, we have used a few platforms. But when test a chemical as Josh already alluded, we also include always in every single run and every single spreading positive controls, this case we used trichostatin A and ciclopirox, and we test runs of chemicals in different subjects. Most of the chemicals that we have designing actually have evaluated in four cell types, MCF7, HepG2, cardiomyocytes, and A549 cells that represent the lung. We have some chemicals are already in other cell types to gain an insight on what is the biological space we need to cover when use us this approach to make sure that we cover all the bases. And this is a question that we haven’t actually answered yet but nonetheless we have tried.

We use three different platforms in the approach we have taken one is 1000. That was offered by geometric but nonetheless it's limited this approach, it only looks directly at about 970 genes. We also use the TempO-seq approach that looks at the entre transctiptome and we have we have also tried to go a screen that is mostly based on the metrics on platform, but it's not actually delivered by Thermo Fisher, and it looks at a broad approach, not as comprehensive a TempO-seq but nonetheless offers about 20,000 and we can look at the changes in the expression level.

Once we've defined a transcriptional profile by the statistical methods, which is this transcriptional profile to compare with transcriptional profiles by other molecules that have been determined and actually placed on the database. And database in the broad sense database called ConnectivityMap. And this database that we have mentioned already now but the idea is that chemicals are the least similar biological activity in any cell type can be compared with the ones we are actually screening, by looking at the top ranking genes meaning operator genes and genes will be shared by different chemistries if they act through the same mode of action or use the same mechanism of action to elicit the response in the cells. We also use this as transitional profile to annotate, and they have different tools, already described some of them the GSA is one of them, and we use the Max Planck Institute ConsensusPath also to annotate the genes and build on the consensus of what pathways have been affected by the chemical exposure. The ultimate goal is to get a better insight on the biological activity that these materials may have or may not have in the system.

We know that chemicals that look alike and then examples here for chemicals that are actually known to be activating the estrogen receptor pathway, Ethinylestradiol is one of them that's very active, genistein that's medium activity compared with Ethinylestradiol that is endogenous hormone, and bisphenol A, although they don't look much similar in physical, chemical properties you look at that but they actually look very similar on the active side of the receptor, the estrogen receptor. And when we compare the response of the transcriptional level, and this is the same representation of genes affected in a robust manner meaning statistical significance and fold change at different times, and different concentrations show here; in blue are down regulated genes and in red are upregulated genes.

At very earlier time points we see a very clear response that is dose dependent for either the active or very active material or the weak estrogen active BPA. And those are equipotent doses so in micromolar range, and this one is in the low microphone range and this in the high microphone range. But the point that I want to make here is that the cells in vitro [inaudible] clearly show a response to the same mode of action of these materials.

When we compare the activity of other type of materials which activity could be very variable, and we do the same analysis. In this case you seen a battle sites, and we completely hear a set of 10 chemicals that are really not related in the structure of chemical properties with the exception of the two phthalates here, one and two here, but they are all known to be liver toxicants. And when we compare that response at the transcriptional level, and those responses I don't show that the response because it's too complex the graph here but these active doses that is comparable across the board. We clearly see that those similarities and difference between these liver toxicants.

And when we do the analyses by pathway analysis and seeing what pathways and what effects what chemicals have, we can group them in at least four categories or three categories, the ones that are known to be peroxisome proliferators, such as the phthalates and psycho phthalates on the wire compound, then two materials that are far away from the others that are

cytotoxic mostly, and those are valproic acid, and acetaminophen at the doses that we tested on an active. And then the other four that are literally just keeping users' in general terms, right, and it's very clear that the ones that showed the most similarity of the closest in the, in the transitional profiling the others but nonetheless, some are very, very similar in the, in the pattern that they elicit, and some are more active than others, even if we compare them at equivalent doses, you would say so. So, not just similar compounds by a physical, chemical properties can be identified using this approach but also materials that are unrelated these chemical properties, but they nonetheless act via the same pathway, or the same biological activity can be identified very clearly using this approach.

So, one of the case studies we already polished as I said by OECD and described here the approach we took was because of the parabens where we use those parabens as preservatives, and they are different types of parabens and I'm showing here the idea of looking at can we group them better. And can we use the data from some of them to support others? And those are so short chain alkyl chain parabens on the side here. This is methyl, ethyl, butyl, propyl, and the compound that is actually released up the metabolism it's relatively fast hydroxybenxoic acid, and we included this one just to make sure that whenever we see effects from these materials can be compared with these ones, too. The assumption was here for this case study that we didn't have any *in vivo* data or any source of data to determine the activity appropriate part of it. And we, the idea was, which one of those saw that parabens can be used to, to help us to define where is the dose response and how we can actually use the data existing there for visual field support the risk assessment of these ones.

This is the summary of the consumer profile elicited in different cell types, and this case I'm only showing the MCF7 cells. For the sake of speed, and a number of genes affected across the different materials, and dose responses. In this case, what I wanted to convey here is the idea that all of them are active and all of them are able to elicit a response, even the metabolite in the cells. And when we look at the principle which one says similar which genes were affected in the same direction by these materials, we find only 45 even including the main metabolite, we will compare only the actives the ones that are used actually, as the, as a parent compound. We found 133 genes being affected robustly in this manner. The idea was also to see can we get one of them better to support these one that's supposed to have no data, and we compare methylparaben with propylparaben than only 260 genes were affected but the ones that they showed the closest similarity by number of genes affected and 19 of the effect is butylparaben and propylparaben 634 genes affected in the same manner, the same direction, and almost the same fold change.

