WEBVTT

NOTE duration:"01:13:16"

NOTE recognizability:0.794

NOTE language:en-us

NOTE Confidence: 0.87841828125

00:00:00.000 --> 00:00:02.695 This is my it's my pleasure to

NOTE Confidence: 0.87841828125

00:00:02.695 --> 00:00:04.460 introduce Jeffrey Townsend as

NOTE Confidence: 0.87841828125

00:00:04.460 --> 00:00:07.060 the today's grand rounds speaker.

NOTE Confidence: 0.87841828125

00:00:07.060 --> 00:00:09.775 Jeff is the Elio professor of

NOTE Confidence: 0.87841828125

00:00:09.775 --> 00:00:11.339 Biostatistics and Professor of

NOTE Confidence: 0.87841828125

00:00:11.339 --> 00:00:13.280 Ecology and evolutionary biology

NOTE Confidence: 0.87841828125

00:00:13.280 --> 00:00:15.317 and the Co leader of the Genetics,

NOTE Confidence: 0.87841828125

00:00:15.320 --> 00:00:17.268 genomics and epigenetics program

NOTE Confidence: 0.87841828125

00:00:17.268 --> 00:00:19.216 at Yale Cancer Center.

NOTE Confidence: 0.87841828125

00:00:19.220 --> 00:00:22.538 He received his PhD in organic chemistry

NOTE Confidence: 0.87841828125

00:00:22.538 --> 00:00:24.853 and evolutionary biology at Harvard

NOTE Confidence: 0.87841828125

00:00:24.853 --> 00:00:26.874 University and in 2019 received

NOTE Confidence: 0.87841828125

00:00:26.874 --> 00:00:28.859 the prestigious membership in the

NOTE Confidence: 0.87841828125

00:00:28.859 --> 00:00:30.850 Connecticut Academy of Sciences and.

NOTE Confidence: 0.87841828125

00:00:30.850 --> 00:00:33.293 Engineering for his work in

NOTE Confidence: 0.87841828125

00:00:33.293 --> 00:00:35.505 developing innovative tools for

NOTE Confidence: 0.87841828125

00:00:35.505 --> 00:00:37.717 to study population biology,

NOTE Confidence: 0.87841828125

00:00:37.720 --> 00:00:39.484 including evolution of

NOTE Confidence: 0.87841828125

00:00:39.484 --> 00:00:40.660 antimicrobial resistance,

NOTE Confidence: 0.87841828125

00:00:40.660 --> 00:00:43.125 disease evolution and transmission and

NOTE Confidence: 0.87841828125

00:00:43.125 --> 00:00:46.470 evolution of of tumor biology tumorigenesis.

NOTE Confidence: 0.87841828125

00:00:46.470 --> 00:00:49.041 His research enabled curtailment of

NOTE Confidence: 0.87841828125

00:00:49.041 --> 00:00:51.274 pathogen evolution outbreak mitigation

NOTE Confidence: 0.87841828125

00:00:51.274 --> 00:00:53.734 and used to inform therapeutic

NOTE Confidence: 0.87841828125

00:00:53.734 --> 00:00:56.050 approaches in cancer metastasis.

NOTE Confidence: 0.87841828125

00:00:56.050 --> 00:00:58.550 So in recognition of his

NOTE Confidence: 0.87841828125

00:00:58.550 --> 00:01:00.054 prominence in the field,

NOTE Confidence: 0.87841828125

00:01:00.054 --> 00:01:02.572 in 2021 Jeff was selected as the Co

NOTE Confidence: 0.87841828125

00:01:02.572 --> 00:01:04.609 Chair elect of the Cancer Revolution

NOTE Confidence: 0.87841828125

00:01:04.609 --> 00:01:07.066 Working Group by the ACR and his

NOTE Confidence: 0.87841828125

00:01:07.138 --> 00:01:09.917 lab is currently working on on many

NOTE Confidence: 0.87841828125

00:01:09.917 --> 00:01:11.786 projects including buying developing

NOTE Confidence: 0.87841828125

00:01:11.786 --> 00:01:14.490 bioinformatics tools for cancer

NOTE Confidence: 0.87841828125

00:01:14.490 --> 00:01:16.518 genetics epigenetics epidemiology.

NOTE Confidence: 0.87841828125

00:01:16.520 --> 00:01:18.692 And nonlinear mathematical

NOTE Confidence: 0.87841828125

00:01:18.692 --> 00:01:21.588 models of disease epidemiology.

NOTE Confidence: 0.87841828125

00:01:21.590 --> 00:01:23.280 So it's my pleasure to give the

NOTE Confidence: 0.87841828125

00:01:23.280 --> 00:01:25.288 podium to Jeff and we look forward

NOTE Confidence: 0.87841828125

00:01:25.288 --> 00:01:26.913 to hearing your your presentation.

NOTE Confidence: 0.953467

00:01:29.150 --> 00:01:31.970 Thank you House for that wonderful

NOTE Confidence: 0.953467

00:01:31.970 --> 00:01:34.830 introduction and thank you and and

NOTE Confidence: 0.817638520909091

00:01:34.830 --> 00:01:37.632 and Ken for the encouragement to

NOTE Confidence: 0.817638520909091

00:01:37.632 --> 00:01:40.350 present today for this audience.

NOTE Confidence: 0.817638520909091

00:01:40.350 --> 00:01:42.975 And thank you all for basically the

NOTE Confidence: 0.817638520909091

00:01:42.975 --> 00:01:45.029 opportunity to present the kind of

NOTE Confidence: 0.817638520909091

00:01:45.029 --> 00:01:47.350 work that we've been doing in my lab.

NOTE Confidence: 0.817638520909091

00:01:47.350 --> 00:01:51.140 The title of. My talk is why me?

NOTE Confidence: 0.817638520909091

00:01:51.140 --> 00:01:53.435 The mutagenic origins of cancer

NOTE Confidence: 0.817638520909091

00:01:53.435 --> 00:01:55.730 for individual tumors and tumor

NOTE Confidence: 0.817638520909091

00:01:55.808 --> 00:01:58.230 types and I'm going to spend some

NOTE Confidence: 0.817638520909091

00:01:58.230 --> 00:01:59.632 time talking about that title.

NOTE Confidence: 0.817638520909091

00:01:59.632 --> 00:02:01.210 But first let me just go

NOTE Confidence: 0.817638520909091

00:02:01.268 --> 00:02:02.480 by my disclosure slide.

NOTE Confidence: 0.817638520909091

00:02:02.480 --> 00:02:05.476 I have done consulting for Black Diamond

NOTE Confidence: 0.817638520909091

00:02:05.476 --> 00:02:07.920 Therapeutics and Agios Pharmaceuticals.

NOTE Confidence: 0.817638520909091

00:02:07.920 --> 00:02:10.848 And so this title, why me?

NOTE Confidence: 0.817638520909091

00:02:10.850 --> 00:02:13.825 Just was inspired by the fact that

NOTE Confidence: 0.817638520909091

00:02:13.830 --> 00:02:16.266 as I started working on this work

NOTE Confidence: 0.817638520909091

00:02:16.270 --> 00:02:18.290 originally largely with Vincent Kantaro,

NOTE Confidence: 0.817638520909091

00:02:18.290 --> 00:02:20.208 who you'll see a picture of later,

00:02:20.210 --> 00:02:22.130 we realized that what we were doing to

00:02:22.130 --> 00:02:24.302 try to understand just what individual

00:02:24.302 --> 00:02:26.347 variants were contributing to cancer.

00:02:26.350 --> 00:02:28.102 Actually to some degree and the

00:02:28.102 --> 00:02:30.090 degree to which it addresses it,

00:02:30.090 --> 00:02:31.588 I'd love for you to think about,

00:02:31.590 --> 00:02:33.816 as I give this talk answers the

00:02:33.816 --> 00:02:36.129 question for an individual patient,

00:02:36.130 --> 00:02:38.110 what the causation of their individual.

00:02:38.110 --> 00:02:39.580 The answer was and I'll go through

00:02:39.580 --> 00:02:41.139 a lot of detail about that,

00:02:41.140 --> 00:02:43.384 but that that gets down to

00:02:43.384 --> 00:02:44.506 the mutagenic origins.

00:02:44.510 --> 00:02:46.560 Again not the physiological origins

00:02:46.560 --> 00:02:49.060 but mutagenic origins of cancer for

00:02:49.060 --> 00:02:50.995 individual tumors and tumor types.

NOTE Confidence: 0.817638520909091

00:02:51.000 --> 00:02:53.241 And I think this is a very it's obviously

NOTE Confidence: 0.817638520909091

00:02:53.241 --> 00:02:55.696 of interest to anyone who studies cancer,

NOTE Confidence: 0.817638520909091

00:02:55.700 --> 00:02:58.082 what the mutagenic origins of cancer

NOTE Confidence: 0.817638520909091

00:02:58.082 --> 00:03:00.809 are and certainly of interest in one

NOTE Confidence: 0.817638520909091

00:03:00.809 --> 00:03:03.294 way or another to patients who have

NOTE Confidence: 0.817638520909091

00:03:03.300 --> 00:03:06.100 have a have come down with cancer.

NOTE Confidence: 0.8764596788

00:03:08.300 --> 00:03:10.108 It has been widely reported that one of

NOTE Confidence: 0.8764596788

00:03:10.108 --> 00:03:12.009 the most difficult questions that patients

NOTE Confidence: 0.8764596788

00:03:12.009 --> 00:03:14.103 and doctors struggle with upon diagnosis

NOTE Confidence: 0.8764596788

00:03:14.153 --> 00:03:16.394 of cancer is the question, why me? Why?

NOTE Confidence: 0.8764596788

00:03:16.394 --> 00:03:18.760 Why was I struck with this ailment?

NOTE Confidence: 0.8764596788

00:03:18.760 --> 00:03:20.804 And it's natural for patients to want

NOTE Confidence: 0.8764596788

00:03:20.804 --> 00:03:22.976 to understand the causes behind their

NOTE Confidence: 0.8764596788

00:03:22.976 --> 00:03:25.652 calamities, and it's difficult to hear only

NOTE Confidence: 0.8764596788

00:03:25.652 --> 00:03:27.879 statistics and probabilities as a response.

00:03:27.880 --> 00:03:29.546 So the traditional way that you answer

00:03:29.546 --> 00:03:31.039 this question of why me is to say,

00:03:31.040 --> 00:03:33.008 well, did you smoke that elevates

00:03:33.008 --> 00:03:33.992 your process probability.

00:03:34.000 --> 00:03:36.094 Do you have this genetic predisposition

00:03:36.094 --> 00:03:37.490 that elevates your probability?

00:03:37.490 --> 00:03:40.773 Um, you know, how old are you?

00:03:40.773 --> 00:03:42.528 What is your ethnic background?

00:03:42.530 --> 00:03:44.015 There's lots of different predictors

00:03:44.015 --> 00:03:45.500 for whether or not someone

00:03:45.549 --> 00:03:46.869 might come down with cancer.

00:03:46.870 --> 00:03:48.070 But those aren't answers about why

00:03:48.070 --> 00:03:49.369 you came down with your cancer.

00:03:49.370 --> 00:03:51.390 Those are answers about

00:03:51.390 --> 00:03:53.410 generalizations about your life.

00:03:53.410 --> 00:03:54.718 So to date,

NOTE Confidence: 0.8764596788

00:03:54.718 --> 00:03:56.898 these statistics and probabilities are

NOTE Confidence: 0.8764596788

00:03:56.898 --> 00:03:59.391 nearly the only answer that science

NOTE Confidence: 0.8764596788

00:03:59.391 --> 00:04:02.140 and medicine has been able to give.

NOTE Confidence: 0.8764596788

00:04:02.140 --> 00:04:04.456 And one answer that's sort of

NOTE Confidence: 0.8764596788

00:04:04.456 --> 00:04:06.780 straightforward and obvious and if you

NOTE Confidence: 0.8764596788

00:04:06.780 --> 00:04:09.265 are proponent of sort of the genetic

NOTE Confidence: 0.8764596788

00:04:09.265 --> 00:04:11.151 evolutionary model of what makes

NOTE Confidence: 0.8764596788

00:04:11.151 --> 00:04:13.353 cancer happen is that mutations happen

NOTE Confidence: 0.8764596788

00:04:13.360 --> 00:04:15.418 and that's why you have your cancer.

NOTE Confidence: 0.8764596788

00:04:15.420 --> 00:04:16.580 It's a very general answer,

NOTE Confidence: 0.8764596788

00:04:16.580 --> 00:04:20.040 though it's not terribly satisfying,

NOTE Confidence: 0.8764596788

00:04:20.040 --> 00:04:21.328 but it can be broken down into a

NOTE Confidence: 0.8764596788

00:04:21.328 --> 00:04:22.888 lot of different kinds of mutations.

NOTE Confidence: 0.8764596788

00:04:22.890 --> 00:04:25.710 So there are clock like endogenous

NOTE Confidence: 0.8764596788

00:04:25.710 --> 00:04:28.134 mutations and processes that fuel

NOTE Confidence: 0.8764596788

00:04:28.134 --> 00:04:31.396 mutation throughout the body over a lifetime.

NOTE Confidence: 0.8764596788

00:04:31.400 --> 00:04:33.140 So as your body ages,

NOTE Confidence: 0.8764596788

00:04:33.140 --> 00:04:35.774 you get these mutations that happen

NOTE Confidence: 0.8764596788

00:04:35.774 --> 00:04:38.047 simply because the cellular processes

NOTE Confidence: 0.8764596788

00:04:38.047 --> 00:04:41.155 that reproduce your DNA are not perfectly

NOTE Confidence: 0.8764596788

00:04:41.160 --> 00:04:43.500 designed to reproduce it perfectly,

NOTE Confidence: 0.8764596788

00:04:43.500 --> 00:04:45.635 and they can't be just because of

NOTE Confidence: 0.8764596788

00:04:45.635 --> 00:04:47.664 the third law of thermodynamics.

NOTE Confidence: 0.8764596788

00:04:47.664 --> 00:04:49.388 So they're endogenous processes

NOTE Confidence: 0.8764596788

00:04:49.388 --> 00:04:51.553 that fuel mutation in your

NOTE Confidence: 0.8764596788

00:04:51.553 --> 00:04:53.205 body throughout your lifetime.

NOTE Confidence: 0.8764596788

00:04:53.210 --> 00:04:56.450 There are also mutational processes

NOTE Confidence: 0.8764596788

00:04:56.450 --> 00:04:58.448 that are fueled by exogenous sources,

NOTE Confidence: 0.8764596788

00:04:58.450 --> 00:05:00.094 such as viral infection

NOTE Confidence: 0.8764596788

00:05:00.094 --> 00:05:01.327 inducing applebach activity.

NOTE Confidence: 0.8764596788

00:05:01.330 --> 00:05:03.400 So viral infection can cause your

NOTE Confidence: 0.8764596788

00:05:03.400 --> 00:05:05.750 cell to react in certain ways,

NOTE Confidence: 0.8764596788

00:05:05.750 --> 00:05:07.514 maybe for cellular defense.

NOTE Confidence: 0.8764596788

00:05:07.514 --> 00:05:09.719 And those mutations that are

NOTE Confidence: 0.8764596788

00:05:09.719 --> 00:05:12.270 brought about as sort of secondary

NOTE Confidence: 0.8764596788

00:05:12.270 --> 00:05:14.788 consequences of your response to viral

NOTE Confidence: 0.8764596788

00:05:14.788 --> 00:05:17.203 infection can also lead to cancer.

NOTE Confidence: 0.8764596788

00:05:17.203 --> 00:05:19.549 And the third category is exogenous

NOTE Confidence: 0.8764596788

00:05:19.549 --> 00:05:21.471 mutagenic sources such as tobacco

NOTE Confidence: 0.8764596788

00:05:21.471 --> 00:05:23.661 smoke that may affect your lungs.

NOTE Confidence: 0.8764596788

00:05:23.670 --> 00:05:26.407 For your head and neck or UV

NOTE Confidence: 0.8764596788

00:05:26.407 --> 00:05:29.078 radiation that can affect your skin.

NOTE Confidence: 0.8764596788

00:05:29.080 --> 00:05:30.968 So these are all sources of mutation that

NOTE Confidence: 0.8764596788

00:05:30.968 --> 00:05:32.712 we know about and probabilistically we

NOTE Confidence: 0.8764596788

00:05:32.712 --> 00:05:34.830 can tell patients about the fact that,

NOTE Confidence: 0.8764596788

00:05:34.830 --> 00:05:35.220 you know,

NOTE Confidence: 0.8764596788

00:05:35.220 --> 00:05:36.585 exposing yourself to a lot of sun

NOTE Confidence: 0.8764596788

00:05:36.585 --> 00:05:38.158 may increase your risk for Melanoma.

NOTE Confidence: 0.8764596788

00:05:38.160 --> 00:05:39.040 If you have a Melanoma,

NOTE Confidence: 0.8764596788

00:05:39.040 --> 00:05:41.182 it may be partly due to the fact that

NOTE Confidence: 0.8764596788

00:05:41.182 --> 00:05:43.354 you exposed yourself to a lot of sun

NOTE Confidence: 0.8764596788

00:05:43.354 --> 00:05:45.557 light at some point during your life.

NOTE Confidence: 0.8764596788

00:05:45.560 --> 00:05:48.360 Now one of the I think quite

NOTE Confidence: 0.8764596788

00:05:48.360 --> 00:05:50.043 revolutionary sort of discoveries

NOTE Confidence: 0.8764596788

00:05:50.043 --> 00:05:53.243 of the of recent times was that we

NOTE Confidence: 0.8764596788

00:05:53.243 --> 00:05:56.127 can actually trace all a lot of

NOTE Confidence: 0.8764596788

00:05:56.127 --> 00:05:58.474 those sources I should say certainly

NOTE Confidence: 0.8764596788

00:05:58.474 --> 00:06:00.706 all the sources I just mentioned,

NOTE Confidence: 0.8764596788

00:06:00.710 --> 00:06:03.916 but many others as well to when we

NOTE Confidence: 0.8764596788

00:06:03.916 --> 00:06:06.660 when we sequence a tumor for instance

NOTE Confidence: 0.796650816111111

00:06:06.749 --> 00:06:10.326 we can trace signatures of those different

NOTE Confidence: 0.796650816111111

00:06:10.326 --> 00:06:12.700 sources in the DNA mutations that happen.