Here is the heat map showing the same results hopefully you can see that, at the closest, an active and most active because this is an equivalent dose, equivalent of 500 micromolar in the MCF7 cells are butylparaben and propylparaben. The other ones are less active and by the, those are comparing the lightness actually this engine recursive, and the color of the, of the description here shows the magnitude of the change very close similarity, however obviously there are differences. The most potent of all, is actually butylparaben, and this is not just us, other people have seen that.

This is just the individual gene expression changes or a few of them. And here is not to read every single gene but the idea is that they are equivalent even in the magnitude they elicit in the changes or a particular gene, the most robust responded jeans for butyl and propyl are very close, and those have been most active of active of the four compounds that we tested and shown in this particular graph.

When we annotate those genes and look at which pathways are many affected by the parabens in general and this is only using the common themes across the board with it for parabens, a combo that that estrogen response is the one that is probably the most possibly engage after exposure to the cells and I want to remind you, these are only the cells were exposed for only six hours so we can see other pathways but the point is that the, the main impact those materials may have in the cell, once they get exposed to these materials is activation of estrogen pathways, and this is shown here. When we compare one to one the most close analogs, butylparaben and propylparaben, we definitely see a higher response in these cases but it's still the ones that come about when we look at the upregulations are estrogen responsive similarly, among other ones have been showing here the ones that are most robustly expressed, or more robustly affected. And also, when we look at the related pathways, they are completely overlapping between the butylparaben and propylparaben. So, we can conclude from both of these data that if we don't have any data for propylparaben, we can rely on the butylparaben to support the risk assessment process, and we show that on the case study as I said already published.

How does it look like with other streams of data? This is actually where we see in the traditional profile mimics where we, other people sees, looking at different message, and we use here, the ToxCast data that exists for parabens. All of these parabens have been evaluated. And those are the assays that have actually been done and what is the common pathway shown here on this table. But what we saw here was that the bioactivity in our case was less for methyl and ethyl, and therefore, and more for propyl and butyl. And almost no hits in this case what identify for the, for the main metabolite.

We look at the data directly to compare a value to derive a value that we can use for risk assessment just looking at the effects on the estrogen pathway, we can use the AC10 median and compare the, the value of the different parabens, and we clearly see that butyl is a very high active, compared with the other ones but then the next one in activity is propylparaben, and the same results that we saw with a transitional profiling and very clearly we can rely on the data of any of them but not most particularly the butylparaben to support the propyl, and it's actually, we consider that it will be a little bit less active propyl than butyl based on all the data that I show you, and some other data shown here.

Moving into the case study for caffeine. In this case, the assumption was we don't have data for caffeine, but the question was, what kind of analogs can we use? Well, what better to use that the main metabolites and the main metabolites showing here, and this is work done by their Gracia-Lor and collaborators published in 2017. As I said, the assumption was that caffeine didn't have any activity or any data to define activity at the final potential valid to using the risk assessment of this active, or this food ingredient. And we used theophylline, theobromine, and paraxanthine to do the analysis of professional profile in the four cell types that I showed you before, this is just one example that the results we saw on the cardiomyocytes, and it’s very clearly those are equivalent doses that we tested here, 1,000 micromolar, that is high but nonetheless it’s the same kind of active that we see across the board with these materials. The most active is theophylline and then the other two analogs that we suggested as compared with caffeine, and the more similar in transcriptional profile as you can see just by the same alignment of the genes here is theophylline to caffeine, the other ones are active, but the activities are not the same and they are more different than the theophylline and caffeine transitional profile.

When we use this transitional profile to look at what activities assuming that we don't know anything about caffeine are actually being identified for this type of materials, and we use the

connectivity map on the Broad Institute to look for materials that have the same transcriptional profile. And this is our response look at the different profiles that we have 50 micrograms per micromolar actually don't show much activity and they have right here, the units I just notice, and I apologize for that but it's micromolar, don't really see too much of a response but when we see 500 or 1000, we definitely see a response, and there is a clear similarity with materials that have been actually evaluated, and most of them are Mtor inhibitors, AGM inhibitors, DNA dependent promoters, [inaudible] inhibitors and others, and we can use this information to build on the activity that we don't know about these material the waters obviously but this is the approach that we take with connectivity map.

When we compare the other materials and the activity at the doses that are active in MCF7 cells, is just an example in MCF7 cells, the three paraxanthines elicit a response but the closest, and the most similar to other type of materials are the least stable response that we saw with caffeine, are theophylline and caffeine, obviously, but there are some overlap so to me because those are also actives at some point and some concentration. From this analysis, and this is just a frame of comparison between the caffeine and theophylline, the dose response analysis, the highest, the highest concentration, the two highest concentrations are very similar in response to their response to the [inaudible] And those are the closest samples that we can use in the risk assessment process if we happen to not have data for caffeine and we just rely on theophylline database. This case has also been published by OECD and you can take a look at that.