NOTE Confidence: 0.796650816111111

00:06:12.700 --> 00:06:15.500 So certain DNA mutations are more frequent.

NOTE Confidence: 0.796650816111111

00:06:15.500 --> 00:06:16.795 And I'll explain this in

NOTE Confidence: 0.796650816111111

00:06:16.795 --> 00:06:18.090 more detail a little later.

NOTE Confidence: 0.796650816111111

00:06:18.090 --> 00:06:20.205 Certain mutations are a little

NOTE Confidence: 0.796650816111111

00:06:20.205 --> 00:06:22.690 more frequent when you have a UV

NOTE Confidence: 0.796650816111111

00:06:22.690 --> 00:06:24.330 mutagenesis and other mutations are

NOTE Confidence: 0.796650816111111

00:06:24.330 --> 00:06:26.394 more frequent when you have just

NOTE Confidence: 0.796650816111111

00:06:26.394 --> 00:06:27.802 simple aging processes, etcetera.

NOTE Confidence: 0.796650816111111

00:06:27.802 --> 00:06:29.986 And it turns out that there are

NOTE Confidence: 0.796650816111111

00:06:29.986 --> 00:06:31.363 enough mutations in the typical

NOTE Confidence: 0.796650816111111

00:06:31.363 --> 00:06:33.150 tumor that you can do a sort of

NOTE Confidence: 0.796650816111111

00:06:33.150 --> 00:06:34.053 machine learning deconvolution.

NOTE Confidence: 0.796650816111111

00:06:34.053 --> 00:06:36.160 And I won't go into the detail

NOTE Confidence: 0.796650816111111

00:06:36.215 --> 00:06:37.713 about that to sort of figure out

NOTE Confidence: 0.796650816111111

00:06:37.713 --> 00:06:39.764 for a given tumor what were the

NOTE Confidence: 0.796650816111111

00:06:39.764 --> 00:06:41.124 different sources that contributed

NOTE Confidence: 0.796650816111111

00:06:41.124 --> 00:06:44.282 these mutations and this is really,

NOTE Confidence: 0.796650816111111

00:06:44.282 --> 00:06:45.518 really extraordinary.

NOTE Confidence: 0.796650816111111

00:06:45.520 --> 00:06:49.085 That we can figure that out the one and

NOTE Confidence: 0.796650816111111

00:06:49.085 --> 00:06:51.380 and just to give you a little bit more

NOTE Confidence: 0.796650816111111

00:06:51.447 --> 00:06:53.820 of a a sort of a a more detail on that.

NOTE Confidence: 0.796650816111111

00:06:53.820 --> 00:06:56.137 So here's S1 which is typically it's

NOTE Confidence: 0.796650816111111

00:06:56.137 --> 00:06:58.187 called emanation of five methyl cytosine

NOTE Confidence: 0.796650816111111

00:06:58.187 --> 00:07:00.504 and that's considered to be sort of

NOTE Confidence: 0.796650816111111

00:07:00.568 --> 00:07:02.662 an endogenous aging process that sort

NOTE Confidence: 0.796650816111111

00:07:02.662 --> 00:07:04.934 of occurs without any particular cause

NOTE Confidence: 0.796650816111111

00:07:04.934 --> 00:07:07.556 other than our other time passing

NOTE Confidence: 0.796650816111111

00:07:07.556 --> 00:07:10.400 for our body through development.

NOTE Confidence: 0.796650816111111

00:07:10.400 --> 00:07:12.542 S2 is is one of two signatures

NOTE Confidence: 0.796650816111111

00:07:12.542 --> 00:07:14.525 that we associate with apobec

NOTE Confidence: 0.796650816111111

13

00:07:14.525 --> 00:07:16.019 activity there's defective.

NOTE Confidence: 0.796650816111111

00:07:16.020 --> 00:07:17.650 From August recombination DNA repair,

NOTE Confidence: 0.796650816111111

00:07:17.650 --> 00:07:20.085 which may be mutation based

NOTE Confidence: 0.796650816111111

00:07:20.085 --> 00:07:21.546 and therefore endogenous,

NOTE Confidence: 0.796650816111111

00:07:21.550 --> 00:07:23.308 but related to a very specific

NOTE Confidence: 0.796650816111111

00:07:23.308 --> 00:07:24.871 process that might be treatable

NOTE Confidence: 0.796650816111111

00:07:24.871 --> 00:07:27.195 tobacco smoke which you can see of

NOTE Confidence: 0.796650816111111

00:07:27.195 --> 00:07:29.050 course largely affects lung cancer,

NOTE Confidence: 0.796650816111111

00:07:29.050 --> 00:07:34.126 but you can also see some for for liver,

NOTE Confidence: 0.796650816111111

00:07:34.130 --> 00:07:35.396 head etcetera, kidney,

NOTE Confidence: 0.796650816111111

00:07:35.396 --> 00:07:37.928 there's some other sources tobacco smoking.

NOTE Confidence: 0.796650816111111

00:07:37.930 --> 00:07:40.170 So an S5 which also you see

NOTE Confidence: 0.796650816111111

00:07:40.170 --> 00:07:41.570 is large circles here.

NOTE Confidence: 0.796650816111111

00:07:41.570 --> 00:07:43.875 That's another signature that has

NOTE Confidence: 0.796650816111111

00:07:43.875 --> 00:07:46.180 been traced essentially to aging.

NOTE Confidence: 0.796650816111111

00:07:46.180 --> 00:07:46.558 Processes.

14

NOTE Confidence: 0.796650816111111

00:07:46.558 --> 00:07:49.204 Although it's a little less well understood

NOTE Confidence: 0.796650816111111

00:07:49.204 --> 00:07:52.116 what the what the underlying basis of it is,

NOTE Confidence: 0.796650816111111

00:07:52.120 --> 00:07:55.600 it's very clear that age is highly correlated

NOTE Confidence: 0.796650816111111

00:07:55.600 --> 00:07:59.048 with the amount of SS5 mutation you get.

NOTE Confidence: 0.796650816111111

00:07:59.050 --> 00:07:59.824 Defective DNA,

NOTE Confidence: 0.796650816111111

00:07:59.824 --> 00:08:00.598 mismatch repair,

NOTE Confidence: 0.796650816111111

00:08:00.598 --> 00:08:01.759 ultraviolet light etcetera.

NOTE Confidence: 0.796650816111111

00:08:01.760 --> 00:08:03.454 And you can see these distribute themselves

NOTE Confidence: 0.796650816111111

00:08:03.454 --> 00:08:04.779 differently for different types of cancer.

NOTE Confidence: 0.796650816111111

00:08:04.780 --> 00:08:06.523 And so again this is very consistent

NOTE Confidence: 0.796650816111111

00:08:06.523 --> 00:08:08.189 what we knew already in a lot,

NOTE Confidence: 0.796650816111111

00:08:08.190 --> 00:08:09.654 very consistent with what we generally

NOTE Confidence: 0.796650816111111

00:08:09.654 --> 00:08:11.109 did which was say predictably like

NOTE Confidence: 0.796650816111111

00:08:11.109 --> 00:08:13.040 if you have a lot of exposure to sun,

NOTE Confidence: 0.796650816111111

00:08:13.040 --> 00:08:15.231 you're more likely to get UV exposure

NOTE Confidence: 0.796650816111111

15

00:08:15.231 --> 00:08:17.745 and that UV exposure then is more

NOTE Confidence: 0.796650816111111

00:08:17.745 --> 00:08:20.007 likely to translate to mutations that

NOTE Confidence: 0.796650816111111

00:08:20.010 --> 00:08:21.865 that may or may not cause Melanoma.

NOTE Confidence: 0.796650816111111

00:08:21.870 --> 00:08:24.158 But but once you have those mutations there,

NOTE Confidence: 0.796650816111111

00:08:24.160 --> 00:08:26.184 you know they may, they may cause that.

NOTE Confidence: 0.796650816111111

00:08:26.190 --> 00:08:28.384 So this is great we've we've got, we've got.

NOTE Confidence: 0.796650816111111

00:08:28.384 --> 00:08:29.444 The ability to see the,

NOTE Confidence: 0.796650816111111

00:08:29.450 --> 00:08:32.257 the sort of the trace or exposure

NOTE Confidence: 0.796650816111111

00:08:32.257 --> 00:08:34.890 in cells to these mutagens.

NOTE Confidence: 0.796650816111111

00:08:34.890 --> 00:08:35.250 The,

NOTE Confidence: 0.796650816111111

00:08:35.250 --> 00:08:37.770 the one thing that's missing though is,

NOTE Confidence: 0.796650816111111

00:08:37.770 --> 00:08:40.688 is that the extent to which each of

NOTE Confidence: 0.796650816111111

00:08:40.688 --> 00:08:42.568 those processes actually contribute to

NOTE Confidence: 0.796650816111111

00:08:42.568 --> 00:08:44.210 tumorigenesis still remains unknown.

NOTE Confidence: 0.796650816111111

00:08:44.210 --> 00:08:45.694 So we can look at what mutations

NOTE Confidence: 0.796650816111111

00:08:45.694 --> 00:08:46.730 are in the genome.

NOTE Confidence: 0.796650816111111

00:08:46.730 --> 00:08:48.539 But if I count up mutations in the genome,

NOTE Confidence: 0.796650816111111

00:08:48.540 --> 00:08:49.252 here's one, here's one,

NOTE Confidence: 0.796650816111111

00:08:49.252 --> 00:08:50.070 here's one, here's one.

NOTE Confidence: 0.796650816111111

00:08:50.070 --> 00:08:51.910 That doesn't tell me how much of those,

NOTE Confidence: 0.796650816111111

00:08:51.910 --> 00:08:53.542 each of those mutations are actually

NOTE Confidence: 0.796650816111111

00:08:53.542 --> 00:08:54.630 contributing to tumor Genesis.

NOTE Confidence: 0.8410616175

00:08:54.630 --> 00:08:56.709 In fact, most of those mutations are

NOTE Confidence: 0.8410616175

00:08:56.709 --> 00:08:58.399 not contributing to tumor Genesis.

NOTE Confidence: 0.8410616175

00:08:58.400 --> 00:09:00.325 And most analysis find that there's only

NOTE Confidence: 0.8410616175

00:09:00.325 --> 00:09:02.014 a few mutations that are contributing

NOTE Confidence: 0.8410616175

00:09:02.014 --> 00:09:03.904 at a significant level sort of at

NOTE Confidence: 0.8410616175

00:09:03.958 --> 00:09:05.678 this SNV single nucleotide variant,

NOTE Confidence: 0.8410616175

00:09:05.680 --> 00:09:09.898 level two to two tumor Genesis.

NOTE Confidence: 0.8410616175

00:09:09.900 --> 00:09:13.140 So we really need to have more in our,

NOTE Confidence: 0.8410616175

00:09:13.140 --> 00:09:15.384 you know, another tool in our

NOTE Confidence: 0.8410616175

00:09:15.384 --> 00:09:16.880 plate to figure out.

NOTE Confidence: 0.8410616175

00:09:16.880 --> 00:09:19.376 What the level each of these

NOTE Confidence: 0.8410616175

00:09:19.380 --> 00:09:20.868 endogenous and exogenous processes

NOTE Confidence: 0.8410616175

00:09:20.868 --> 00:09:23.100 are contributing to a given cancer,

NOTE Confidence: 0.8410616175

00:09:23.100 --> 00:09:24.330 and here's just a schematic

NOTE Confidence: 0.8410616175

00:09:24.330 --> 00:09:25.560 for this right you know,

NOTE Confidence: 0.8410616175

00:09:25.560 --> 00:09:27.516 so the mutation 1, mutation 2,

NOTE Confidence: 0.8410616175

00:09:27.520 --> 00:09:29.557 up to mutation N However many there

NOTE Confidence: 0.8410616175

00:09:29.557 --> 00:09:31.958 are that are really affecting cancer,

NOTE Confidence: 0.8410616175

00:09:31.960 --> 00:09:34.235 they can cause increased cellular

NOTE Confidence: 0.8410616175

00:09:34.235 --> 00:09:35.600 proliferation and survival.

NOTE Confidence: 0.8410616175

00:09:35.600 --> 00:09:39.209 And sunlight may be contributing to UV

NOTE Confidence: 0.8410616175

00:09:39.209 --> 00:09:41.063 radiation may be contributing to some

NOTE Confidence: 0.8410616175

00:09:41.063 --> 00:09:43.063 of those mutations more than others

NOTE Confidence: 0.8410616175

00:09:43.063 --> 00:09:45.103 because certain mutations are caused by

NOTE Confidence: 0.8410616175

00:09:45.103 --> 00:09:47.229 sunlight and other ones are not similarly.

NOTE Confidence: 0.8410616175

00:09:47.230 --> 00:09:48.718 Aging may contribute to some of

NOTE Confidence: 0.8410616175

00:09:48.718 --> 00:09:50.060 those mutations more than others.

NOTE Confidence: 0.8410616175

00:09:50.060 --> 00:09:52.307 And what I've got right here is,

NOTE Confidence: 0.8410616175

00:09:52.310 --> 00:09:52.924 you know,

NOTE Confidence: 0.8410616175

00:09:52.924 --> 00:09:55.380 if you take nothing else from this lecture,

NOTE Confidence: 0.8410616175

00:09:55.380 --> 00:09:56.668 this is the main thing that I

NOTE Confidence: 0.8410616175

00:09:56.668 --> 00:09:57.832 want to emphasize is that there's

NOTE Confidence: 0.8410616175

00:09:57.832 --> 00:09:58.960 sort of two stages to this.

NOTE Confidence: 0.8410616175

00:09:58.960 --> 00:10:00.148 One is, you know,

NOTE Confidence: 0.8410616175

00:10:00.148 --> 00:10:01.930 what mutagens have you been exposed

NOTE Confidence: 0.8410616175

00:10:01.987 --> 00:10:04.024 to and contributing to the set of

NOTE Confidence: 0.8410616175

00:10:04.024 --> 00:10:06.078 mutations that are causing your cancer?

NOTE Confidence: 0.8410616175

00:10:06.080 --> 00:10:08.360 And the 2nd is how much do each of those

NOTE Confidence: 0.8410616175

00:10:08.424 --> 00:10:10.764 mutations actually contribute to the

NOTE Confidence: 0.8410616175

00:10:10.764 --> 00:10:12.636 increased cellular proliferation and

NOTE Confidence: 0.8410616175

19

00:10:12.636 --> 00:10:15.000 survival that is the phenotype of cancer.

NOTE Confidence: 0.88202142375

00:10:17.260 --> 00:10:19.716 And there's a way to figure this out.

NOTE Confidence: 0.88202142375

00:10:19.720 --> 00:10:21.176 But to figure it out we we need

NOTE Confidence: 0.88202142375

00:10:21.176 --> 00:10:22.720 to sort of deconvolve something.

NOTE Confidence: 0.88202142375

00:10:22.720 --> 00:10:24.256 And this is an old idea and I'm

NOTE Confidence: 0.88202142375

00:10:24.256 --> 00:10:25.800 going to go through it in some

NOTE Confidence: 0.88202142375

00:10:25.800 --> 00:10:27.479 detail just to make sure that it's

NOTE Confidence: 0.88202142375

00:10:27.479 --> 00:10:29.403 clear to everyone that cancers are

NOTE Confidence: 0.88202142375

00:10:29.403 --> 00:10:31.308 the outcome of an evolutionary

NOTE Confidence: 0.88202142375

00:10:31.308 --> 00:10:33.299 process that's driven by mutation,

NOTE Confidence: 0.88202142375

00:10:33.300 --> 00:10:34.515 consequent genetic variation

NOTE Confidence: 0.88202142375

00:10:34.515 --> 00:10:36.135 created by that mutation,

NOTE Confidence: 0.88202142375

00:10:36.140 --> 00:10:37.900 and natural selection for

NOTE Confidence: 0.88202142375

00:10:37.900 --> 00:10:39.660 the more oncogenic variants.

NOTE Confidence: 0.88202142375

00:10:39.660 --> 00:10:42.250 This is from Peter Knowles

NOTE Confidence: 0.88202142375

00:10:42.250 --> 00:10:43.012 1976 science article,

NOTE Confidence: 0.88202142375

00:10:43.012 --> 00:10:44.790 a very well known article where he

NOTE Confidence: 0.88202142375

00:10:44.837 --> 00:10:46.289 just went through the idea that,

NOTE Confidence: 0.88202142375

00:10:46.290 --> 00:10:47.430 you know, it's an evolutionary.

NOTE Confidence: 0.88202142375

00:10:47.430 --> 00:10:50.086 Process that actually produces

NOTE Confidence: 0.88202142375

00:10:50.086 --> 00:10:53.406 malignancies and in this depiction

NOTE Confidence: 0.88202142375

00:10:53.406 --> 00:10:56.760 you can see a cellular lineages

NOTE Confidence: 0.88202142375

00:10:56.760 --> 00:10:59.340 differentiating and dividing.

NOTE Confidence: 0.88202142375

00:10:59.340 --> 00:11:01.332 You see a lot of lineages that are

NOTE Confidence: 0.88202142375

00:11:01.332 --> 00:11:03.247 hashed out here meaning they go

NOTE Confidence: 0.88202142375

00:11:03.247 --> 00:11:04.937 extinct and that's the selective

NOTE Confidence: 0.88202142375

00:11:04.937 --> 00:11:06.060 process in operation.

NOTE Confidence: 0.88202142375

00:11:06.060 --> 00:11:07.775 You know most of our our cells

NOTE Confidence: 0.88202142375

00:11:07.775 --> 00:11:10.101 are all dying at the same rate as

NOTE Confidence: 0.88202142375

00:11:10.101 --> 00:11:11.631 we're as they're dividing typically

NOTE Confidence: 0.88202142375

00:11:11.696 --> 00:11:12.696 as in as an adult.

NOTE Confidence: 0.88202142375

00:11:12.700 --> 00:11:15.139 So there's a lot of death going on we

NOTE Confidence: 0.88202142375

00:11:15.139 --> 00:11:17.436 don't usually emphasize that but but.