Now the question is can we use the same approach transcriptional profiling evaluation to really assess the activity of complex mixtures and in this case, in particular botanical extracts? The answer is yes. But it's not an easy task to do. So, when we go for a mixture of botanicals that could be one single plant or could be multiple plants that we alternate botanical extracts that we can do some point as a supplement the question is what activities really those extracts may have if we don’t know much about them?

The first thing we need to do is some analytical analysis of the materials to have an insight of what is in there, and there are two publications that I hadn't put forward, and they are here to describe really what is needed. The really bottom line is we need to know what main components and the chemical labor are part of this mixture to get an insight of what kind of activities we should be looking at, if any, if we don't have any information and can we still use the traditional profiling approach. Together, to find potential biological activity. The answer is yes. So, we also using this case. Another approach not just the transcriptional profiling to identify biological activity, we also use like bioaffinity assays to better provide a focus on molecular level, these activities. Those are the materials we have tested and we decided to test those because those are actually not recommended for pregnant women or women that are lactating for various reasons, they may have activity on the reproductive and developmental endpoint in particular. So, I'm going to just focus on ashwagandha, and this is the material that I mentioned but there are other botanicals that we have tested and this work has also been published and there are references you can take a look if you want to.

This is the summary of the response of various cell lines, in this case we created in Ishikawa cells because as I said, this type of material may have an impact on the reproductive system. So, we include the Ishikawa cells because are cells derived from the uterus of a human. And ashwagandha was tested at this concentration and the units are micrograms per ml not milligrams. And those have the genes affected in a dose-responsive manner. That was really neat effect that low doses and we reach a significant does we significantly see changes in a robust manner at almost every single cell line. And this is important because we can select the

most sensitive cell line when we just look at number of genes affected in a robust manner, or we can look at all of them and see if there is any overlap in the activity and we have done both parts.

When we use the transcriptional profile in a particular setup to compare with the database that is in the, in the system. And the Broad Institute database, and we look at one materials actually elicit similar transition profile than these mixtures, we identify some of them, and the score means the highest score. The closest the similarity between the materials that we're testing and the materials that we look at in the database, and two of them come out here and there are two and different places because they have been tested in MCF7 cells, too, but at different concentrations and this is with withaferin A in this case and there are other ones, but there are many more materials that have been tested obviously in database than in individual as individual chemicals.

And when we take a look at that. Why is that withaferin A there in the connectivity as a one of the materials that elicit a similar response. Here is the response why: the level of concentration in the mixture that we have one of the analyze that console, has a high concentration productively high compared with others is actually withaferin A, among others. If any of their material that is actually part of the, of the chemistry that we are testing here would have been evaluated in any cell and would be database, we would actually be able to identify it. But the only one that has been tested as an individual chemical was this one, and we clearly identify that in the Connectivity Map approach that we took.

Not just that, obviously, we compare the other cell lines, the response of the other cell lines at the same concentration with these materials, the connectivity map, and in all cases, the one that comes to the very top, a similar material that has been. This is withaferin A, and we can help us to define what is the main biological activity of these materials and we should expect when we analyze it, [inaudible] eliciting the cell lines that we tested. So, there are many other activities there. Some of them will make actually sense when we look at the individual pathways being affected, and the top 10 are listed here, actually more than 10 but they are tested here, test, show here in this table you can take a look at that for the sake of time.

When we look at more specific pathways because you can look at the entire, let's say, transcriptome part was having an affiliate, or you can focus on a specific pathway, you have a reason to believe that some of the pathway should be actually modify there based on what is known about the material that we're testing, and we know that in this case of ashwagandha, and it has been used by traditional medicine in different to treat different illnesses or different conditions [inaudible] of these particular set of chemicals.

The question here is, does what people think or people use this extract for make sense with different sort of why we see as biological activity answer the question is, how can we be sure of that? So, the main biological activity of having [inaudible] has been listed here and inflammatory anti-tumor and antioxidant etc. And when we look at the vivo pathways really it does make sense with the transcriptional profile, these activities are supported by this specific pathways that we look at. When we use the profile for the analysis. So, definitely this approach can tell us much about the activity of these materials, and actually support particular activities that have been associated with the use of this composition.

To make sure to ensure that what we found out the transcriptional level can be actually translated in in a usable manner. We also use the ligand affinity assays together with an insight on, on the molecules of the sector. And we have used, not just this approach the

pharmacological profiling score, and this is the reference to that approach. We also use the characterization of individual, analyze that that the extract has run through a framework that we have developing in house to define potential chemicals with the potential to elicit developmental toxicity and the references also there. Using one approach is a way to gain more insight and better define potencies and activities of these materials. The screening assay that was used was the SafetyScreen 44 offered by Eurofins that looks at different agonists, receptor binding assays, ionic channels, and they are listed here.

For the dose that we choose we selected a particular those based on what is known. Cmax of these ingredients of this botanical extracts but also, we use the activity that is actually used to define the equivalent of a batch of extract to another. And in this case, we also included the potential metabolism effect on the extract because we didn't include that on the cells, when we use the transitional profile. Those cells may have some metabolic capabilities, but it's limited that we know. So, in this case we deliberately actually use human as an infraction to evaluate the response that we see on the individual molecule targets here and work on that. And ashwagandha as a marker we use this particular compound to define the dose we use at the highest dose. Ver nice dose response analysis for every target that we analyze.

Here's the summary of the main molecules that we look at the most responsive of the assays that we see and the arrows here point out which ones was affected by ashwagandha, and this is without metabolic activation. And this is when we do actually a metabolic activation for 90 minutes. The most active of all or the most say affected by all is actually progesterone receptor. When we metabolize the mixture. And then we do the assay, progesterone receptor activation is gone. So, can we use this information to build on that reference as possible using these data?

The answer is yes. And here is what we think could be an approach to do that. Obviously, this is not the only one, but this is the most conservative. There is a reason to believe ashwagandha may have an impact on progesterone system based on the profile, and also on the response that we see at the individual assay that we, we, we developed. And we use the IC50 for progesterone is shown here on this graph that is just showing the different [inaudible] in a linear dose kind of response here. These IC50 now and it's 2.5 nanomolar for progesterone, this is endogenous hormone. When we look at the same activity ashwagandha without metabolic activation without metabolism. The IC50 56.7 micromolar or the slightly higher obviously. So, ashwagandha once is metabolized, the IC50 is way above the level that we described here. So, we use this particular level or this particular IC50 to derive a reference value to support the use of the ashwagandha we actually not concerns for the safety of the user. We will be in a very good shape using this approach. And it's also a validated by the transcriptional profiling data that we find no concerns for other patterns that could be more susceptible and ensure the data but it’s there. It’s also published this work so you can take a look at that.

So, in summary, biological activity of cosmetic ingredients can be characterized through in vitro systems using transcriptional profile. We know that gene expression analysis in particular is comprehensive, it's a lot better and use with the connectivity map approach can help us better define biologic activity of these materials. It has a better coverage that *in vivo* assays for individual targets. It can help us very well for functional read-across through pathway analysis of connectivity map or just by looking directly on the materials that we tested, if they're compatible or not to define which would be a better suitable analog to read across in a case

study where we don't have data or sufficient data we have data gaps, for some of the materials that we went to use.

The other approach that can be taken is the Cerep or the ToxCast data that is plenty already for many chemicals that are more than 800 assays individual assays for different molecules in ToxCast database, and can also help be helpful. And then we use them together, obviously can help us to better define biological spaces are covering, and the biological activity that the materials are testing they have, and at the same time help us o define if there are false positives that could happen, obviously, in some of the instances.

So, this is not done. And there are plenty of work to be done. And one of the main things that we are trying to address now is what kind of materials should be tested and where and here I'm showing you some of the chemicals we have tested, most of them in these four cell lines. The long cells had to do the MC7 [inaudible] or the cardiomyocytes, and in some instances, we have tested these materials in other cell lines, and those are showing here. These are the chemicals at different dosages, and I showed you the dose in micromolar. But the point here in every single cell is some numbers, those numbers are genes that are affected when we look at transcriptome, and those genes are affected robustly.

That's fine when the materials are active robustly. Any of these four cell lines show some response, and we tested in other cell types, they seem to be also very responsive, and they show a robust response by number of gene says looking directly at those numbers, but what happened we have chemicals that don't show any activity in these four types that we have been using. And just happened then we test the time we see a response when we say nothing in the others.

And in some instances, it seems that they don't really, they're not responsive, at least in a large relatively large number of cells. So, the question here is, when we present these data to the regulators, the question is, what did you cover all the biological space when you put the testing? Is there a potential for a cell type that you didn't include in your evaluation that could actually be responsive to this class of material or this chemical and why we don't have the answer yet, and we need to get it at some point we want to use this approach to support the safety of the materials that we're putting the cosmetic so in any product that we use on the market, or in the food supply. Either way, we need to get that information.

So, thank you all for your attention. These are the participants that have been helping us along the line, this is the people from P&G, Yuqing and Xiaohong doing the work in the lab, Nadira and Catherine guiding us, and George obviously leading us with some of the work that we need to do. Part of this work has been done in collaboration, via Cosmetics Europe, and those of the participants, different companies represented here by the people that are listed here and hopefully I didn't miss anyone, the consortia is very supportive obviously of this work because, as I said, we need to support ingredients that we cannot be testing anymore in animals, and also BioSpyder that somehow gets cut here on the slide for helping us because this is the platform that we're using the most now because it's a comprehensive one.

And as I said, there are some references here for you to read if you want to and get more details. And I thank you all for your attention.

Xiugong Gao: Alright, thank you so much, Dr. Naciff, for the presentation. And I really appreciate the fact of the composite case studies of particular interest to the FDA audience. So, again, since we are behind schedule so we will save all the questions for you to answer

during the roundtable session. So, with that, I'll pass on to Steve will take us to the roundtable discussion. Thank you. Thank you, Steve.

Stephen Edwards: And I'd like to thank all of our speakers and ask the speakers if they wouldn't mind turning their videos on. And then I'm going to hand over to Jason Aungst from the FDA who will actually be moderating our panel discussion.