NOTE Confidence: 0.88202142375

00:11:17.440 --> 00:11:19.184 So that death may be going on and

NOTE Confidence: 0.88202142375

00:11:19.184 --> 00:11:20.786 what happens is that at some point

NOTE Confidence: 0.88202142375

00:11:20.786 --> 00:11:22.204 you get lineages that are reproducing

NOTE Confidence: 0.88202142375

00:11:22.204 --> 00:11:23.989 a lot more than they are dying.

NOTE Confidence: 0.88202142375

00:11:23.990 --> 00:11:25.367 And those ones,

NOTE Confidence: 0.88202142375

00:11:25.367 --> 00:11:28.121 in the case that they cause

NOTE Confidence: 0.88202142375

00:11:28.121 --> 00:11:30.155 difficulties for your life are

NOTE Confidence: 0.88202142375

00:11:30.155 --> 00:11:32.280 usually referred to as malignancies,

NOTE Confidence: 0.88202142375

00:11:32.280 --> 00:11:33.960 especially if they can then

NOTE Confidence: 0.88202142375

00:11:33.960 --> 00:11:35.304 migrate to other locations.

NOTE Confidence: 0.88202142375

00:11:35.310 --> 00:11:36.582 And this.

NOTE Confidence: 0.88202142375

00:11:36.582 --> 00:11:39.573 So these later evolved lineages are

NOTE Confidence: 0.88202142375

00:11:39.573 --> 00:11:41.739 usually the product of a series

NOTE Confidence: 0.88202142375

00:11:41.739 --> 00:11:44.082 of mutations that come along

22

NOTE Confidence: 0.88202142375

00:11:44.082 --> 00:11:45.990 during this evolutionary process

NOTE Confidence: 0.88202142375

00:11:45.990 --> 00:11:48.190 and what's happening with those.

NOTE Confidence: 0.88202142375

00:11:48.190 --> 00:11:51.070 Patience is they're actually enabling the

NOTE Confidence: 0.88202142375

00:11:51.070 --> 00:11:53.290 cells to survive and proliferate better,

NOTE Confidence: 0.88202142375

00:11:53.290 --> 00:11:55.180 so they're selected as the terminology

NOTE Confidence: 0.88202142375

00:11:55.232 --> 00:11:56.627 we use in evolutionary biology,

NOTE Confidence: 0.88202142375

00:11:56.630 --> 00:11:59.138 and they persist.

NOTE Confidence: 0.88202142375

00:11:59.140 --> 00:12:02.236 And that arising of those mutations

NOTE Confidence: 0.88202142375

00:12:02.236 --> 00:12:04.300 within individual cells within

NOTE Confidence: 0.88202142375

00:12:04.386 --> 00:12:06.605 cancer lineages is what we need to

NOTE Confidence: 0.88202142375

00:12:06.605 --> 00:12:07.650 sort of understand because there's

NOTE Confidence: 0.88202142375

00:12:07.694 --> 00:12:08.799 two processes going on here.

NOTE Confidence: 0.88202142375

00:12:08.800 --> 00:12:10.630 One is the appearance of these

NOTE Confidence: 0.88202142375

00:12:10.630 --> 00:12:12.620 mutations and then there's the amount

NOTE Confidence: 0.88202142375

00:12:12.620 --> 00:12:14.410 that they actually increase the

NOTE Confidence: 0.88202142375

00:12:14.410 --> 00:12:16.240 survival and replication of the cells.

NOTE Confidence: 0.88202142375

00:12:16.240 --> 00:12:18.700 So to quantify cancer effect size,

NOTE Confidence: 0.88202142375

00:12:18.700 --> 00:12:20.404 which is what I typically call

NOTE Confidence: 0.88202142375

00:12:20.404 --> 00:12:21.540 this the effect on,

NOTE Confidence: 0.88202142375

00:12:21.540 --> 00:12:22.316 you know,

NOTE Confidence: 0.88202142375

00:12:22.316 --> 00:12:25.032 on cells of actually leading to cancer,

NOTE Confidence: 0.88202142375

00:12:25.040 --> 00:12:27.320 which in evolutionary biology we

NOTE Confidence: 0.88202142375

00:12:27.320 --> 00:12:29.600 just call a selection coefficient.

NOTE Confidence: 0.88202142375

00:12:29.600 --> 00:12:31.472 We need to understand what the

NOTE Confidence: 0.88202142375

00:12:31.472 --> 00:12:32.720 prevalence in a population,

NOTE Confidence: 0.88202142375

00:12:32.720 --> 00:12:34.442 patient population is of a tumor

NOTE Confidence: 0.88202142375

00:12:34.442 --> 00:12:36.420 and we need to deconvolve that

NOTE Confidence: 0.88202142375

00:12:36.420 --> 00:12:38.340 prevalence into two factors because

NOTE Confidence: 0.88202142375

00:12:38.340 --> 00:12:40.531 when we see a certain mutation

NOTE Confidence: 0.88202142375

00:12:40.531 --> 00:12:42.792 very commonly in a kind of cancer,

NOTE Confidence: 0.88202142375

00:12:42.800 --> 00:12:44.380 that doesn't mean it's causing

NOTE Confidence: 0.88202142375

00:12:44.380 --> 00:12:45.960 a lot of the cancer.

NOTE Confidence: 0.88202142375

00:12:45.960 --> 00:12:47.479 It may just be that the mutation

NOTE Confidence: 0.88202142375

00:12:47.479 --> 00:12:49.324 rate is very high and we've seen

NOTE Confidence: 0.88202142375

00:12:49.324 --> 00:12:50.749 that very frequently in instances

NOTE Confidence: 0.88202142375

00:12:50.749 --> 00:12:52.418 where we have genes that are very

NOTE Confidence: 0.88202142375

00:12:52.418 --> 00:12:53.860 large or have very high mutation

NOTE Confidence: 0.88202142375

00:12:53.860 --> 00:12:55.235 rates that show up frequently

NOTE Confidence: 0.88202142375

00:12:55.235 --> 00:12:56.430 when we sequence tumors,

NOTE Confidence: 0.88202142375

00:12:56.430 --> 00:12:58.985 but are not significant causes of cancer.

NOTE Confidence: 0.88202142375

00:12:58.990 --> 00:13:00.726 And so we really need to understand,

NOTE Confidence: 0.88202142375

00:13:00.730 --> 00:13:01.182 you know,

NOTE Confidence: 0.88202142375

00:13:01.182 --> 00:13:02.312 which ones are actually contributing

NOTE Confidence: 0.88202142375

00:13:02.312 --> 00:13:03.492 cancer and which ones are

NOTE Confidence: 0.88202142375

00:13:03.492 --> 00:13:04.412 just typically contributing to

NOTE Confidence: 0.88202142375

00:13:04.412 --> 00:13:05.332 prevalence because of an

NOTE Confidence: 0.765286207826087

00:13:05.371 --> 00:13:06.268 underlying mutation rate.

NOTE Confidence: 0.765286207826087

00:13:06.270 --> 00:13:08.629 So to quantify the cancer effective size,

NOTE Confidence: 0.765286207826087

00:13:08.630 --> 00:13:10.510 we have to do a fairly straightforward thing,

NOTE Confidence: 0.765286207826087

00:13:10.510 --> 00:13:11.630 which is take that prevalence,

NOTE Confidence: 0.765286207826087

00:13:11.630 --> 00:13:13.304 how frequent we see them in

NOTE Confidence: 0.765286207826087

00:13:13.304 --> 00:13:14.806 patients and deconvolve it into

NOTE Confidence: 0.765286207826087

00:13:14.806 --> 00:13:16.130 the baseline mutation rate.

NOTE Confidence: 0.765286207826087

00:13:16.130 --> 00:13:18.345 How frequently the mutations are

NOTE Confidence: 0.765286207826087

00:13:18.345 --> 00:13:21.025 occurring in the lineage and into

NOTE Confidence: 0.765286207826087

00:13:21.025 --> 00:13:23.335 the degree of selection for that

NOTE Confidence: 0.765286207826087

00:13:23.335 --> 00:13:25.659 mutation in the cancer lineage.

NOTE Confidence: 0.765286207826087

00:13:25.660 --> 00:13:27.660 And if we can differentiate those two things,

NOTE Confidence: 0.765286207826087

00:13:27.660 --> 00:13:29.214 then we can better understand how

NOTE Confidence: 0.765286207826087

00:13:29.214 --> 00:13:31.126 much is because how much is that

NOTE Confidence: 0.765286207826087

00:13:31.126 --> 00:13:32.770 mutation there is because of the

NOTE Confidence: 0.765286207826087

00:13:32.770 --> 00:13:34.307 underlying mutations that are happening

26

00:13:34.307 --> 00:13:36.125 and across your whole genome that

00:13:36.130 --> 00:13:37.455 aren't necessarily relevant and how

00:13:37.455 --> 00:13:39.204 much is due to those individual

00:13:39.204 --> 00:13:40.684 mutations actually increasing the

00:13:40.684 --> 00:13:42.820 proliferation and survival of the cell.

00:13:42.820 --> 00:13:44.738 So here's just a schematic of that.

00:13:44.740 --> 00:13:48.010 This is just basic evolutionary biology.

00:13:48.010 --> 00:13:49.874 one-on-one mutation creates variation

00:13:49.874 --> 00:13:52.204 symbolized by the different shades

00:13:52.204 --> 00:13:54.650 of Gray there unfavorable mutations

00:13:54.650 --> 00:13:56.087 are selected against.

00:13:56.090 --> 00:13:57.978 Reproduction and mutation occur,

00:13:57.978 --> 00:14:00.810 and the favorable mutations are more

00:14:00.880 --> 00:14:03.070 likely to survive and reproduce,

00:14:03.070 --> 00:14:05.494 and the point of this is that it

00:14:05.494 --> 00:14:07.835 both the mutation rate and the

00:14:07.835 --> 00:14:09.905 extent to which they contribute

00:14:09.905 --> 00:14:13.420 to survival and reproduction.

00:14:13.420 --> 00:14:16.028 Contribute to what you see at as an

00:14:16.028 --> 00:14:19.249 end product of the process of cellular

00:14:19.249 --> 00:14:22.489 differentiation, especially into cancers.

00:14:22.490 --> 00:14:22.848 All right.

00:14:22.848 --> 00:14:23.922 So how do we figure out

00:14:23.922 --> 00:14:24.850 that baseline mutation rate?

00:14:24.850 --> 00:14:27.706 Well, it's already been done for me anyway.

00:14:27.710 --> 00:14:29.480 It was a lot of the work was already done,

00:14:29.480 --> 00:14:32.570 which is really great by by

00:14:32.570 --> 00:14:34.630 Lawrence and and others.

00:14:34.630 --> 00:14:36.415 This is a 2013 paper quite a

00:14:36.415 --> 00:14:38.325 while ago where they showed that

00:14:38.325 --> 00:14:40.125 mutation rate varies widely across

NOTE Confidence: 0.765286207826087

00:14:40.125 --> 00:14:42.507 the genome and correlates with DNA

NOTE Confidence: 0.765286207826087

00:14:42.507 --> 00:14:44.507 replication time and expression level.

NOTE Confidence: 0.765286207826087

00:14:44.510 --> 00:14:45.358 So there's these covariates.

NOTE Confidence: 0.765286207826087

00:14:45.358 --> 00:14:47.037 I'm not going to go into a lot

NOTE Confidence: 0.765286207826087

00:14:47.037 --> 00:14:47.757 of detail about this.

NOTE Confidence: 0.765286207826087

00:14:47.760 --> 00:14:49.585 I've talked about this before

NOTE Confidence: 0.765286207826087

00:14:49.585 --> 00:14:51.498 with various audiences here, but.

NOTE Confidence: 0.765286207826087

00:14:51.498 --> 00:14:54.186 That mutation rate varies and correlates

NOTE Confidence: 0.765286207826087

00:14:54.186 --> 00:14:56.450 with DNA replication time and

NOTE Confidence: 0.765286207826087

00:14:56.450 --> 00:14:58.272 expression level with heterochromatin marks.

NOTE Confidence: 0.765286207826087

00:14:58.272 --> 00:15:00.942 A bunch of other correlates that we can

NOTE Confidence: 0.765286207826087

00:15:00.942 --> 00:15:02.697 actually get about individual tumors.

NOTE Confidence: 0.765286207826087

00:15:02.700 --> 00:15:04.416 Those allow us to ask questions

NOTE Confidence: 0.765286207826087

00:15:04.416 --> 00:15:06.430 about you know a given gene and

NOTE Confidence: 0.765286207826087

00:15:06.430 --> 00:15:08.685 whether or not it's got a very high

NOTE Confidence: 0.765286207826087

29

00:15:08.685 --> 00:15:10.876 mutation rate or a low mutation rate.

NOTE Confidence: 0.765286207826087

00:15:10.880 --> 00:15:13.169 By using those correlates to help us

NOTE Confidence: 0.765286207826087

00:15:13.169 --> 00:15:14.906 predict that along with synonymous

NOTE Confidence: 0.765286207826087

00:15:14.906 --> 00:15:17.321 changes in the genome which we can

NOTE Confidence: 0.765286207826087

00:15:17.321 --> 00:15:19.565 presume don't have any effect on the

NOTE Confidence: 0.765286207826087

00:15:19.565 --> 00:15:21.995 proliferation and survival of of cells.

NOTE Confidence: 0.765286207826087

00:15:22.000 --> 00:15:22.996 So for instance,

NOTE Confidence: 0.765286207826087

00:15:22.996 --> 00:15:23.660 olfactory receptors,

NOTE Confidence: 0.765286207826087

00:15:23.660 --> 00:15:25.612 which early on were this bugaboo that would

NOTE Confidence: 0.765286207826087

00:15:25.612 --> 00:15:27.826 show up when we did these tumor sequencing,

NOTE Confidence: 0.765286207826087

00:15:27.830 --> 00:15:28.964 happened to be in a part of

NOTE Confidence: 0.765286207826087

00:15:28.964 --> 00:15:29.949 the genome that gets a very,

NOTE Confidence: 0.765286207826087

00:15:29.950 --> 00:15:31.326 very high mutation rate.

NOTE Confidence: 0.765286207826087

00:15:31.326 --> 00:15:32.702 It's basically not expressed

NOTE Confidence: 0.765286207826087

00:15:32.702 --> 00:15:34.020 and not expressed.

NOTE Confidence: 0.765286207826087

00:15:34.020 --> 00:15:35.448 Parts of the genome don't have

NOTE Confidence: 0.765286207826087

00:15:35.448 --> 00:15:36.162 transcription enabled repair,

NOTE Confidence: 0.765286207826087

00:15:36.170 --> 00:15:36.548 etcetera.

NOTE Confidence: 0.765286207826087

00:15:36.548 --> 00:15:38.816 CSMD 3 is another example where

NOTE Confidence: 0.765286207826087

00:15:38.816 --> 00:15:40.470 there's very high levels,

NOTE Confidence: 0.765286207826087

00:15:40.470 --> 00:15:41.845 high correlates and also very

NOTE Confidence: 0.765286207826087

00:15:41.845 --> 00:15:42.670 high mutation rate.

NOTE Confidence: 0.765286207826087

00:15:42.670 --> 00:15:43.985 And typically it's not considered

NOTE Confidence: 0.765286207826087

00:15:43.985 --> 00:15:46.100 to be a driver even though you see

NOTE Confidence: 0.765286207826087

00:15:46.100 --> 00:15:47.836 it a lot in cancer tumor sequencing

NOTE Confidence: 0.765286207826087

00:15:47.892 --> 00:15:49.290 and you can do regressions on

NOTE Confidence: 0.765286207826087

00:15:49.290 --> 00:15:51.142 this and then I'm just going to

NOTE Confidence: 0.765286207826087

00:15:51.142 --> 00:15:52.358 very quickly mention that.

NOTE Confidence: 0.841321872857143

00:15:52.360 --> 00:15:54.299 This wonderful work was done by Lawrence,

NOTE Confidence: 0.841321872857143

00:15:54.300 --> 00:15:56.575 but typically that work was only applied

NOTE Confidence: 0.841321872857143

00:15:56.575 --> 00:15:59.078 to the question of whether or not

NOTE Confidence: 0.841321872857143

00:15:59.078 --> 00:16:00.913 genes were overburdened with mutations.

NOTE Confidence: 0.841321872857143

00:16:00.920 --> 00:16:01.984 So in other words, they got these

NOTE Confidence: 0.841321872857143

00:16:01.984 --> 00:16:02.909 mutation rates and they just said,

NOTE Confidence: 0.841321872857143

00:16:02.910 --> 00:16:04.709 well, is it more than we expect.

NOTE Confidence: 0.841321872857143

00:16:04.710 --> 00:16:06.285 And then they calculated P value for

NOTE Confidence: 0.841321872857143

00:16:06.285 --> 00:16:07.985 whether we should put this gene in the

NOTE Confidence: 0.841321872857143

00:16:07.985 --> 00:16:09.457 category of mutated or not and then

NOTE Confidence: 0.841321872857143

00:16:09.457 --> 00:16:10.807 they leave behind that mutation rate

NOTE Confidence: 0.841321872857143

00:16:10.807 --> 00:16:12.525 and then just look at prevalence in

NOTE Confidence: 0.841321872857143

00:16:12.525 --> 00:16:14.301 most of the analysis that were done

NOTE Confidence: 0.841321872857143

00:16:14.301 --> 00:16:18.153 from 2013 through 2018 or so. So.

NOTE Confidence: 0.841321872857143

00:16:18.153 --> 00:16:20.204 So typically that was sort of left

NOTE Confidence: 0.841321872857143

00:16:20.204 --> 00:16:22.497 behind at that point and that's what.

NOTE Confidence: 0.841321872857143

00:16:22.500 --> 00:16:24.012 Vincent Kintaro and I in 2018 sort

NOTE Confidence: 0.841321872857143

00:16:24.012 --> 00:16:25.488 of picked up on and said look,

NOTE Confidence: 0.841321872857143

00:16:25.490 --> 00:16:27.038 this mutation rate is more important

NOTE Confidence: 0.841321872857143

00:16:27.038 --> 00:16:28.489 than for just calculating P values.

NOTE Confidence: 0.841321872857143

00:16:28.490 --> 00:16:30.010 It's actually important for

NOTE Confidence: 0.841321872857143

00:16:30.010 --> 00:16:31.150 calculating the effect.