Jason Aungst: Okay. Thank you, Steve. We did have a few questions in the question and answer in the chat, I think, some of them partially answered during the discussion here but we can go through them again and get maybe a broader discussion. So, the first one was: are the cell based assays typically done with neoplastic cells, like MTF7 cells, and have we studied whether malignant cells are similar to a normal healthy cells in response to chemicals. And that's a good and easy shortcut for transcriptomic studies.

Jorge Naciff: I will just say something very fast, not just in this type of cells that are derived obviously from tumors or from cancer patients. We have also tested primary cells. And when we compare for example the response of HepRG cells versus primary human, the response is very similar. And is there is a huge redundancy on the transcriptional profile that we see with chemicals that affect hepatocytes in this particular case. A collaborator in the lab has been trying microphysiological systems. Those are derived from stem cells, in some cases, and so many from tissues that have donated and they are built in microphysiological systems, and we see some similarities with some of the individual cell types that we have tested too. The problem with a microphysiological systems is that they don't, they don't have them in production in sufficient amounts in a robust manner that we can use for example, for cases study when we use a few chemicals with multiple doses and multiple times. There is a huge variability from one set of microphysiological systems to the next. But nonetheless, yes, those are also responsive, and they share quite significant similarity with what we see in in the cells we have tested in our case.

Jason Aungst: Okay. Thank you. The next question was, if there was no information about possible toxic effects on a particular chemical substance, how could talk toxicogenomics be used to support hazard assessment of that chemical? I believe Jorge was addressing this somewhat with the extent of types of cells.

Jorge Naciff: Yes, I didn't want to say anything I wanted to give everyone a chance to say something. We can use it in two ways. One is if we don't know anything about the chemical, we can do toxicogenomic profiling in a few cell types and see if that transitional profile matches something database. The database has some bias to materials that are pharmacologically active. And that's part of the Broad Institute's work, obviously, but EPA is looking at other instances with materials that we can compare with and see, is there any data that can guide us, or what is the potential toxicity to these materials, we don't have access to that either. We can rely on the transitional profile and look at pathways, what pathways were affected? And we do the dose response analysis of this translational profile in the most susceptible cell type that we probably can find, then we use that information to define value that will ensure that no transcription happens and the change in the transcriptome happens at a given dose and use that as a reference those to derive a value, where we can then expose to the self can get exposed without expecting a response.

Russell Thomas: Yeah, and I think we, we have a similar concept right in our traditional testing right where we have a limit dose, so you don't see an adverse effect of that particular dose it's then you know you're assuming that you know that you're testing up to that, but

there's no toxicity at that level right and assuming that you have done the appropriate quality control analysis on the chemical that you're testing in your in vitro study, that you know what the chemical is and, and that it is, you know, being the cells are being exposed to the right, chemical, and I think we have at least a product across a broader speed of cell type there's, you know, we're building a reasonable confidence that has meaning has the right that that negative does have meaning and that's the one thing that we need to continue to work on is what are the negatives mean and what do we need to do to ensure ourselves that that negative value does mean what we think it means.

Jason Aungst: Right, thank you. So, can we talk about ingredients and food additives, we're talking about an oral route of exposure. How is metabolism considered in an extrapolation from toxicogenomic data to apical endpoints and *in vivo* data?

Joshua A. Harrill: I think the answer to that might, might be related to the same type of answer you would ask for another type of high throughput screening assay like a targeting screening assay. If the metabolic capacity is absent from that system, then, then that is a big black box as far as extrapolating from an *in vivo* situation. But there are technologies coming online now, I think Jorge mentioned one in his talk, Rusty mentioned another, important you can actually in the in vitro situation do in vitro metabolism of a particular compound and mixture and apply those metabolites to your test system, which in some ways accounts for the type of metabolic profile you see *in vivo*. So, it's kind of an indirect way to look at that, rather than a computational way.

Carole L. Yauk: And it's one of the reasons that we're moving towards these organoid models that have more metabolic competency than the two-dimensional formats of the same cell types and the non-metabolically competent immortalized cancer cell lines and when you compare the metabolic proficiency of those spheroids, it's really closely approximating primary human hepatocytes when you're talking about hepatocytes which is really the gold standard. So, I think, I think we're doing a much better job these days with dealing with metabolism but you know for other models we have a ways to go.

Jason Aungst: Okay. Two more questions from the chat. Can you comment on the criteria for selection of the *in vitro* experimental model and dose selection?

Carole L. Yauk: I guess, you know, one of the things that I would point out, and Dr. Naciff's presentation really showed this, is that these technologies are getting more and more high throughput and more and more cost effective. And so, a lot of cell models are being used in studies that you can see all of the cell models that he used. So, we're not limited to one or two cell lines anymore and as we move forward, I think we'll see a broader biological space. I think you have to consider what the question that you're asking is, if you're working in on a group of chemicals that you're hypothesizing operate through a specific mode of action, for example, then that would govern what cell model, you'd prioritize for use, but it doesn't restrict you, you know, for example, we're working on bisphenols, and so we need cells that express the estrogen receptor, or we're working in in an area where we're, we're interested in genotoxicity, and we need to ensure that the cells have a p53, an intact p53 response, so those things would govern your cell type selection but, but for broad screening. The question is pretty open still which cell lines to use and how many to use and are they going to be 2D or 3D. And so, those are interesting discussions.