NOTE Confidence: 0.841321872857143

00:16:31.150 --> 00:16:33.226 You know in the biostatisticians mind

NOTE Confidence: 0.841321872857143

00:16:33.226 --> 00:16:36.050 P value is sort of a secondary thing.

NOTE Confidence: 0.841321872857143

00:16:36.050 --> 00:16:37.472 First you calculate the effect of

NOTE Confidence: 0.841321872857143

00:16:37.472 --> 00:16:38.870 the thing you're looking at and

NOTE Confidence: 0.841321872857143

00:16:38.870 --> 00:16:40.094 then you calculate that you see

NOTE Confidence: 0.841321872857143

00:16:40.094 --> 00:16:41.568 whether you should trust that effect.

NOTE Confidence: 0.841321872857143

00:16:41.570 --> 00:16:43.766 And so that's what Vincent Cantara

NOTE Confidence: 0.841321872857143

00:16:43.766 --> 00:16:46.424 and I did and just here's a sort of a

NOTE Confidence: 0.841321872857143

00:16:46.424 --> 00:16:48.137 brief introduction to how we do that

NOTE Confidence: 0.841321872857143

00:16:48.137 --> 00:16:49.382 calculation by convolving the gene

NOTE Confidence: 0.841321872857143

00:16:49.382 --> 00:16:51.370 based rates from the silent sites and

NOTE Confidence: 0.841321872857143

00:16:51.370 --> 00:16:53.020 covariates with they're trying to die.

NOTE Confidence: 0.841321872857143

00:16:53.020 --> 00:16:53.280 Context.

NOTE Confidence: 0.841321872857143

00:16:53.280 --> 00:16:55.100 So you can just go through tumor

NOTE Confidence: 0.841321872857143

00:16:55.157 --> 00:16:57.124 sequence data and you can look at

NOTE Confidence: 0.841321872857143

00:16:57.124 --> 00:16:58.684 what the underlying mutation rate

NOTE Confidence: 0.841321872857143

00:16:58.684 --> 00:17:00.658 is using basically that Lawrence at

NOTE Confidence: 0.841321872857143

00:17:00.658 --> 00:17:02.428 all approach that I talked about

NOTE Confidence: 0.841321872857143

00:17:02.428 --> 00:17:03.259 with the covariance,

NOTE Confidence: 0.841321872857143

00:17:03.260 --> 00:17:04.420 you can sort of look at every gene

NOTE Confidence: 0.841321872857143

00:17:04.420 --> 00:17:05.583 in the genome and calculate what

NOTE Confidence: 0.841321872857143

00:17:05.583 --> 00:17:06.415 the mutation rate is.

NOTE Confidence: 0.841321872857143

00:17:06.420 --> 00:17:08.060 And this is just one of these plots

NOTE Confidence: 0.841321872857143

00:17:08.060 --> 00:17:09.638 that's just scatter plot on one axis

NOTE Confidence: 0.841321872857143

00:17:09.638 --> 00:17:11.240 of what the different gene rates are.

NOTE Confidence: 0.841321872857143

00:17:11.240 --> 00:17:13.074 And you can see there's quite a

NOTE Confidence: 0.841321872857143

00:17:13.074 --> 00:17:13.860 wide range here.

NOTE Confidence: 0.841321872857143

00:17:13.860 --> 00:17:15.729 And I think that's the most important

NOTE Confidence: 0.841321872857143

00:17:15.729 --> 00:17:17.745 lesson of this little image is that

NOTE Confidence: 0.841321872857143

00:17:17.745 --> 00:17:19.200 the mutation rate varies quite

NOTE Confidence: 0.841321872857143

00:17:19.200 --> 00:17:21.314 extensively from gene to gene from 10

NOTE Confidence: 0.841321872857143

00:17:21.314 --> 00:17:23.269 to the minus two to 10 to the minus 4.

NOTE Confidence: 0.841321872857143

00:17:23.270 --> 00:17:24.566 In this particular instance,

NOTE Confidence: 0.841321872857143

00:17:24.566 --> 00:17:26.510 so that's two orders of magnitude

NOTE Confidence: 0.841321872857143

00:17:26.564 --> 00:17:28.289 rate variation in mutation rates.

NOTE Confidence: 0.841321872857143

00:17:28.290 --> 00:17:29.450 So when you see,

NOTE Confidence: 0.841321872857143

00:17:29.450 --> 00:17:31.564 you know one gene mutated in a

NOTE Confidence: 0.841321872857143

00:17:31.564 --> 00:17:33.388 cancer tumor pop cohort at 100,

NOTE Confidence: 0.841321872857143

00:17:33.390 --> 00:17:33.904 you know,

NOTE Confidence: 0.841321872857143

00:17:33.904 --> 00:17:35.703 100 copies out of 1000 and another

NOTE Confidence: 0.841321872857143

00:17:35.703 --> 00:17:36.839 at 10 out of 1000,

NOTE Confidence: 0.841321872857143

00:17:36.840 --> 00:17:38.544 that's only one order of magnitude

NOTE Confidence: 0.841321872857143

00:17:38.544 --> 00:17:39.955 difference in prevalence and you

NOTE Confidence: 0.841321872857143

00:17:39.955 --> 00:17:41.305 can explain that by just half

NOTE Confidence: 0.841321872857143

00:17:41.305 --> 00:17:42.619 of this mutation rate diagram.

NOTE Confidence: 0.841321872857143

00:17:42.620 --> 00:17:43.487 In other words,

NOTE Confidence: 0.841321872857143

00:17:43.487 --> 00:17:45.510 mutation rate can explain a lot of

NOTE Confidence: 0.841321872857143

00:17:45.572 --> 00:17:47.516 the differences in how prevalent genes

NOTE Confidence: 0.841321872857143

00:17:47.516 --> 00:17:50.259 are when you look in a patient population.

NOTE Confidence: 0.841321872857143

00:17:50.260 --> 00:17:52.065 So you shouldn't take that

NOTE Confidence: 0.841321872857143

00:17:52.065 --> 00:17:53.509 prevalence as an indicator.

NOTE Confidence: 0.841321872857143

00:17:53.510 --> 00:17:55.772 As a strong indicator of how

NOTE Confidence: 0.841321872857143

00:17:55.772 --> 00:17:58.469 important a gene is in the cancer,

NOTE Confidence: 0.841321872857143

00:17:58.470 --> 00:18:01.098 you really need to basically understand

NOTE Confidence: 0.841321872857143

00:18:01.098 --> 00:18:03.429 this underlying mutation rate as well.

NOTE Confidence: 0.841321872857143

00:18:03.430 --> 00:18:06.027 And so then you can take different

NOTE Confidence: 0.841321872857143

00:18:06.027 --> 00:18:08.739 genes that are on that that diagram

NOTE Confidence: 0.841321872857143

00:18:08.740 --> 00:18:11.008 and you can look at each individual

NOTE Confidence: 0.841321872857143

00:18:11.008 --> 00:18:13.518 tumor and you can map out what

NOTE Confidence: 0.841321872857143

00:18:13.518 --> 00:18:15.353 the trinucleotide rate rates are.

NOTE Confidence: 0.817349678

00:18:15.360 --> 00:18:16.340 So. So this rate is,

NOTE Confidence: 0.817349678

00:18:16.340 --> 00:18:18.300 the rate above is just the rate at

NOTE Confidence: 0.817349678

00:18:18.300 --> 00:18:20.239 which the gene itself gets mutated.

NOTE Confidence: 0.817349678

00:18:20.240 --> 00:18:22.600 But if we want to look at every given site,

NOTE Confidence: 0.817349678

00:18:22.600 --> 00:18:24.814 the important thing is that the

NOTE Confidence: 0.817349678

00:18:24.814 --> 00:18:26.671 different mutational processes that I

NOTE Confidence: 0.817349678

00:18:26.671 --> 00:18:28.591 mentioned earlier in this talk affect

NOTE Confidence: 0.817349678

00:18:28.591 --> 00:18:30.820 different sites at different frequencies.

NOTE Confidence: 0.817349678

00:18:30.820 --> 00:18:31.700 Have a question right there.

NOTE Confidence: 0.4676921275

00:18:33.790 --> 00:18:36.910 2nd normalized for length.

NOTE Confidence: 0.4676921275

00:18:36.910 --> 00:18:39.400 Is the mutation rate itself?

NOTE Confidence: 0.4676921275

00:18:39.400 --> 00:18:42.780 In this case it is, yes.

NOTE Confidence: 0.4676921275

00:18:42.780 --> 00:18:46.495 So, so, so these different mutational

NOTE Confidence: 0.4676921275

00:18:46.495 --> 00:18:48.539 processes contribute to differently.

NOTE Confidence: 0.4676921275

00:18:48.540 --> 00:18:49.288 So in this case,

NOTE Confidence: 0.4676921275

00:18:49.288 --> 00:18:50.223 I'm looking at lung cancer,

NOTE Confidence: 0.4676921275

00:18:50.230 --> 00:18:52.351 which is why we can be carriers

NOTE Confidence: 0.4676921275

00:18:52.351 --> 00:18:53.620 and EGFR highlighted here.

NOTE Confidence: 0.4676921275

00:18:53.620 --> 00:18:54.704 And in lung cancer,

NOTE Confidence: 0.4676921275

00:18:54.704 --> 00:18:57.053 you get a lot of these RCA mutations

NOTE Confidence: 0.4676921275

00:18:57.053 --> 00:18:59.356 that are preceded by a T and

NOTE Confidence: 0.4676921275

00:18:59.356 --> 00:19:01.647 followed by an A and also ones that

NOTE Confidence: 0.4676921275

00:19:01.647 --> 00:19:03.057 are preceded by and followed by

NOTE Confidence: 0.4676921275

00:19:03.057 --> 00:19:05.080 an A&amp;C and an A and an A and an A.

NOTE Confidence: 0.4676921275

00:19:05.080 --> 00:19:08.120 So, so, so all of these bright red

NOTE Confidence: 0.4676921275

00:19:08.120 --> 00:19:09.568 trinucleotide context get much

NOTE Confidence: 0.4676921275

00:19:09.568 --> 00:19:11.378 more mutation than other ones.

NOTE Confidence: 0.4676921275

00:19:11.380 --> 00:19:12.886 And again I just want to

NOTE Confidence: 0.4676921275

00:19:12.886 --> 00:19:13.890 emphasize that the coloration.

NOTE Confidence: 0.4676921275

00:19:13.890 --> 00:19:16.473 Here is scaled to how often we see it.

NOTE Confidence: 0.4676921275

00:19:16.480 --> 00:19:18.838 And so you see almost an order of magnitude,

NOTE Confidence: 0.4676921275

00:19:18.840 --> 00:19:19.670 sometimes more,

NOTE Confidence: 0.4676921275

00:19:19.670 --> 00:19:22.160 with some cancer types of variation,

NOTE Confidence: 0.4676921275

00:19:22.160 --> 00:19:25.110 again in how frequently given

NOTE Confidence: 0.4676921275

00:19:25.110 --> 00:19:27.260 sites get mutated over other sites.

NOTE Confidence: 0.4676921275

00:19:27.260 --> 00:19:28.500 So when you combine this

NOTE Confidence: 0.4676921275

00:19:28.500 --> 00:19:30.120 plus the gene by gene rates,

NOTE Confidence: 0.4676921275

00:19:30.120 --> 00:19:32.437 you're talking about 3 orders of magnitude,

NOTE Confidence: 0.4676921275

00:19:32.440 --> 00:19:34.336 maybe even four in some cases,

NOTE Confidence: 0.4676921275

00:19:34.340 --> 00:19:35.570 between a given site and

NOTE Confidence: 0.4676921275

00:19:35.570 --> 00:19:36.800 another site in the genome,

NOTE Confidence: 0.4676921275

00:19:36.800 --> 00:19:37.970 and how frequently gets mutated.

NOTE Confidence: 0.4676921275

00:19:37.970 --> 00:19:39.386 So this is a really important

NOTE Confidence: 0.4676921275

00:19:39.386 --> 00:19:40.670 factor to take into consideration

NOTE Confidence: 0.4676921275

00:19:40.670 --> 00:19:42.095 when wondering whether or not

NOTE Confidence: 0.4676921275

00:19:42.095 --> 00:19:43.960 a given site is important for.

NOTE Confidence: 0.4676921275

00:19:43.960 --> 00:19:45.328 Driving cancer and what you can

NOTE Confidence: 0.4676921275

00:19:45.328 --> 00:19:47.047 do is you can basically tape this

NOTE Confidence: 0.4676921275

00:19:47.047 --> 00:19:48.790 map and look at each gene and

NOTE Confidence: 0.4676921275

00:19:48.843 --> 00:19:49.859 here I've just look,

NOTE Confidence: 0.4676921275

00:19:49.860 --> 00:19:52.028 I'm looking at like an excerpt of a

NOTE Confidence: 0.4676921275

00:19:52.028 --> 00:19:54.130 tiny little part of the of the genome.

NOTE Confidence: 0.4676921275

00:19:54.130 --> 00:19:54.551 Sorry.

NOTE Confidence: 0.4676921275

00:19:54.551 --> 00:19:58.950 This is this is site 850 to 870 and EGFR,

NOTE Confidence: 0.4676921275

00:19:58.950 --> 00:20:02.560 here's site 1 to 20 in K Ras and here's

NOTE Confidence: 0.4676921275

00:20:02.661 --> 00:20:05.629 site 30 to 50 in cutting and B1.

NOTE Confidence: 0.4676921275

00:20:05.630 --> 00:20:07.166 And I just want to mention that if

NOTE Confidence: 0.4676921275

00:20:07.166 --> 00:20:08.693 you you know you take these rates

NOTE Confidence: 0.4676921275

00:20:08.693 --> 00:20:10.425 and then you make sure that the

NOTE Confidence: 0.4676921275

00:20:10.425 --> 00:20:11.885 individual site rates are accommodated

NOTE Confidence: 0.4676921275

00:20:11.885 --> 00:20:14.990 by ensuring that you know TCA is much more.

NOTE Confidence: 0.4676921275

00:20:14.990 --> 00:20:18.986 Frequent then see CCG chaining 2

NOTE Confidence: 0.4676921275

00:20:18.986 --> 00:20:22.432 and A and and and do all of the math

NOTE Confidence: 0.4676921275

00:20:22.432 --> 00:20:23.694 that's very straightforward here

NOTE Confidence: 0.4676921275

00:20:23.694 --> 00:20:26.358 but a bit of a lot of accounting

NOTE Confidence: 0.4676921275

00:20:26.428 --> 00:20:28.618 bioinformatics ally and then map it

NOTE Confidence: 0.4676921275

00:20:28.618 --> 00:20:30.666 through the the actual genetic code.

NOTE Confidence: 0.4676921275

00:20:30.666 --> 00:20:32.424 So you're looking at every single

NOTE Confidence: 0.4676921275

00:20:32.424 --> 00:20:34.429 site in that protein and saying well

NOTE Confidence: 0.4676921275

00:20:34.429 --> 00:20:36.398 how likely is this 850 histidine to

NOTE Confidence: 0.4676921275

00:20:36.398 --> 00:20:38.337 change based on its three code on

NOTE Confidence: 0.4676921275

00:20:38.337 --> 00:20:40.552 sites into a tyrosine or a proline

NOTE Confidence: 0.4676921275

00:20:40.552 --> 00:20:42.420 or a phenylalanine etcetera etcetera.

NOTE Confidence: 0.4676921275

00:20:42.420 --> 00:20:44.088 And some sites of course some

NOTE Confidence: 0.4676921275

00:20:44.088 --> 00:20:45.578 changes of course can't really

NOTE Confidence: 0.4676921275

00:20:45.578 --> 00:20:46.878 happen through a single.

NOTE Confidence: 0.4676921275

00:20:46.880 --> 00:20:47.744 Nucleotide mutation,

00:20:47.744 --> 00:20:50.098 others can in multiple ways, etcetera.

00:20:50.098 --> 00:20:52.030 So there's a lot of addition to add up here.

00:20:52.030 --> 00:20:52.968 But once you add it all up,

00:20:52.970 --> 00:20:55.034 this diagram tells you how likely

00:20:55.034 --> 00:20:57.182 each different change is to happen

00:20:57.182 --> 00:20:58.250 by neutral mutation.

00:20:58.250 --> 00:20:59.811 That is when we just expect new

00:20:59.811 --> 00:21:01.355 mutations to be sprayed on there

00:21:01.355 --> 00:21:03.005 and have no difference in the

00:21:03.005 --> 00:21:03.950 replication and survival.

00:21:03.950 --> 00:21:06.094 So then we get this diagram of how

00:21:06.094 --> 00:21:08.176 much each amino acid position would

00:21:08.176 --> 00:21:10.402 be expected to be mutation mutated,

00:21:10.410 --> 00:21:12.096 and then we can compare that

00:21:12.096 --> 00:21:13.220 to what's actually observed.

NOTE Confidence: 0.700787762

00:21:18.480 --> 00:21:21.774 Um, what's actually observed is much,

NOTE Confidence: 0.700787762

00:21:21.774 --> 00:21:23.658 much more rarified set of mutations

NOTE Confidence: 0.700787762

00:21:23.658 --> 00:21:25.480 than what you actually expect

NOTE Confidence: 0.700787762

00:21:25.480 --> 00:21:27.116 based on neutral evolution.

NOTE Confidence: 0.700787762

00:21:27.120 --> 00:21:28.618 And that's because when we sample tumors,

NOTE Confidence: 0.700787762

00:21:28.620 --> 00:21:30.504 we're sampling tumors that have been

NOTE Confidence: 0.700787762

00:21:30.504 --> 00:21:32.160 under selection for very specific

NOTE Confidence: 0.700787762

00:21:32.160 --> 00:21:34.200 mutations and because right here I've

NOTE Confidence: 0.700787762

00:21:34.200 --> 00:21:36.582 selected sites that actually do have an

NOTE Confidence: 0.700787762

00:21:36.582 --> 00:21:38.212 effect on proliferation and survival.

NOTE Confidence: 0.700787762

00:21:38.220 --> 00:21:42.060 So on the top EGFR 858,

NOTE Confidence: 0.700787762

00:21:42.060 --> 00:21:45.020 Lucine is a very well known mutational site.

NOTE Confidence: 0.700787762

00:21:45.020 --> 00:21:47.700 The KSG 12 is also a very well known one.

NOTE Confidence: 0.700787762

00:21:47.700 --> 00:21:48.864 And then this.

NOTE Confidence: 0.700787762

00:21:48.864 --> 00:21:50.316 Part of continuing 1B1,

NOTE Confidence: 0.700787762

43

00:21:50.316 --> 00:21:52.604 it's a domain that is known to be

NOTE Confidence: 0.700787762

00:21:52.604 --> 00:21:54.380 oncogenic when it gets mutated slightly

NOTE Confidence: 0.700787762

00:21:54.380 --> 00:21:55.990 lower level in terms of the others.

NOTE Confidence: 0.700787762

00:21:55.990 --> 00:21:58.958 But the whole region across here is sort

NOTE Confidence: 0.700787762

00:21:58.958 --> 00:22:02.356 of known to be important to to oncogenesis.

NOTE Confidence: 0.700787762

00:22:02.360 --> 00:22:04.336 And So what you can basically do is

NOTE Confidence: 0.700787762

00:22:04.336 --> 00:22:06.338 take the prevalence that we see and and

NOTE Confidence: 0.700787762

00:22:06.338 --> 00:22:08.997 this is in a very crude terms but and

NOTE Confidence: 0.700787762

00:22:08.997 --> 00:22:10.994 there's some corrections that are involved,

NOTE Confidence: 0.700787762

00:22:10.994 --> 00:22:13.906 I'm not going to go into but essentially

NOTE Confidence: 0.700787762

00:22:13.906 --> 00:22:15.851 divide the expectations the observed

NOTE Confidence: 0.700787762

00:22:15.851 --> 00:22:18.760 here by the expected block on the same.

NOTE Confidence: 0.700787762

00:22:18.760 --> 00:22:20.821 On the same plot on the left and that

NOTE Confidence: 0.700787762

00:22:20.821 --> 00:22:22.979 gives you a metric for the cancer factor.

NOTE Confidence: 0.700787762

00:22:22.980 --> 00:22:25.200 How strongly that that site

NOTE Confidence: 0.700787762

00:22:25.200 --> 00:22:26.976 is mutated that sorry,

NOTE Confidence: 0.700787762

00:22:26.980 --> 00:22:29.305 how strongly that site is

NOTE Confidence: 0.700787762

00:22:29.305 --> 00:22:31.630 selected once it is mutated.

NOTE Confidence: 0.700787762

00:22:31.630 --> 00:22:33.742 And as I said these are well known

NOTE Confidence: 0.700787762

00:22:33.742 --> 00:22:35.870 sites in these particular cancers.

NOTE Confidence: 0.700787762

00:22:35.870 --> 00:22:37.542 And if you do that across all the

NOTE Confidence: 0.700787762

00:22:37.542 --> 00:22:38.816 different sites that you can look at

NOTE Confidence: 0.700787762

00:22:38.816 --> 00:22:40.323 what you see is a is a distribution

NOTE Confidence: 0.700787762

00:22:40.323 --> 00:22:41.842 that looks like this where on the

NOTE Confidence: 0.700787762

00:22:41.850 --> 00:22:44.769 X axis is the cancer effect size.

NOTE Confidence: 0.700787762

00:22:44.770 --> 00:22:47.042 It ranges from 10 to the zero to 10 to the

NOTE Confidence: 0.700787762

00:22:47.042 --> 00:22:49.186 6th maybe even a little bit more typically.

NOTE Confidence: 0.700787762

00:22:49.190 --> 00:22:50.350 And why is that?

NOTE Confidence: 0.700787762

00:22:50.350 --> 00:22:50.640 Why,

NOTE Confidence: 0.700787762

00:22:50.640 --> 00:22:52.570 what does this range mean?

NOTE Confidence: 0.700787762

00:22:52.570 --> 00:22:54.145 The range is what it is because

NOTE Confidence: 0.700787762

00:22:54.145 --> 00:22:55.566 that has it's it's complicated and

NOTE Confidence: 0.700787762

00:22:55.566 --> 00:22:57.763 I don't want to go into a lot of

NOTE Confidence: 0.700787762

00:22:57.763 --> 00:22:58.823 detail but population genetically

NOTE Confidence: 0.700787762

00:22:58.823 --> 00:23:00.715 it has to do with the population.

NOTE Confidence: 0.700787762

00:23:00.715 --> 00:23:01.815 Size of the cancer,

NOTE Confidence: 0.700787762

00:23:01.820 --> 00:23:03.024 the reproductive population size.

NOTE Confidence: 0.700787762

00:23:03.024 --> 00:23:04.830 How many cells in the cancer

NOTE Confidence: 0.700787762

00:23:04.884 --> 00:23:06.099 could possibly reproduce?

NOTE Confidence: 0.700787762

00:23:06.100 --> 00:23:07.234 I'm not going to go into more

NOTE Confidence: 0.700787762

00:23:07.234 --> 00:23:07.720 saying about that,

NOTE Confidence: 0.700787762

00:23:07.720 --> 00:23:09.904 but that's why it exists across

NOTE Confidence: 0.700787762

00:23:09.904 --> 00:23:10.996 this wide range.

NOTE Confidence: 0.700787762

00:23:11.000 --> 00:23:12.808 The density here is just I'm just going

NOTE Confidence: 0.700787762

00:23:12.808 --> 00:23:14.711 to density a plot across cancer effect

NOTE Confidence: 0.700787762

00:23:14.711 --> 00:23:16.111 size of these different mutations.

NOTE Confidence: 0.700787762

00:23:16.120 --> 00:23:18.272 So most of the mutations lie at this

NOTE Confidence: 0.700787762

00:23:18.272 --> 00:23:20.667 very low range where it's not even

NOTE Confidence: 0.700787762

00:23:20.667 --> 00:23:22.911 clear necessarily if they're under any

NOTE Confidence: 0.700787762

00:23:22.911 --> 00:23:25.836 selection below say 10 to the four or so.

NOTE Confidence: 0.700787762

00:23:25.840 --> 00:23:28.080 And in blue I show you the

NOTE Confidence: 0.700787762

00:23:28.080 --> 00:23:29.862 synonymous mutations and in red

NOTE Confidence: 0.700787762

00:23:29.862 --> 00:23:31.370 the non synonymous mutations.

NOTE Confidence: 0.700787762

00:23:31.370 --> 00:23:33.200 So there's just a slight,

NOTE Confidence: 0.700787762

00:23:33.200 --> 00:23:36.164 a slight bias over the synonymous

NOTE Confidence: 0.700787762

00:23:36.164 --> 00:23:37.646 mutations of NONSYNONYMOUS

NOTE Confidence: 0.700787762

00:23:37.646 --> 00:23:39.339 mutations to be oncogenic.

NOTE Confidence: 0.700787762

00:23:39.340 --> 00:23:40.460 But the really important mutations

NOTE Confidence: 0.700787762

00:23:40.460 --> 00:23:42.039 are all out on this tail here,

NOTE Confidence: 0.700787762

00:23:42.040 --> 00:23:43.656 and I've just shown 2 here for reference.

NOTE Confidence: 0.700787762

00:23:43.660 --> 00:23:46.019 Here's a P53 mutation that's quite common.

NOTE Confidence: 0.700787762

00:23:46.020 --> 00:23:48.228 Here's an NF2L2 mutation is quite

NOTE Confidence: 0.700787762

00:23:48.228 --> 00:23:50.740 common in lung squamous cell carcinoma.

NOTE Confidence: 0.700787762

00:23:50.740 --> 00:23:52.144 So these tail mutations are the

NOTE Confidence: 0.700787762

00:23:52.144 --> 00:23:53.080 ones that are important.

NOTE Confidence: 0.700787762

00:23:53.080 --> 00:23:54.235 And this harks back to

NOTE Confidence: 0.700787762

00:23:54.235 --> 00:23:55.390 what I was saying earlier

NOTE Confidence: 0.817316937692308

00:23:55.442 --> 00:23:57.332 when we say, oh, lots of mutations

NOTE Confidence: 0.817316937692308

00:23:57.332 --> 00:23:59.300 are happening in the genome because

NOTE Confidence: 0.817316937692308

00:23:59.363 --> 00:24:01.379 of UV light or something like that.

NOTE Confidence: 0.817316937692308

00:24:01.380 --> 00:24:02.765 If they're not these key

NOTE Confidence: 0.817316937692308

00:24:02.765 --> 00:24:04.490 mutations out here on the tail,

NOTE Confidence: 0.817316937692308

00:24:04.490 --> 00:24:06.266 they're not contributing much to cancer.

NOTE Confidence: 0.817316937692308

00:24:06.270 --> 00:24:08.942 So we really need that component to be

NOTE Confidence: 0.817316937692308

00:24:08.942 --> 00:24:11.690 included if we want to ask the question

NOTE Confidence: 0.817316937692308

00:24:11.690 --> 00:24:14.238 what is causing cancer in an individual.

NOTE Confidence: 0.817316937692308

00:24:14.240 --> 00:24:16.740 Tumor in a digital patient.

NOTE Confidence: 0.817316937692308

00:24:16.740 --> 00:24:17.750 You can do this diagram

NOTE Confidence: 0.817316937692308

00:24:17.750 --> 00:24:18.760 not just for lung cancer,

NOTE Confidence: 0.817316937692308

00:24:18.760 --> 00:24:20.896 but for lots of different cancers,

NOTE Confidence: 0.817316937692308

00:24:20.900 --> 00:24:22.964 and we see very much the same pattern.

NOTE Confidence: 0.834948496

00:24:31.370 --> 00:24:33.474 OK. Now just to provide you a little

NOTE Confidence: 0.834948496

00:24:33.474 --> 00:24:35.682 bit of perhaps validation that this

NOTE Confidence: 0.834948496

00:24:35.682 --> 00:24:38.106 cancer effect you know is meaningful,

NOTE Confidence: 0.834948496

00:24:38.110 --> 00:24:39.445 probably many of you are

NOTE Confidence: 0.834948496

00:24:39.445 --> 00:24:40.513 familiar with GLENVAR variants,

NOTE Confidence: 0.834948496

00:24:40.520 --> 00:24:43.663 variants that have been attributed over

NOTE Confidence: 0.834948496

00:24:43.663 --> 00:24:46.928 time with some clinical significance.

NOTE Confidence: 0.834948496

00:24:46.930 --> 00:24:48.690 And these by the way these are Clint

NOTE Confidence: 0.834948496

00:24:48.690 --> 00:24:50.400 Barbarians that were attributed significance,

NOTE Confidence: 0.834948496

00:24:50.400 --> 00:24:52.865 not potential, not ones that

NOTE Confidence: 0.834948496

00:24:52.865 --> 00:24:54.344 weren't attributed significance.

NOTE Confidence: 0.834948496

00:24:54.350 --> 00:24:57.650 And on the X axis we've sort of divided them,

NOTE Confidence: 0.834948496

00:24:57.650 --> 00:24:58.810 those Clint Barbarians up and

NOTE Confidence: 0.834948496

00:24:58.810 --> 00:24:59.970 some categories I'll talk about,

NOTE Confidence: 0.834948496

00:24:59.970 --> 00:25:01.290 but on the Y axis.

NOTE Confidence: 0.834948496

00:25:01.290 --> 00:25:03.040 Is the scale selection coefficient,

NOTE Confidence: 0.834948496

00:25:03.040 --> 00:25:04.784 and generally there's basically

NOTE Confidence: 0.834948496

00:25:04.784 --> 00:25:05.656 2 comparisons.

NOTE Confidence: 0.834948496

00:25:05.660 --> 00:25:06.818 I really want to emphasize here.

NOTE Confidence: 0.834948496

00:25:06.820 --> 00:25:09.207 If we look at glenvar single nucleotide

NOTE Confidence: 0.834948496

00:25:09.207 --> 00:25:11.131 variants that are recurrent within

NOTE Confidence: 0.834948496

00:25:11.131 --> 00:25:13.795 cancer type and compare it to other

NOTE Confidence: 0.834948496

00:25:13.795 --> 00:25:15.255 single nucleotide variants that

NOTE Confidence: 0.834948496

00:25:15.255 --> 00:25:17.477 are recurrent within cancer type,

NOTE Confidence: 0.834948496

00:25:17.480 --> 00:25:19.676 we see that the GLENVAR variants have a much,

NOTE Confidence: 0.834948496

00:25:19.680 --> 00:25:21.325 much higher distribution of selection

NOTE Confidence: 0.834948496

00:25:21.325 --> 00:25:23.700 coefficient than the ones that are other SNV.

NOTE Confidence: 0.834948496

00:25:23.700 --> 00:25:26.100 So in other words, there's,

NOTE Confidence: 0.834948496

00:25:26.100 --> 00:25:27.768 you know this literally,

NOTE Confidence: 0.834948496

00:25:27.768 --> 00:25:29.853 this is saying that Glenvar

NOTE Confidence: 0.834948496

00:25:29.853 --> 00:25:31.300 predicts cancer effect.

NOTE Confidence: 0.834948496

00:25:31.300 --> 00:25:32.889 But the opposite is true and I'll

NOTE Confidence: 0.834948496

00:25:32.889 --> 00:25:34.488 show you that in the next slide.

NOTE Confidence: 0.834948496

00:25:34.490 --> 00:25:37.050 And then we can also compare Glenvar STD's

NOTE Confidence: 0.834948496

00:25:37.050 --> 00:25:39.896 that are a single hit within a cancer type.

NOTE Confidence: 0.834948496

00:25:39.900 --> 00:25:42.879 That is ones that we only see once when

NOTE Confidence: 0.834948496

00:25:42.879 --> 00:25:45.550 they're clean bar single nukite variance

NOTE Confidence: 0.834948496

00:25:45.550 --> 00:25:48.385 versus other SNB's that are single hit.

NOTE Confidence: 0.834948496

00:25:48.390 --> 00:25:50.679 And you can see that the cancer

NOTE Confidence: 0.834948496

00:25:50.679 --> 00:25:53.361 affect size of those ones that are

NOTE Confidence: 0.834948496

00:25:53.361 --> 00:25:55.797 you know known oncogenic are believed

NOTE Confidence: 0.834948496

00:25:55.868 --> 00:25:58.342 oncogenic variants have a much higher

NOTE Confidence: 0.834948496

00:25:58.342 --> 00:26:00.130 cancer effect than the ones that

NOTE Confidence: 0.834948496

00:26:00.190 --> 00:26:02.116 are not believed to be oncogenic.

NOTE Confidence: 0.834948496

00:26:02.120 --> 00:26:04.353 And this is a highly significant from

NOTE Confidence: 0.834948496

00:26:04.353 --> 00:26:06.698 a a statistical science point of view.

NOTE Confidence: 0.834948496

00:26:06.700 --> 00:26:08.392 By the way, this is work of Jeffrey Mandel,

NOTE Confidence: 0.834948496

00:26:08.400 --> 00:26:09.597 who's sitting over here in the audience,

NOTE Confidence: 0.834948496

00:26:09.600 --> 00:26:12.066 a grad student in my lab.

NOTE Confidence: 0.834948496

00:26:12.070 --> 00:26:16.888 And so that should be reassuring.

NOTE Confidence: 0.834948496

00:26:16.890 --> 00:26:17.278 Furthermore,

NOTE Confidence: 0.834948496

00:26:17.278 --> 00:26:20.382 if you take the mean or the top

NOTE Confidence: 0.834948496

00:26:20.382 --> 00:26:22.830 cancer effect of a given variant,

NOTE Confidence: 0.834948496

00:26:22.830 --> 00:26:25.065 they're much stronger predictions of

NOTE Confidence: 0.834948496

00:26:25.065 --> 00:26:27.790 glenvar status than the SIFT score,

NOTE Confidence: 0.834948496

00:26:27.790 --> 00:26:29.550 the Polyphen 2 score,

NOTE Confidence: 0.834948496

00:26:29.550 --> 00:26:30.870 or variant prevalence.

NOTE Confidence: 0.834948496

00:26:30.870 --> 00:26:32.436 Any of these measures that are

NOTE Confidence: 0.834948496

00:26:32.436 --> 00:26:34.284 typically used to try to say whether

NOTE Confidence: 0.834948496

00:26:34.284 --> 00:26:35.790 a variant is important or not.

NOTE Confidence: 0.834948496

00:26:35.790 --> 00:26:36.158 So really,

NOTE Confidence: 0.834948496

00:26:36.158 --> 00:26:37.630 you should be using cancer effects if you

NOTE Confidence: 0.834948496

00:26:37.671 --> 00:26:39.087 want to know whether variance important.

NOTE Confidence: 0.75101091

00:26:41.280 --> 00:26:42.785 This is also work by Jeff Mandell.

NOTE Confidence: 0.755946486363636

00:26:45.150 --> 00:26:47.537 OK, so hopefully it persuaded you that

NOTE Confidence: 0.755946486363636

00:26:47.537 --> 00:26:49.641 cancer effect is a measure that you

NOTE Confidence: 0.755946486363636

00:26:49.641 --> 00:26:51.543 should be thoughtful about and use

NOTE Confidence: 0.755946486363636

00:26:51.543 --> 00:26:53.826 and and in the research you're doing.

NOTE Confidence: 0.755946486363636

00:26:53.830 --> 00:26:55.430 But what we wanted to get to from

NOTE Confidence: 0.755946486363636

00:26:55.430 --> 00:26:56.624 the beginning of this talk was

NOTE Confidence: 0.755946486363636

00:26:56.624 --> 00:26:58.169 the extent to which each of those

NOTE Confidence: 0.755946486363636

00:26:58.169 --> 00:26:59.669 processes contribute to tumorigenesis.

NOTE Confidence: 0.755946486363636

00:26:59.670 --> 00:27:01.278 So if you'll if you'll at least walk

NOTE Confidence: 0.755946486363636

00:27:01.278 --> 00:27:03.237 with me on the idea that cancer affect

NOTE Confidence: 0.755946486363636

00:27:03.237 --> 00:27:05.294 quantifies the degree to which a given

NOTE Confidence: 0.755946486363636

00:27:05.294 --> 00:27:06.626 variant contributes to tumorigenesis,

NOTE Confidence: 0.755946486363636

00:27:06.630 --> 00:27:09.276 then that apply that gives us the

NOTE Confidence: 0.755946486363636

00:27:09.276 --> 00:27:11.145 key to finish that association.