Joshua A. Harrill: It's a challenging question and we look, we've looked at it from a slightly different lens. In the context of the second situation you described Carole where you don't

really have information in the chemical ahead of time. So, where do you start, what cell line you start with? And so, with no information, it's not, it's not a crapshoot, you basically look for a model where there may be somewhat similar information in some other assay system and that's where you start. We put a lot of effort into understanding, okay, if we're going to test these on them chemicals across the cell line battery, how can we, how can we minimize redundancy across the cell lines that we're testing, how can we make sure the battery has complementary biology? So, you're not wasting all the time and effort looking at the same components twice or three times in very related cell models. So, you can start anywhere on that spectrum, but from point A, point B needs to be quite different, right? So, that's our creative answer to that for now.

Russell Thomas: And what Josh is getting at is some efforts that we have in collaboration with [inaudible], what combination of cell types given a fixed budget, what combination of cell types maximizes your biological space that you're evaluating? And while minimizing as Josh was saying the redundancy in gene and pathways that are being evaluated and so that maximization, you know, could be, you know, a breast cell type so tight with liver cell type and a lung cell type right, if you can only test three, what three cell types, maximize that biological space and build it from there.

Jorge Naciff: We have used in silico tools a priori, right, and we have the molecule, we know which molecule we deal with one individual chemical at the time is like, is there any concern for a particular structure based on just physical, chemical properties, potential metabolites, etc. and use that information to guide us, do we should be concerned about for example particular response in a cell type, then, if that is the case, like for example, we deal with phytoestrogens that we use frequently in our company. We look at systems that will be responsive to these type of materials, and the cell type selection would be that obviously we've tried to cover it other options because the activity is not unique to just estrogen system but also other potential targets and using the tools that are available to look at side change or moieties or activity etc. can help us also guide us where to look at, because As Rusty said we cannot do every single cell type, we don't have the capability and the money do that. And the best way to start doing that is what will be the best cell type or cell types that we need to evaluate to get a good insight on activity?

Jason Aungst: Okay, thank you. I think the discussion we just had somewhat answers the next question that was in line here about what level of AOPs would need to be built and what type of data would be needed to provide a large margin of safety and whether or not science is still getting there on, you know, a comprehensive assessment of AOPs in way to support an assessment without having to do the *in vivo* data. And we talk about structure analysis and other structure activity analysis. So, I think we have somewhat addressed that question. One more question. This came in. Is toxicogenomic data subject to multiple interpretations? What would be the impact?

Carole L. Yauk: I guess, I would think it depends on the context of the application; if you're looking for a transcriptomic point of departure, for example, I think that there will be, you know, people can derive transcriptomic points of departure in a variety of different ways. But the work that I'm seeing emerging from our own studies as well as colleagues from around the world is that, you know, different ways of deriving the transcriptomic point of departure produce a similar point of departure, so if the if the application is, I want to know where this what concentration or dose there's biological activity to know if there's any potential risk from that, then I think there could be a great deal of consistency. Also, the use of biomarkers sort of removes that subjective piece about interpreting, for example, pathway perturbations and other

things and that's why my group has been moving more towards the use of these biomarkers, the transcriptomic profiles, the signatures are more reproducible across technologies and cell types as well as, you know, trying to identify differentially expressed gene or enrich pathway or something, we see a higher level of concordance and agreement in biomarkers like signature-based approach like C mapping those kinds of patterns that you're looking for.

Russell Thomas: I would agree that we're beginning to converge on and becoming mature and how we're interpreting data so that we get that consistency. I think that was one of the you know the challenges that were pointed out in those multiple documents that I highlighted in the history of toxicogenomics at EPA, that a consistency that the regulatory agencies are, are able to interpret the same data in a very similar way, which similar conclusions right, it's still a sometimes still a challenge in our traditional tox testing that. But I think we're still at least in a toxicogenomics end beginning to converge on a subset of methods that can begin to answer those either mechanistic questions or those more quantitative point of departure type questions as well.

Jorge Naciff: One of the points that no one mentioned, but already was discussed was what Josh described on how to present the data, right, how do we publish the data and we report data? If we report the data in a way that is consistent across the board. Anyone that take a look at data can grab the raw data and process their own way should then come very close to the similar conclusion that the original proposal of the data is, and this is not something that we can change it just, that's great. We have done some work that some other people have taken and they asked us with the raw data and do the analysis and come to the same conclusions. So, yes, they can interpret the data different way but if we use the same set of parameters so how would report and how we use it, and describe what tools we use, we should be able to get a good concordance on conclusions from these type of data.

Carole L. Yauk: I think that's a great point and OECD's omics reporting frameworks is the first step on the more technical aspects. And really the important work to come is those application reporting modules where, where you're describing exactly what you said, how you took the output from those data analysis pipelines and used it and applying it to making a decision about the chemicals and then there'll be transparency in how that was done, and others can compare how they would do it against that but the full transparency is what's required for that understanding and uptake.