NOTE Confidence: 0.755946486363636

00:27:11.145 --> 00:27:13.718 I said. So we know, you know,

NOTE Confidence: 0.755946486363636

00:27:13.718 --> 00:27:15.140 from Alexandra's work.

NOTE Confidence: 0.755946486363636

00:27:15.140 --> 00:27:16.525 The degree to which, no, sorry.

NOTE Confidence: 0.755946486363636

00:27:16.525 --> 00:27:18.910 We know from this work the degree to which

NOTE Confidence: 0.755946486363636

00:27:18.965 --> 00:27:21.125 mutations contribute to the increased

NOTE Confidence: 0.755946486363636

00:27:21.125 --> 00:27:22.853 cellular perforation and survival.

NOTE Confidence: 0.755946486363636

00:27:22.860 --> 00:27:25.636 And we know from Alexandra's work and others,

NOTE Confidence: 0.755946486363636

00:27:25.640 --> 00:27:28.358 some strain in Xanal and others

NOTE Confidence: 0.755946486363636

00:27:28.360 --> 00:27:30.980 what the contribution of various

NOTE Confidence: 0.755946486363636

00:27:30.980 --> 00:27:32.932 mutagenic processes toward creating

NOTE Confidence: 0.755946486363636

00:27:32.932 --> 00:27:34.396 those mutations are.

NOTE Confidence: 0.755946486363636

00:27:34.400 --> 00:27:36.296 And so by putting those two things together,

NOTE Confidence: 0.755946486363636

00:27:36.300 --> 00:27:38.832 we can understand the relationship between

NOTE Confidence: 0.755946486363636

00:27:38.832 --> 00:27:40.520 these increased cellular proliferation

NOTE Confidence: 0.755946486363636

00:27:40.573 --> 00:27:42.883 and survival and the actual processes

NOTE Confidence: 0.755946486363636

00:27:42.883 --> 00:27:44.423 underlying these mutational effects.

NOTE Confidence: 0.830300483333333

00:27:46.450 --> 00:27:49.058 So just going back again to Alexandra's work,

NOTE Confidence: 0.830300483333333

00:27:49.058 --> 00:27:50.678 we know each signature contributes

NOTE Confidence: 0.830300483333333

00:27:50.678 --> 00:27:51.939 differentially to mutation counts

NOTE Confidence: 0.830300483333333

00:27:51.939 --> 00:27:53.169 observed in each cancer type.

NOTE Confidence: 0.830300483333333

00:27:53.170 --> 00:27:55.447 I showed this slide earlier and

NOTE Confidence: 0.830300483333333

00:27:55.447 --> 00:27:56.932 here's here's the slide where

NOTE Confidence: 0.830300483333333

00:27:56.932 --> 00:27:58.979 you can you can sort of like.

NOTE Confidence: 0.830300483333333

00:27:58.980 --> 00:28:01.260 Fade out for a moment if you want,

NOTE Confidence: 0.830300483333333

00:28:01.260 --> 00:28:02.436 and then come back in a moment.

NOTE Confidence: 0.830300483333333

00:28:02.440 --> 00:28:03.289 It's only saying.

NOTE Confidence: 0.830300483333333

00:28:03.289 --> 00:28:05.615 What it's it's this is I'm going to

NOTE Confidence: 0.830300483333333

00:28:05.615 --> 00:28:07.313 narrate through for those of you

NOTE Confidence: 0.830300483333333

00:28:07.313 --> 00:28:09.439 who are really interested how we

NOTE Confidence: 0.830300483333333

00:28:09.439 --> 00:28:11.059 actually calculate this process.

NOTE Confidence: 0.830300483333333

00:28:11.060 --> 00:28:12.632 But if if you've understood everything

NOTE Confidence: 0.830300483333333

00:28:12.632 --> 00:28:14.308 before, there's nothing new here.

NOTE Confidence: 0.830300483333333

00:28:14.308 --> 00:28:16.797 It's just the bookkeeping of how we

NOTE Confidence: 0.830300483333333

00:28:16.797 --> 00:28:18.657 calculate this process and the the

NOTE Confidence: 0.830300483333333

00:28:18.657 --> 00:28:21.450 point is that forget for each for each.

NOTE Confidence: 0.830300483333333

00:28:21.450 --> 00:28:23.230 A source of mutation.

NOTE Confidence: 0.830300483333333

00:28:23.230 --> 00:28:26.068 Here's deamination with age apobec tobacco,

NOTE Confidence: 0.830300483333333

00:28:26.070 --> 00:28:28.008 and then unload clock like signature,

NOTE Confidence: 0.830300483333333

00:28:28.010 --> 00:28:29.924 which were the four sources that

NOTE Confidence: 0.830300483333333

00:28:29.924 --> 00:28:32.016 came out of the deconvolution for

NOTE Confidence: 0.830300483333333

00:28:32.016 --> 00:28:34.620 a particular tumor in the TCA data

NOTE Confidence: 0.830300483333333

00:28:34.620 --> 00:28:37.220 set that turned out to be useful

NOTE Confidence: 0.830300483333333

00:28:37.220 --> 00:28:38.656 for illustration of this.

NOTE Confidence: 0.830300483333333

00:28:38.660 --> 00:28:40.420 For each of those processes,

NOTE Confidence: 0.830300483333333

00:28:40.420 --> 00:28:42.382 there's a weight of mutation that

NOTE Confidence: 0.830300483333333

00:28:42.382 --> 00:28:44.460 they contribute to given trying time

NOTE Confidence: 0.830300483333333

00:28:44.460 --> 00:28:46.578 nucleotides that are listed down here.

NOTE Confidence: 0.830300483333333

00:28:46.580 --> 00:28:48.650 So deamination with age really

NOTE Confidence: 0.830300483333333

00:28:48.650 --> 00:28:51.819 focuses on these AC to TG mutations.

NOTE Confidence: 0.830300483333333

00:28:51.820 --> 00:28:53.500 That's what they cause for the most part.

NOTE Confidence: 0.830300483333333

00:28:53.500 --> 00:28:55.131 But then there's a few other ones

NOTE Confidence: 0.830300483333333

00:28:55.131 --> 00:28:56.629 here that are quite frequent.

NOTE Confidence: 0.830300483333333

00:28:56.630 --> 00:29:01.058 Apobec really focuses on TCA or

NOTE Confidence: 0.830300483333333

00:29:01.060 --> 00:29:04.980 TCC or TCG or TCT changing to T,

NOTE Confidence: 0.830300483333333

00:29:04.980 --> 00:29:07.740 and tobacco has a broader distribution

NOTE Confidence: 0.830300483333333

00:29:07.740 --> 00:29:08.660 of neurogenic.

NOTE Confidence: 0.830300483333333

00:29:08.660 --> 00:29:10.886 In fact and this unknown clock

NOTE Confidence: 0.830300483333333

00:29:10.886 --> 00:29:11.628 like signature,

NOTE Confidence: 0.830300483333333

57

00:29:11.630 --> 00:29:14.978 there's another aging signature has a

NOTE Confidence: 0.830300483333333

00:29:14.978 --> 00:29:17.910 generally quite broad distribution as well.

NOTE Confidence: 0.830300483333333

00:29:17.910 --> 00:29:21.795 So we deconvolve that tumor into these

NOTE Confidence: 0.830300483333333

00:29:21.795 --> 00:29:23.376 different signatures to understand

NOTE Confidence: 0.830300483333333

00:29:23.376 --> 00:29:25.428 how much each one is contributing.

NOTE Confidence: 0.830300483333333

00:29:25.430 --> 00:29:27.302 That gives us a signature weight

NOTE Confidence: 0.830300483333333

00:29:27.302 --> 00:29:28.550 for every signature here.

NOTE Confidence: 0.830300483333333

00:29:28.550 --> 00:29:29.870 And I'm just emphasizing that,

NOTE Confidence: 0.830300483333333

00:29:29.870 --> 00:29:30.658 you know,

NOTE Confidence: 0.830300483333333

00:29:30.658 --> 00:29:33.810 we can do lots of uncertainty analysis by.

NOTE Confidence: 0.830300483333333

00:29:33.810 --> 00:29:34.734 Bootstrapping the signature,

NOTE Confidence: 0.830300483333333

00:29:34.734 --> 00:29:35.042 deconvolution,

NOTE Confidence: 0.830300483333333

00:29:35.042 --> 00:29:37.175 that's what all these dots are many

NOTE Confidence: 0.830300483333333

00:29:37.175 --> 00:29:38.385 bootstraps on and given tumor,

NOTE Confidence: 0.830300483333333

00:29:38.390 --> 00:29:39.896 just saying how much of that

NOTE Confidence: 0.830300483333333

00:29:39.896 --> 00:29:41.270 signature do we really believe

NOTE Confidence: 0.830300483333333

00:29:41.270 --> 00:29:42.870 is contributing to that cancer.

NOTE Confidence: 0.830300483333333

00:29:42.870 --> 00:29:45.086 So we do do that and then you

NOTE Confidence: 0.830300483333333

00:29:45.086 --> 00:29:46.570 can also and then in addition to

NOTE Confidence: 0.830300483333333

00:29:46.617 --> 00:29:47.981 understanding how much signature

NOTE Confidence: 0.830300483333333

00:29:47.981 --> 00:29:49.345 is contributed to cancer,

NOTE Confidence: 0.830300483333333

00:29:49.350 --> 00:29:51.168 we look at the probability that

NOTE Confidence: 0.830300483333333

00:29:51.168 --> 00:29:52.870 each source created each variant.

NOTE Confidence: 0.830300483333333

00:29:52.870 --> 00:29:54.130 And we know that because we know

NOTE Confidence: 0.830300483333333

00:29:54.130 --> 00:29:55.404 what the sources are and we can

NOTE Confidence: 0.830300483333333

00:29:55.404 --> 00:29:56.442 just look at the relative height

NOTE Confidence: 0.830300483333333

00:29:56.480 --> 00:29:57.740 of these bars essentially to give

NOTE Confidence: 0.830300483333333

00:29:57.740 --> 00:29:59.320 us the probability that each source

NOTE Confidence: 0.830300483333333

00:29:59.320 --> 00:30:01.540 contributed to a given variant and

NOTE Confidence: 0.830300483333333

00:30:01.540 --> 00:30:04.339 then that probability comes out of that.

NOTE Confidence: 0.830300483333333

00:30:04.340 --> 00:30:06.265 Just by multiplying those together

NOTE Confidence: 0.830300483333333

00:30:06.265 --> 00:30:08.236 essentially and that gives US4P53

NOTE Confidence: 0.830300483333333

00:30:08.236 --> 00:30:10.792 here KF5 and this odorant receptor

NOTE Confidence: 0.830300483333333

00:30:10.792 --> 00:30:12.740 which doesn't have much cancer effect,

NOTE Confidence: 0.830300483333333

00:30:12.740 --> 00:30:14.750 what the probably each source

NOTE Confidence: 0.830300483333333

00:30:14.750 --> 00:30:16.760 contributed to creating each variant.

NOTE Confidence: 0.830300483333333

00:30:16.760 --> 00:30:18.265 And then we take that effect size

NOTE Confidence: 0.830300483333333

00:30:18.265 --> 00:30:19.778 that I just described to you,

NOTE Confidence: 0.830300483333333

00:30:19.780 --> 00:30:22.900 which is very high for this P53 variant,

NOTE Confidence: 0.830300483333333

00:30:22.900 --> 00:30:24.466 quite a bit lower for KF5,

NOTE Confidence: 0.830300483333333

00:30:24.470 --> 00:30:26.498 but still there and is basically

NOTE Confidence: 0.830300483333333

00:30:26.498 --> 00:30:28.413 nonexistent for the odorant receptor

NOTE Confidence: 0.830300483333333

00:30:28.413 --> 00:30:29.181 mutation. So.

NOTE Confidence: 0.830300483333333

00:30:29.181 --> 00:30:31.008 So this is a really important variant,

NOTE Confidence: 0.830300483333333

00:30:31.010 --> 00:30:32.354 this is a less important variant

NOTE Confidence: 0.830300483333333

00:30:32.354 --> 00:30:33.770 and this is not important.

NOTE Confidence: 0.823364442333333

00:30:33.770 --> 00:30:35.930 Fall and then we can just sort of

NOTE Confidence: 0.823364442333333

00:30:35.930 --> 00:30:37.848 multiply through each variant by the

NOTE Confidence: 0.823364442333333

00:30:37.848 --> 00:30:39.498 probability that each source created

NOTE Confidence: 0.823364442333333

00:30:39.498 --> 00:30:41.532 that variant and that gives us this

NOTE Confidence: 0.823364442333333

00:30:41.532 --> 00:30:43.420 final thing which is the proportional

NOTE Confidence: 0.823364442333333

00:30:43.420 --> 00:30:44.980 mutation source effect size.

NOTE Confidence: 0.823364442333333

00:30:44.980 --> 00:30:45.997 That's a mouthful.

NOTE Confidence: 0.823364442333333

00:30:45.997 --> 00:30:48.370 But what we're just trying to say

NOTE Confidence: 0.823364442333333

00:30:48.440 --> 00:30:50.748 is how much of this given variant

NOTE Confidence: 0.823364442333333

00:30:50.748 --> 00:30:53.418 was caused by the particular

NOTE Confidence: 0.823364442333333

00:30:53.418 --> 00:30:55.540 mutational process and or sorry,

NOTE Confidence: 0.823364442333333

00:30:55.540 --> 00:30:57.360 how much of the selection for oncogenesis

NOTE Confidence: 0.823364442333333

00:30:57.360 --> 00:30:59.276 was caused by that particular process.

NOTE Confidence: 0.823364442333333

00:30:59.280 --> 00:31:03.004 So the TP50 were bar the TP 53 bars.

NOTE Confidence: 0.823364442333333

00:31:03.004 --> 00:31:05.021 Are much higher than the ones in

NOTE Confidence: 0.823364442333333

00:31:05.021 --> 00:31:07.121 KF5 and are those are way higher

NOTE Confidence: 0.823364442333333

00:31:07.121 --> 00:31:08.928 than anything in order receptor

NOTE Confidence: 0.823364442333333

00:31:08.928 --> 00:31:10.828 because the odorant receptor in

NOTE Confidence: 0.823364442333333

00:31:10.828 --> 00:31:13.222 fact doesn't do anything for cancer.

NOTE Confidence: 0.823364442333333

00:31:13.222 --> 00:31:16.130 So the average then you can then

NOTE Confidence: 0.823364442333333

00:31:16.130 --> 00:31:18.790 you can look across all of those,

NOTE Confidence: 0.823364442333333

00:31:18.790 --> 00:31:19.710 all of the variants,

NOTE Confidence: 0.823364442333333

00:31:19.710 --> 00:31:21.641 not just these ones to look at what

NOTE Confidence: 0.823364442333333

00:31:21.641 --> 00:31:22.886 the average attributable effect size

NOTE Confidence: 0.823364442333333

00:31:22.886 --> 00:31:25.273 is in a given tumor and you get this

NOTE Confidence: 0.823364442333333

00:31:25.273 --> 00:31:27.100 distribution which says oh for this

NOTE Confidence: 0.823364442333333

00:31:27.100 --> 00:31:29.795 tumor you know most of the oncogenic

NOTE Confidence: 0.823364442333333

00:31:29.795 --> 00:31:32.598 cause came from deamination with age.

NOTE Confidence: 0.823364442333333

00:31:32.600 --> 00:31:35.348 And for this tumor?

NOTE Confidence: 0.823364442333333

00:31:35.350 --> 00:31:37.422 You know the second most common process

NOTE Confidence: 0.823364442333333

00:31:37.422 --> 00:31:39.589 that was creating mutations that led to

NOTE Confidence: 0.823364442333333

00:31:39.589 --> 00:31:41.425 oncogenesis was this light Gray which

NOTE Confidence: 0.823364442333333

00:31:41.484 --> 00:31:43.458 is this unknown clock like SIEGENER 5.

NOTE Confidence: 0.823364442333333

00:31:43.460 --> 00:31:43.752 So.

NOTE Confidence: 0.823364442333333

00:31:43.752 --> 00:31:45.796 So this is a largely aging driven

NOTE Confidence: 0.823364442333333

00:31:45.796 --> 00:31:47.559 tumor and there's a little bit

NOTE Confidence: 0.823364442333333

00:31:47.559 --> 00:31:49.637 of Apple back here and a little

NOTE Confidence: 0.823364442333333

00:31:49.637 --> 00:31:51.760 bit of tobacco smoke and and you

NOTE Confidence: 0.823364442333333

00:31:51.760 --> 00:31:53.320 can follow it through like that.

NOTE Confidence: 0.823364442333333

00:31:53.320 --> 00:31:55.412 So this is one example for a

NOTE Confidence: 0.823364442333333

00:31:55.412 --> 00:31:56.828 given tumor and then that result,

NOTE Confidence: 0.823364442333333

00:31:56.830 --> 00:31:58.713 you know it basically tells you what

NOTE Confidence: 0.823364442333333

00:31:58.713 --> 00:32:00.702 at least with the knowledge we have

NOTE Confidence: 0.823364442333333

00:32:00.702 --> 00:32:02.750 right now what the effect size by

NOTE Confidence: 0.823364442333333

00:32:02.750 --> 00:32:04.598 mutational source for this tumor was,

NOTE Confidence: 0.823364442333333

00:32:04.600 --> 00:32:05.830 this is a lung cancer tumor.

NOTE Confidence: 0.823364442333333

00:32:05.830 --> 00:32:07.438 By the way.

NOTE Confidence: 0.823364442333333

00:32:07.440 --> 00:32:10.392 Now you can look at this not just at,

NOTE Confidence: 0.823364442333333

00:32:10.400 --> 00:32:10.972 you know,

NOTE Confidence: 0.823364442333333

00:32:10.972 --> 00:32:12.688 you can sort of understand that

NOTE Confidence: 0.823364442333333

00:32:12.688 --> 00:32:13.999 for a given site,

NOTE Confidence: 0.823364442333333

00:32:14.000 --> 00:32:18.072 but then you can also look at

NOTE Confidence: 0.823364442333333

00:32:18.072 --> 00:32:20.689 what a set of sites all look like.

NOTE Confidence: 0.823364442333333

00:32:20.690 --> 00:32:22.420 So this is just a diagram where we do that.

NOTE Confidence: 0.823364442333333

00:32:22.420 --> 00:32:22.744 Again,

NOTE Confidence: 0.823364442333333

00:32:22.744 --> 00:32:24.040 a little bit complex,

NOTE Confidence: 0.823364442333333

00:32:24.040 --> 00:32:25.980 but hopefully this everyone can

NOTE Confidence: 0.823364442333333

00:32:25.980 --> 00:32:27.532 follow along directly with.

NOTE Confidence: 0.823364442333333

00:32:27.540 --> 00:32:29.316 If you look across the genome,

NOTE Confidence: 0.823364442333333

00:32:29.320 --> 00:32:30.820 there's an average mutational weight.

NOTE Confidence: 0.823364442333333

00:32:30.820 --> 00:32:32.536 So tobacco smoke is causing a

NOTE Confidence: 0.823364442333333

00:32:32.536 --> 00:32:34.300 certain number of the mutations,

NOTE Confidence: 0.823364442333333

00:32:34.300 --> 00:32:35.530 certain proportion of the mutations

64

NOTE Confidence: 0.823364442333333

00:32:35.530 --> 00:32:37.120 and then a number of others.

NOTE Confidence: 0.823364442333333

00:32:37.120 --> 00:32:38.108 And in these diagrams,

NOTE Confidence: 0.823364442333333

00:32:38.108 --> 00:32:39.909 I've sort of put the major mutagenic

NOTE Confidence: 0.823364442333333

00:32:39.909 --> 00:32:41.743 cause on the left and then stacked

NOTE Confidence: 0.823364442333333

00:32:41.743 --> 00:32:43.636 all the other causes on the right,

NOTE Confidence: 0.823364442333333

00:32:43.640 --> 00:32:45.590 just because it helps you really

NOTE Confidence: 0.823364442333333

00:32:45.590 --> 00:32:46.890 see the differential effect

NOTE Confidence: 0.823364442333333

00:32:46.948 --> 00:32:48.520 of these different processes.

NOTE Confidence: 0.823364442333333

00:32:48.520 --> 00:32:51.502 So tobacco smoking is the major cause

NOTE Confidence: 0.823364442333333

00:32:51.502 --> 00:32:55.532 of of loads in general in terms of

NOTE Confidence: 0.823364442333333

00:32:55.532 --> 00:32:57.660 the underlying genomic mutation.

NOTE Confidence: 0.823364442333333

00:32:57.660 --> 00:33:00.558 But if you look at from site to site,

NOTE Confidence: 0.823364442333333

00:33:00.560 --> 00:33:02.756 each site has a different probability

NOTE Confidence: 0.823364442333333

00:33:02.756 --> 00:33:05.199 of being caused by tobacco smoke.

NOTE Confidence: 0.823364442333333

00:33:05.200 --> 00:33:06.660 So here's.

NOTE Confidence: 0.823364442333333

00:33:06.660 --> 00:33:09.201 KSG 12C very, very,

NOTE Confidence: 0.823364442333333

00:33:09.201 --> 00:33:11.307 very strong caused by tobacco smoke,

NOTE Confidence: 0.823364442333333

00:33:11.310 --> 00:33:13.092 maybe that's not surprising it in

NOTE Confidence: 0.823364442333333

00:33:13.092 --> 00:33:14.910 lung cancer, we see that variant very,

NOTE Confidence: 0.823364442333333

00:33:14.910 --> 00:33:15.826 very frequently.

NOTE Confidence: 0.823364442333333

00:33:15.826 --> 00:33:18.780 We very rarely see care SG12C in

NOTE Confidence: 0.823364442333333

00:33:18.780 --> 00:33:20.430 other cancers like pancreatic cancer,

NOTE Confidence: 0.823364442333333

00:33:20.430 --> 00:33:21.890 other K rosterman cancers.

NOTE Confidence: 0.823364442333333

00:33:21.890 --> 00:33:23.350 So why is that?

NOTE Confidence: 0.839534726

00:33:23.350 --> 00:33:25.450 Well, it's just because that site is

NOTE Confidence: 0.839534726

00:33:25.450 --> 00:33:28.189 hit a lot more in terms of mutations.

NOTE Confidence: 0.839534726

00:33:28.190 --> 00:33:31.007 It's not a doesn't appear to have anything to

NOTE Confidence: 0.839534726

00:33:31.007 --> 00:33:34.012 do from our calculations with its particular

NOTE Confidence: 0.839534726

00:33:34.012 --> 00:33:36.810 cancer effect relative to other variants.

NOTE Confidence: 0.839534726

00:33:36.810 --> 00:33:40.814 And, and in contrast, here's EGFR LA58R.

NOTE Confidence: 0.839534726

00:33:40.814 --> 00:33:42.998 It's a long known fact that you

NOTE Confidence: 0.839534726

00:33:42.998 --> 00:33:45.382 rarely see those in individuals who

NOTE Confidence: 0.839534726

00:33:45.382 --> 00:33:46.850 are non-smokers, are smokers.

NOTE Confidence: 0.839534726

00:33:46.850 --> 00:33:49.160 You see that in non-smokers and the

NOTE Confidence: 0.839534726

00:33:49.160 --> 00:33:51.357 reason is it's not caused by smoking.

NOTE Confidence: 0.839534726

00:33:51.360 --> 00:33:54.564 So when you see a patient with these Fr

NOTE Confidence: 0.839534726

00:33:54.564 --> 00:33:57.020 mutation, they typically are not smoker.

NOTE Confidence: 0.839534726

00:33:57.020 --> 00:33:59.708 There's a lot more individuals coming

NOTE Confidence: 0.839534726

00:33:59.708 --> 00:34:02.897 in with EGFR who are not smokers then

NOTE Confidence: 0.839534726

00:34:02.897 --> 00:34:04.684 are smokers relatively speaking.

NOTE Confidence: 0.839534726

00:34:04.684 --> 00:34:08.700 So you can do this for lung adenocarcinoma.

NOTE Confidence: 0.839534726

00:34:08.700 --> 00:34:10.513 You can look at other variants of

NOTE Confidence: 0.839534726

00:34:10.513 --> 00:34:12.060 course lung squamous cell carcinoma.

NOTE Confidence: 0.839534726

00:34:12.060 --> 00:34:15.492 Here you see PI3 kinase largely driven

NOTE Confidence: 0.839534726

00:34:15.492 --> 00:34:18.053 in lung squamous cell carcinoma by other

NOTE Confidence: 0.839534726

00:34:18.053 --> 00:34:21.870 effects, mostly apobec but in fact.

NOTE Confidence: 0.839534726

00:34:21.870 --> 00:34:23.235 Not at all driven by tobacco smoke.

NOTE Confidence: 0.839534726

00:34:23.240 --> 00:34:24.880 Again, that's an empirical observation

NOTE Confidence: 0.839534726

00:34:24.880 --> 00:34:26.892 that people have noted many times

NOTE Confidence: 0.839534726

00:34:26.892 --> 00:34:28.507 that individuals with lung squamous

NOTE Confidence: 0.839534726

00:34:28.507 --> 00:34:30.475 cell carcinoma who have PI3 kinase

NOTE Confidence: 0.839534726

00:34:30.475 --> 00:34:32.353 mutation are rarely are less frequently

NOTE Confidence: 0.839534726

00:34:32.353 --> 00:34:34.408 smokers than than other mutations.

NOTE Confidence: 0.839534726

00:34:34.408 --> 00:34:36.892 And P3 mutations, on the other hand,

NOTE Confidence: 0.839534726

00:34:36.892 --> 00:34:37.418 are diverse,

NOTE Confidence: 0.839534726

00:34:37.420 --> 00:34:39.420 some of them likely to be created by

NOTE Confidence: 0.839534726

00:34:39.420 --> 00:34:42.300 tobacco smoke, some of them less likely.

NOTE Confidence: 0.839534726

00:34:42.300 --> 00:34:43.848 OK, so we can look at this on an

NOTE Confidence: 0.839534726

00:34:43.848 --> 00:34:45.045 individual basis and then we can

NOTE Confidence: 0.839534726

00:34:45.045 --> 00:34:46.025 look at some other cancers.

NOTE Confidence: 0.839534726

00:34:46.030 --> 00:34:48.557 So here's bladder cancer and cervical cancer.

NOTE Confidence: 0.839534726

00:34:48.560 --> 00:34:51.400 I just added these because.

00:34:51.400 --> 00:34:53.584 Maybe this is a little less well known,

00:34:53.590 --> 00:34:55.982 but a lot of the mutation in both

00:34:55.982 --> 00:34:57.568 bladder cancer and cervical cancer

00:34:57.568 --> 00:34:59.710 is is caused at least by the.

00:34:59.710 --> 00:35:01.734 This deconvolution approach appears

00:35:01.734 --> 00:35:05.055 to be attributable to APOBEC mutation.

00:35:05.055 --> 00:35:08.980 Apobec is this apolipoprotein B.

00:35:08.980 --> 00:35:11.490 Gene that enzymatically we know

00:35:11.490 --> 00:35:14.300 mutates DNA and appears to be a

00:35:14.300 --> 00:35:15.380 viral defense protein.

00:35:15.380 --> 00:35:17.666 And what we see is that a lot of

00:35:17.666 --> 00:35:19.762 the mutagenic cause in the in

00:35:19.762 --> 00:35:22.100 the genome is created by apobec,

00:35:22.100 --> 00:35:24.067 some of it's by aging and bladder

00:35:24.067 --> 00:35:25.340 cancer and cervical cancer.

00:35:25.340 --> 00:35:27.713 There's a little bit of defective homologous

00:35:27.713 --> 00:35:29.920 recombination as a source there as well.

00:35:29.920 --> 00:35:31.608 But as you can see for a number

00:35:31.608 --> 00:35:32.620 of these mutations,

00:35:32.620 --> 00:35:37.104 the some P3 mutations for FGFR 3 for KSG 12D,

00:35:37.104 --> 00:35:39.218 we see almost no cause from APOBEC.

00:35:39.220 --> 00:35:39.940 But on the other hand,

00:35:39.940 --> 00:35:42.250 this other FGFR 3 mutation very

00:35:42.250 --> 00:35:44.630 likely to be caused by apobec,

00:35:44.630 --> 00:35:45.264 PI3 kinase,

00:35:45.264 --> 00:35:47.166 again very likely to be caused

00:35:47.166 --> 00:35:48.410 by APOBEC mutation.

00:35:48.410 --> 00:35:49.970 Cervical cancer are the same thing.

00:35:49.970 --> 00:35:50.368 All right,

00:35:50.368 --> 00:35:52.200 so we can look at the interval variance here.

00:35:52.200 --> 00:35:54.690 Let's get back to the main theme

NOTE Confidence: 0.839534726

00:35:54.690 --> 00:35:57.190 that this talk hopefully is.

NOTE Confidence: 0.839534726

00:35:57.190 --> 00:35:58.003 Presenting to you,

NOTE Confidence: 0.839534726

00:35:58.003 --> 00:36:00.201 which is that once we understand for every

NOTE Confidence: 0.839534726

00:36:00.201 --> 00:36:01.958 one of these variants what the causes

NOTE Confidence: 0.839534726

00:36:01.958 --> 00:36:04.088 are and how much they're causing cancer,

NOTE Confidence: 0.839534726

00:36:04.090 --> 00:36:06.478 we can then look at tumor

NOTE Confidence: 0.839534726

00:36:06.478 --> 00:36:08.070 causation by tumor type.

NOTE Confidence: 0.839534726

00:36:08.070 --> 00:36:10.059 And this isn't the best way to contrast them,

NOTE Confidence: 0.839534726

00:36:10.060 --> 00:36:11.242 I'll show you another that maybe

NOTE Confidence: 0.839534726

00:36:11.242 --> 00:36:12.480 contrast it a little bit better.

NOTE Confidence: 0.839534726

00:36:12.480 --> 00:36:15.014 But here we have all the different

NOTE Confidence: 0.839534726

00:36:15.014 --> 00:36:17.435 signatures on the Y axis and all

NOTE Confidence: 0.839534726

00:36:17.435 --> 00:36:19.319 the different cancers on the X

NOTE Confidence: 0.828217057666666

00:36:19.389 --> 00:36:21.901 axis and the red is the amount that

NOTE Confidence: 0.828217057666666

00:36:21.901 --> 00:36:24.830 the tumor type is caused by that

NOTE Confidence: 0.828217057666666

71

00:36:24.830 --> 00:36:27.130 particular signature and the Gray.

NOTE Confidence: 0.828217057666666

00:36:27.130 --> 00:36:29.392 It is or black is, the amount that you

NOTE Confidence: 0.828217057666666

00:36:29.392 --> 00:36:32.200 see mutation for due to that signature.

NOTE Confidence: 0.828217057666666

00:36:32.200 --> 00:36:33.500 And there's some big differences,

NOTE Confidence: 0.828217057666666

00:36:33.500 --> 00:36:36.461 say in signature 5 here for thyroid

NOTE Confidence: 0.828217057666666

00:36:36.461 --> 00:36:39.444 cancer where you see an enormous amount

NOTE Confidence: 0.828217057666666

00:36:39.444 --> 00:36:41.774 of cause but much less mutation.

NOTE Confidence: 0.828217057666666

00:36:41.774 --> 00:36:44.780 But it's a little hard to read that dot plot.

NOTE Confidence: 0.828217057666666

00:36:44.780 --> 00:36:46.820 Down below we have just these

NOTE Confidence: 0.828217057666666

00:36:46.820 --> 00:36:48.650 bar plots showing the can't,

NOTE Confidence: 0.828217057666666

00:36:48.650 --> 00:36:51.134 the weight of mutation.

NOTE Confidence: 0.828217057666666

00:36:51.134 --> 00:36:53.550 How much? Mutation was caught,

NOTE Confidence: 0.828217057666666

00:36:53.550 --> 00:36:56.648 which of the mutation in the genome was

NOTE Confidence: 0.828217057666666

00:36:56.648 --> 00:36:58.354 caused by a given mutational process.

NOTE Confidence: 0.828217057666666

00:36:58.354 --> 00:37:00.490 And on the right the effects and these

NOTE Confidence: 0.828217057666666

00:37:00.540 --> 00:37:02.122 may look pretty similar but let I'll

00:37:02.122 --> 00:37:03.914 show you the contrast that shows you

00:37:03.914 --> 00:37:05.516 how they're different in a moment.

00:37:05.520 --> 00:37:07.368 The thing I want to emphasize right

00:37:07.368 --> 00:37:09.460 now is we've given colors for all

00:37:09.460 --> 00:37:11.005 of those exogenous sources that

00:37:11.005 --> 00:37:12.805 may in principle be things that

00:37:12.805 --> 00:37:14.942 we could interfere on to stop.

00:37:14.942 --> 00:37:17.750 So UV light, defective,

00:37:17.750 --> 00:37:18.689 homologous recombination, presumably,

00:37:18.689 --> 00:37:21.180 maybe there be a way to do that,

00:37:21.180 --> 00:37:22.512 apobec perhaps if we.

00:37:22.512 --> 00:37:25.412 You know, avoided viral infection,

00:37:25.412 --> 00:37:28.388 tobacco certainly interventional alcohol,

00:37:28.390 --> 00:37:30.130 definitely something we can intervention on.

00:37:30.130 --> 00:37:31.574 Mutagenic chemical exposures definitely

00:37:31.574 --> 00:37:34.130 something we can do intervention on anyway.

NOTE Confidence: 0.828217057666666

00:37:34.130 --> 00:37:35.930 All those interventional ones are the

NOTE Confidence: 0.828217057666666

00:37:35.930 --> 00:37:38.205 colored ones and the aging ones are the

NOTE Confidence: 0.828217057666666

00:37:38.205 --> 00:37:40.300 Gray ones and then the unknown processes.

NOTE Confidence: 0.828217057666666

00:37:40.300 --> 00:37:42.244 The process is that we haven't figured out

NOTE Confidence: 0.828217057666666

00:37:42.244 --> 00:37:44.057 what they're associated with are in black.

NOTE Confidence: 0.828217057666666

00:37:44.060 --> 00:37:45.968 So this diagram actually tells you

NOTE Confidence: 0.828217057666666

00:37:45.968 --> 00:37:48.610 a lot about what you can do now

NOTE Confidence: 0.828217057666666

00:37:48.610 --> 00:37:50.240 to understand more about cancer,

NOTE Confidence: 0.828217057666666

00:37:50.240 --> 00:37:52.543 right because or to intervene we can

NOTE Confidence: 0.828217057666666

00:37:52.543 --> 00:37:55.186 intervene a lot on these cancer on these

NOTE Confidence: 0.828217057666666

00:37:55.186 --> 00:37:58.040 cancer types for which we see a lot of color.

NOTE Confidence: 0.828217057666666

00:37:58.040 --> 00:37:59.246 We there's much less we can

NOTE Confidence: 0.828217057666666

00:37:59.246 --> 00:38:00.519 do for the ones we don't.

NOTE Confidence: 0.828217057666666

00:38:00.520 --> 00:38:02.823 So for instance glioma very a lot

NOTE Confidence: 0.828217057666666

00:38:02.823 --> 00:38:05.460 of aging not a lot of other things,

NOTE Confidence: 0.828217057666666

00:38:05.460 --> 00:38:07.260 thyroid cancer, a lot of apobec,

NOTE Confidence: 0.828217057666666

00:38:07.260 --> 00:38:09.921 but other than that aging glioblastoma.

NOTE Confidence: 0.828217057666666

00:38:09.921 --> 00:38:13.182 Again, a lot of aging prostate cancer,

NOTE Confidence: 0.828217057666666

00:38:13.182 --> 00:38:14.286 a lot of aging,

NOTE Confidence: 0.828217057666666

00:38:14.290 --> 00:38:17.146 just a little bit of apobec and

NOTE Confidence: 0.828217057666666

00:38:17.146 --> 00:38:18.370 defective homologous recombination.

NOTE Confidence: 0.828217057666666

00:38:18.370 --> 00:38:20.197 So there's some we don't have much

NOTE Confidence: 0.828217057666666

00:38:20.197 --> 00:38:21.808 way to intervene on skin cancer,

NOTE Confidence: 0.828217057666666

00:38:21.810 --> 00:38:24.500 extremely easy to intervene to

NOTE Confidence: 0.828217057666666

00:38:24.500 --> 00:38:27.190 reduce the number of mutations,

NOTE Confidence: 0.828217057666666

00:38:27.190 --> 00:38:29.626 lung cancer, a lot of tobacco,

NOTE Confidence: 0.828217057666666

00:38:29.630 --> 00:38:30.818 a lot of defective,

NOTE Confidence: 0.828217057666666

00:38:30.818 --> 00:38:32.006 longest recombination and Applejack.