Jason Aungst: Okay, thank you. I think we might have discussed this a little bit in the, in the discussion. So, what applications does this panel think are ready for prime time, so where can we start putting this in and replacing maybe some *in vivo* studies or other studies that are currently more traditionally used?

Carole L. Yauk: I would say that, you know, for data poor chemicals. When you have no other data, this is ripe for application, you know, when nothing else is available the transcriptomic data can tell you if there's any transcriptomic activity, and that transcriptomic activity as we've seen in these presentations is associated after chronic exposures with endpoints that are adverse. So, in terms of determining whether there's a potential risk posed by these chemicals for data poor chemical I think that's really ready for prime time. And, and, and then prioritizing of course if there is activity what the next test should be. Integration into, you know, integrated approaches to testing and assessment, I think there's excellent examples from the OECD as to how and other and other publications as to how we can integrate omics data directly into what we're already doing in terms of decision making and how that enriches and enhances the weight of evidence evaluation. So, these are clear prime time applications in my opinion. And

the potency comparisons you know for chemical grouping. To determine similarities between that the highly toxic chemicals within the group and the substitutes is also, I think very close to ready for adoption.

Russell Thomas: That covers it. I think that those three particularly for data poor, the three that Carole, point of departure, whether it's *in vitro* or *in vivo* and depending on the decision context. Right, either a protective or a predictive point of departure. Right. And those are, could potentially two different things and to evaluate that you with that. depending on the type of decision you need to make. Signatures in certain cases for, you know, supporting a particular mode of action or, or key event. And likely that can be in different weight of evidence contexts as well so it really, it's a broad range of decision context, at least in my opinion. Transcriptomics and, particularly with these new reporting frameworks and becoming, being able to, you know, be more transparent about how the data was generated and how it's analyzed that we're, we're certainly converging on a, on a broader application of these tools in a different, in a range of different decisions.

Jason Aungst: Okay. Thank you. Steve, do you have another question?

Stephen Edwards: Yeah, I'll throw out a thought for the discussion. So, and it's kind of a loop back to Scott Auerbach's comments and questions about the AOPs because I too think a lot about the time required to develop these things and, you know, covering all of biological spaces is daunting, but you know one thing that I thought Carole's PFAS did a great job of highlighting is, you know, kind of the tiered approaches and how you can learn from one to the next, and the other thing that that I was struck by all four presentations, was all of you incorporated kind of this component of, well, if we know if we know related chemicals, right, we can we can do a lot of stuff. And whether it's Read Across or grouping, or you know any number of examples that you showed. And so, you know, I like to think about what if your if your tiered testing incorporated that, right? So, you, you already on the front end as Jorge mentioned, you know you're going to know something about the structure of the chemical and what that means in terms of what that chemical might do, when you do your initial round of screening you're going to learn a little bit more, right, does that chemical look, does the biological signature of that chemical look like those related structures or not, right? And so, it seems like at each step of your testing paradigm you're either going to get confirmation, yeah, what I thought going in is correct, or no, things are looking a little strange and so it would seem like that would give you a really good indication of which chemicals need to keep moving down that road versus which chemicals you'd feel safe to kind of off ramp and say, yeah, we know enough, so I'd love to hear you know everyone's opinion on kind of how that could work.

Joshua A. Harrill: In the diagram I showed in my talk, kind of the flow chart from the CompTox blueprint that application point within tier one, the comparison of profiles across chemicals that we've tested sits right at that junction. So, one of the main things we use to make inferences about chemical bioactivity from and subsystems indeed whether the profile looks like something we know more about. And in my mind, that pushes you more toward the lefthand side of that chart where you want to confirm that bioactivity doing something else and flow into an AOP system, or the lefthand side where, okay, this chemical doesn't really look like anything that has a mark of toxicity we know about or looks more like a general cell stressor. Can we just set our, you know, our molecular point of departure based on the most sensitive perturbations, so that branch point is where a lot I see a lot of the profile comparisons that first come in [inaudible].

Carole L. Yauk: That profile—oh, sorry.

Russell Thomas: No, it's okay, Carole. Go ahead.

Carole L. Yauk: I was just going to say that profile paired with the extent of bioactivity. So, if it's doing a lot, then you can assume that it's probably not going to be great.

Russell Thomas: Yeah, and building on what Josh said is that, you know, in order to, at that junction point, and in order to begin to make decisions on those points, we don't have to know all AOPs and have signatures for every target. Right. You know, I think we can still begin to make decisions on that bioactivity point of departure and accumulate those signatures and accumulate that AOP knowledge over time. And as that knowledge base grows we can make and identify more and more of the selective chemicals, of which we can derive more mode of action based and mechanistic based points of departure but until then we can still, I believe in my, in my opinion, begin to make quantitative evaluation based on that bioactivity point of departure and I will point out at least in an EPA space where I would say most of the chemicals are likely to be non-selective anyways, you know, my guess, with the experience that I have is that less than half and probably more like 20 or 30% of the chemicals that EPA is out there in the environment, it's under the EPA's jurisdictions are probably selective anyways, that the vast majority are these non-selective chemicals in that bioactivity point of departure is probably going to be as good as you get, right, that it's only for a minority of chemicals that you have a true predominant mode of action and molecular target. And that you're going to have to be able to use that signature approach to identify what that mechanism is, right. And that's my own judgment, but that that junction point is that a critical decision point where you're trying to discriminate between those types of chemicals.

Stephen Edwards: And it would seem that you could then use that, you know, to even guide your AOP development, right, in cases where you need more understanding. Then, you know, that's where you need to develop your AOPs, and in cases where, you know, the decisions are pretty straightforward, you know, those are cases, so rather than thinking of it as we've got to have an AOP for everything, or we don't need AOPs for anything. You know, I just, you know, even as an AOP evangelist, right, I feel like there's, there's a role, and we need to understand what that role is and not be held to, everything's got to be, you know, fully characterized before we move forward.

Carole L. Yauk: Yeah, I would agree, Stephen, and in fact the AOP framework itself can help in the interpretation, just applying that framework itself if you have a dose response in a time series experiment you can use that to interpret your toxicogenomic data to enable, you know, development of AOPs, like you said, in real time, or for predictive toxicology if you can, if you can find those key events that align with the, with the pathways, or genes or signatures that you're seeing.

Jason Aungst: Okay, we've received one more question. So, regarding reproducibility challenge, how much information is there to characterize the extent of variability in T points of departure, or other measures of potency from repeats of the HTTr experiments? Separate experiments, loss of cells, etc., not on a plate, reproducibility only?

Russell Thomas: Josh, you want to take that one because a lot of your reproducibility is actually separate experiments?

Joshua A. Harrill: Yeah, yeah, I can, I was thinking the same thing, so within the within the context of our HTTr or high throughput screening paradigm, as the question mentioned, we do

have, we do measure reproducibility from plate to plate, but those screening campaigns are conducted over many months, using separate biological replicates or independent cultures and different screening blocks. So, at least, it is limited by the fact that is still within a laboratory, in this case, but in screening campaigns we've run more recently, the experimental design has changed in that, instead of including triplicates of reference chemicals on the plate, we actually including a full concentration series on our plates. And so, you can actually evaluate the reproducibility of your transcriptomics point of departure for a small set of reference chemicals across many months or many, many different independent cultures in the same, the same cell type in potentially, there's sometimes we use the same chemicals in different cell types, so you can look at reproducibility there. But again, that that's still limited to, in that particular instance, data from one laboratory. You know, over just a, what you could call a short period of time, months to a year. So, the interlaboratory reproducibility, I think that's a much larger question, and I'm not sure there's a lot of data out there at this point at the moment. Could be wrong.

Carole L. Yauk: And Josh, are you seeing good reproducibility over time of your t PODs?

Joshua A. Harrill: Oh absolutely, yeah, we usually get them within one Walton order of magnitude, which is what you'd expect for any *in vitro* screening assay. I'm sure, I'm sure if we dig hard enough, there are other instances where you can find at least a chemical that was tested in a similar test system in some concentration series from different points in time and what happens. I don't have direct knowledge of that.

Carole L. Yauk: That's a really good question though and maybe there are data sets that we could look at to compare some of your tPODs pods to our own tPODs or to others, you know. For example, the points of departure for HepaRG spheroids versus the points of departure for InSphero spheroids would be interesting to compare those if we have some, some chemicals in common.

Joshua A. Harrill: Yeah, I agree, I think, of course, it becomes easier if don't say POD and just say profile. Are profiles comparable across labs, technologies, platforms, time. I think the answer there is yes.

Russell Thomas: And also the reference material become so important, right, that, you know, that was availability of those reference chemicals to make sure that you know that not only the analysis paradigm, but also whatever way you're measuring the gene expression changes are comparable as well, so I think that that's a key component in that overall confidence building type of exercise.

Stephen Edwards: Well, with that, we're coming to the end of our time. So, I would like to take a moment and thank all of our speakers slash panelists and Jason, and Xiugong, my co-chair, for a really great colloquium. I want to thank all of the participants for coming and all of the great questions and comments that were put into the Q&A pod and the chat. I think this has been a really productive discussion.

And if I can get the slides to advance, I would like to make people aware that we do have two more coming up in the 20-23 year, so we will have one of one more towards the end of 2022, and another one in spring of 2023. Right now, the next two topics are Toxicological Approaches for Assessment of Low-Level Process Constituents and Immunotoxicology in Food and Ingredient Safety. These are still under discussion and may be tweaked a little bit, but those are the tentative topics for our next two colloquia.

And then, just as a reminder, this was mentioned at the beginning, but all the information from this colloquium will be available on the web. After the editing has taken place. I'll also note that Josh has graciously agreed to re-record his talk for us, since we had some audio problems during the live talk. So, if anyone would like to hear what they missed in some of the audio cut outs, please stay tuned and check out the recording once it's available on the web.

And with that I will reiterate my thanks to both the speakers and all of you. And you will be receiving a survey after this, and we really, really appreciate your input. And I'm not only on the chairing this session, but I'm also on the organizing group, and I can tell you that we take those surveys very seriously and we do appreciate the time it takes for you to fill those out.

Thank you again. And with that, I believe we can adjourn.

Jorge Naciff: Thank you.

Carole L. Yauk: Thank you very much.