NOTE Confidence: 0.828217057666666

00:38:32.010 --> 00:38:33.792 So there's a lot we can do in terms

NOTE Confidence: 0.828217057666666

00:38:33.792 --> 00:38:35.713 of stopping those and then also where

NOTE Confidence: 0.828217057666666

00:38:35.713 --> 00:38:37.609 there's a lot more to understand.

NOTE Confidence: 0.828217057666666

00:38:37.610 --> 00:38:38.850 So for instance, breast cancer,

NOTE Confidence: 0.828217057666666

00:38:38.850 --> 00:38:40.440 ER 9 minus breast cancer.

NOTE Confidence: 0.828217057666666

00:38:40.440 --> 00:38:41.958 Like nearly half of the mutations,

NOTE Confidence: 0.828217057666666

00:38:41.960 --> 00:38:43.718 we don't know why they're being

NOTE Confidence: 0.828217057666666

00:38:43.718 --> 00:38:44.597 caused process wise.

NOTE Confidence: 0.828217057666666

00:38:44.600 --> 00:38:47.155 So this is something to be investigated

NOTE Confidence: 0.828217057666666

00:38:47.160 --> 00:38:48.238 because if we could figure it out,

NOTE Confidence: 0.828217057666666

00:38:48.240 --> 00:38:50.700 maybe there are these interminable

NOTE Confidence: 0.828217057666666

00:38:50.700 --> 00:38:52.296 processes that we could do something about.

NOTE Confidence: 0.902923374

00:38:54.630 --> 00:38:55.560 Etcetera. So you can sort

NOTE Confidence: 0.902923374

00:38:55.560 --> 00:38:56.490 of look at the black.

NOTE Confidence: 0.902923374

00:38:56.490 --> 00:38:58.138 That gives you an idea of how much

NOTE Confidence: 0.902923374

00:38:58.138 --> 00:38:59.870 we still need to learn and the Gray

NOTE Confidence: 0.902923374

00:38:59.870 --> 00:39:01.389 tells you and the idea of like,

NOTE Confidence: 0.902923374

00:39:01.390 --> 00:39:04.138 how much more.

76

NOTE Confidence: 0.902923374

00:39:04.140 --> 00:39:06.276 How much aging versus other processes

NOTE Confidence: 0.902923374

00:39:06.276 --> 00:39:08.919 seem to be causing that given cancer?

NOTE Confidence: 0.902923374

00:39:08.920 --> 00:39:10.630 And of course the cancers that

NOTE Confidence: 0.902923374

00:39:10.630 --> 00:39:12.083 are most age-related are at

NOTE Confidence: 0.902923374

00:39:12.083 --> 00:39:13.378 the bottom of this diagram,

NOTE Confidence: 0.902923374

00:39:13.380 --> 00:39:14.694 and the ones that are least

NOTE Confidence: 0.902923374

00:39:14.694 --> 00:39:16.059 age-related tend to be higher up.

NOTE Confidence: 0.852804418

00:39:18.160 --> 00:39:20.552 So this is just a bigger diagram of

NOTE Confidence: 0.852804418

00:39:20.552 --> 00:39:22.956 of that same picture in case Umm,

NOTE Confidence: 0.852804418

00:39:22.960 --> 00:39:24.590 you can see it better.

NOTE Confidence: 0.852804418

00:39:24.590 --> 00:39:26.137 And then I'm going to show you,

NOTE Confidence: 0.852804418

00:39:26.140 --> 00:39:27.352 I'm not going to show you

NOTE Confidence: 0.852804418

00:39:27.352 --> 00:39:28.160 the actual cancer types,

NOTE Confidence: 0.852804418

00:39:28.160 --> 00:39:29.990 but just an animation that actually

NOTE Confidence: 0.852804418

00:39:29.990 --> 00:39:31.565 Vincent Cantero made that helps

NOTE Confidence: 0.852804418

00:39:31.565 --> 00:39:33.065 you see the difference between

NOTE Confidence: 0.852804418

00:39:33.065 --> 00:39:34.720 the cancer mutation and effects.

NOTE Confidence: 0.852804418

00:39:34.720 --> 00:39:36.748 So this just varies between how

NOTE Confidence: 0.852804418

00:39:36.748 --> 00:39:38.772 much mutation is causing the given

NOTE Confidence: 0.852804418

00:39:38.772 --> 00:39:40.886 cancer and how much of the cancer

NOTE Confidence: 0.852804418

00:39:40.886 --> 00:39:42.913 affect by those mutations is causing

NOTE Confidence: 0.852804418

00:39:42.913 --> 00:39:45.036 the cancer and allows you to sort

NOTE Confidence: 0.852804418

00:39:45.036 --> 00:39:46.404 of see how different they are.

NOTE Confidence: 0.801817787142857

00:39:49.480 --> 00:39:51.699 You know, for ones like skin cancer,

NOTE Confidence: 0.801817787142857

00:39:51.700 --> 00:39:53.184 it doesn't change that much because nearly

NOTE Confidence: 0.801817787142857

00:39:53.184 --> 00:39:54.830 all the mutations are caused by UV anyway.

NOTE Confidence: 0.92253798

00:39:57.030 --> 00:39:59.850 All right. So as we said,

NOTE Confidence: 0.92253798

00:39:59.850 --> 00:40:00.694 as I said earlier,

NOTE Confidence: 0.92253798

00:40:00.694 --> 00:40:01.960 the extent to which the processes

NOTE Confidence: 0.92253798

00:40:02.010 --> 00:40:02.934 contribute determines tumor

NOTE Confidence: 0.92253798

00:40:02.934 --> 00:40:04.166 Genesis has been unknown.

NOTE Confidence: 0.92253798

00:40:04.170 --> 00:40:06.872 But now we can link it together

NOTE Confidence: 0.92253798

00:40:06.872 --> 00:40:08.030 with this process.

NOTE Confidence: 0.92253798

00:40:08.030 --> 00:40:09.678 And now I wanted to go back to

NOTE Confidence: 0.92253798

00:40:09.678 --> 00:40:11.127 this slide because I'm going to

NOTE Confidence: 0.92253798

00:40:11.127 --> 00:40:13.060 show you a bunch of diagrams and

NOTE Confidence: 0.92253798

00:40:13.060 --> 00:40:14.668 they're pretty complicated diagrams.

NOTE Confidence: 0.92253798

00:40:14.670 --> 00:40:16.446 But on the left hand side is going

NOTE Confidence: 0.92253798

00:40:16.446 --> 00:40:18.474 to be a bar plot that's respect

NOTE Confidence: 0.92253798

00:40:18.474 --> 00:40:20.440 reflecting like how much each process

NOTE Confidence: 0.92253798

00:40:20.440 --> 00:40:22.335 is contributing to the mutations.

NOTE Confidence: 0.92253798

00:40:22.340 --> 00:40:23.864 It's this left hand sign sign

NOTE Confidence: 0.92253798

00:40:23.864 --> 00:40:25.674 and on the right of each plot

NOTE Confidence: 0.92253798

00:40:25.674 --> 00:40:27.144 is going to be another bar.

NOTE Confidence: 0.92253798

00:40:27.150 --> 00:40:29.467 Plot that shows you how much each

NOTE Confidence: 0.92253798

00:40:29.467 --> 00:40:31.817 mutation is contributing to the increased

NOTE Confidence: 0.92253798

00:40:31.817 --> 00:40:33.545 cellular proliferation and survival.

NOTE Confidence: 0.92253798

00:40:33.550 --> 00:40:35.610 For four different cancers,

NOTE Confidence: 0.92253798

00:40:35.610 --> 00:40:37.670 here's primary skin cancer.

NOTE Confidence: 0.92253798

00:40:37.670 --> 00:40:42.128 Sorry, primary and metastatic skin cancer.

NOTE Confidence: 0.92253798

00:40:42.130 --> 00:40:45.230 Colorectal cancer.

NOTE Confidence: 0.92253798

00:40:45.230 --> 00:40:46.310 Actually this color,

NOTE Confidence: 0.92253798

00:40:46.310 --> 00:40:47.030 colon cancer,

NOTE Confidence: 0.92253798

00:40:47.030 --> 00:40:48.425 HPV negative head,

NOTE Confidence: 0.92253798

00:40:48.425 --> 00:40:51.438 neck cancer and thyroid cancer and the

NOTE Confidence: 0.92253798

00:40:51.438 --> 00:40:53.846 diagrams are this bar versus this bar.

NOTE Confidence: 0.92253798

00:40:53.850 --> 00:40:55.794 So the bar on the left is how

NOTE Confidence: 0.92253798

00:40:55.794 --> 00:40:57.934 much of a gift for a specific

NOTE Confidence: 0.92253798

00:40:57.934 --> 00:40:59.937 tumor was contributed by a given

NOTE Confidence: 0.92253798

00:40:59.937 --> 00:41:02.422 process and then how much of the

NOTE Confidence: 0.92253798

00:41:02.422 --> 00:41:04.073 oncogenesis for that tumor was

NOTE Confidence: 0.92253798

00:41:04.073 --> 00:41:05.628 caused by that particular process.

NOTE Confidence: 0.92253798

00:41:05.630 --> 00:41:07.905 And I've lined these up so that

NOTE Confidence: 0.92253798

00:41:07.905 --> 00:41:09.924 what I'm showing you is just

NOTE Confidence: 0.92253798

00:41:09.924 --> 00:41:11.874 five examples here from TCJ and

NOTE Confidence: 0.92253798

00:41:11.874 --> 00:41:14.029 this is actually from some data

NOTE Confidence: 0.92253798

00:41:14.030 --> 00:41:16.256 gathered here at Yale on Melanoma.

NOTE Confidence: 0.92253798

00:41:16.260 --> 00:41:17.100 But anyway,

NOTE Confidence: 0.92253798

00:41:17.100 --> 00:41:18.780 we've looked across these

NOTE Confidence: 0.92253798

00:41:18.780 --> 00:41:20.040 different cancer type,

NOTE Confidence: 0.92253798

00:41:20.040 --> 00:41:21.315 these different tumors and the

NOTE Confidence: 0.92253798

00:41:21.315 --> 00:41:22.590 question is are these different

NOTE Confidence: 0.92253798

00:41:22.635 --> 00:41:23.499 or are these similar?

NOTE Confidence: 0.92253798

00:41:23.500 --> 00:41:25.180 Like is the basic mutagenic effect

NOTE Confidence: 0.92253798

00:41:25.180 --> 00:41:27.123 and the cancer effect similar or is

NOTE Confidence: 0.92253798

00:41:27.123 --> 00:41:28.908 it different and you see they're very

NOTE Confidence: 0.92253798

00:41:28.965 --> 00:41:30.615 similar for two these two tumors,

NOTE Confidence: 0.92253798

00:41:30.620 --> 00:41:32.145 very similar for this third

NOTE Confidence: 0.92253798

00:41:32.145 --> 00:41:33.670 getting a little different here

NOTE Confidence: 0.92253798

00:41:33.730 --> 00:41:35.440 and getting quite different here.

NOTE Confidence: 0.92253798

00:41:35.440 --> 00:41:38.164 And these are arranged at the

NOTE Confidence: 0.92253798

00:41:38.164 --> 00:41:39.980 quartiles of the distribution.

NOTE Confidence: 0.92253798

00:41:39.980 --> 00:41:41.408 So it sort of represents the

NOTE Confidence: 0.92253798

00:41:41.408 --> 00:41:43.240 range of what you see in patients.

NOTE Confidence: 0.92253798

00:41:43.240 --> 00:41:44.956 So most of the time mutagenic

NOTE Confidence: 0.92253798

00:41:44.956 --> 00:41:46.100 effect and cancer causation.

NOTE Confidence: 0.92253798

00:41:46.100 --> 00:41:47.714 Are aligned very closely in primary

NOTE Confidence: 0.92253798

00:41:47.714 --> 00:41:49.695 skin cancer and that's because UV is

NOTE Confidence: 0.92253798

00:41:49.695 --> 00:41:51.110 causing almost all these mutations

NOTE Confidence: 0.92253798

00:41:51.110 --> 00:41:52.925 and changing things in colon cancer.

NOTE Confidence: 0.92253798

00:41:52.925 --> 00:41:54.755 As you extend from the more

NOTE Confidence: 0.92253798

00:41:54.755 --> 00:41:56.407 similar to the more different,

NOTE Confidence: 0.92253798

00:41:56.410 --> 00:41:58.312 you see a lot more heterogeneity

NOTE Confidence: 0.92253798

00:41:58.312 --> 00:42:00.521 from patient to patient in terms of

NOTE Confidence: 0.92253798

00:42:00.521 --> 00:42:02.285 whether or not the causative factors

NOTE Confidence: 0.92253798

00:42:02.285 --> 00:42:04.507 are the same as the myogenic factors.

NOTE Confidence: 0.92253798

00:42:04.510 --> 00:42:06.286 And that gets even more extreme

NOTE Confidence: 0.92253798

00:42:06.286 --> 00:42:07.470 with HPV negative head,

NOTE Confidence: 0.92253798

00:42:07.470 --> 00:42:10.085 neck cancer and even more

NOTE Confidence: 0.92253798

00:42:10.085 --> 00:42:12.177 extreme with thyroid cancer.

NOTE Confidence: 0.92253798

00:42:12.180 --> 00:42:15.414 So, but let me just emphasize again,

NOTE Confidence: 0.92253798

00:42:15.420 --> 00:42:17.130 these measures are for individual patients.

NOTE Confidence: 0.92253798

00:42:17.130 --> 00:42:19.404 So in principle this calculation can

NOTE Confidence: 0.92253798

00:42:19.404 --> 00:42:22.334 be done on any tumor sequence from

NOTE Confidence: 0.92253798

00:42:22.334 --> 00:42:25.010 an individual patient to tell you.

NOTE Confidence: 0.92253798

00:42:25.010 --> 00:42:27.652 What the causation of their cancer was,

NOTE Confidence: 0.92253798

00:42:27.652 --> 00:42:30.396 at least to the level that we

NOTE Confidence: 0.92253798

00:42:30.400 --> 00:42:31.486 are able to analyze this now,

NOTE Confidence: 0.92253798

00:42:31.490 --> 00:42:32.906 there's a bunch of things that

NOTE Confidence: 0.92253798

00:42:32.906 --> 00:42:34.652 are that we would love to also

NOTE Confidence: 0.92253798

00:42:34.652 --> 00:42:36.158 be able to incorporate into this.

NOTE Confidence: 0.92253798

00:42:36.160 --> 00:42:37.900 This is only single nucleotide mutations.

NOTE Confidence: 0.92253798

00:42:37.900 --> 00:42:39.490 It doesn't take into account

NOTE Confidence: 0.92253798

00:42:39.490 --> 00:42:40.444 copy number variation.

NOTE Confidence: 0.92253798

00:42:40.450 --> 00:42:41.755 It doesn't take into account

NOTE Confidence: 0.92253798

00:42:41.755 --> 00:42:42.277 epigenetic changes.

NOTE Confidence: 0.92253798

00:42:42.280 --> 00:42:43.596 And as I said at the outset,

NOTE Confidence: 0.8360394

00:42:43.600 --> 00:42:45.536 none of this has to do with physiological

NOTE Confidence: 0.8360394

00:42:45.536 --> 00:42:47.355 things like whether you exercise and have

NOTE Confidence: 0.8360394

00:42:47.355 --> 00:42:49.419 good autophagy in your you know it doesn't.

NOTE Confidence: 0.8360394

00:42:49.420 --> 00:42:50.508 It's not that physiological

NOTE Confidence: 0.8360394

00:42:50.508 --> 00:42:52.140 question of why you got cancer,

NOTE Confidence: 0.8360394

00:42:52.140 --> 00:42:53.375 but it is the mutagenic

NOTE Confidence: 0.8360394

00:42:53.375 --> 00:42:55.060 answer of why you got cancer.

NOTE Confidence: 0.8360394

00:42:55.060 --> 00:42:59.853 Down at the SMV level. Uh. And.

NOTE Confidence: 0.8360394

00:42:59.853 --> 00:43:03.692 So it reveals that and so that, so.

NOTE Confidence: 0.8360394

00:43:03.692 --> 00:43:05.386 So I think we're very good there.

NOTE Confidence: 0.8360394

00:43:05.390 --> 00:43:07.371 I would argue that the logic behind

NOTE Confidence: 0.8360394

00:43:07.371 --> 00:43:09.779 this is right and that we can apply

NOTE Confidence: 0.8360394

00:43:09.779 --> 00:43:11.485 that same logic to epigenetics to

NOTE Confidence: 0.8360394

00:43:11.485 --> 00:43:12.666 to copy number changes etcetera.

NOTE Confidence: 0.8360394

00:43:12.666 --> 00:43:14.220 There's just a lot of understanding.

NOTE Confidence: 0.8360394

00:43:14.220 --> 00:43:16.166 We still need of the basic underlying

NOTE Confidence: 0.8360394

00:43:16.166 --> 00:43:17.573 mutation rate for those things

NOTE Confidence: 0.8360394

00:43:17.573 --> 00:43:19.348 in order to actually do that and

NOTE Confidence: 0.8360394

00:43:19.348 --> 00:43:20.874 we're trying to work on that now.

NOTE Confidence: 0.873409281333333

00:43:23.100 --> 00:43:24.752 Now I just this, this is basically

NOTE Confidence: 0.873409281333333

00:43:24.752 --> 00:43:26.746 the the the end of the major talk but

NOTE Confidence: 0.873409281333333

00:43:26.746 --> 00:43:28.469 I just want to emphasize that this

NOTE Confidence: 0.873409281333333

00:43:28.469 --> 00:43:30.443 doesn't just apply to the origin of

NOTE Confidence: 0.873409281333333

00:43:30.450 --> 00:43:33.040 cancer in the early tumor genesis the

NOTE Confidence: 0.873409281333333

00:43:33.040 --> 00:43:35.867 same the same processes are going on

NOTE Confidence: 0.873409281333333

00:43:35.867 --> 00:43:38.286 in patients as we treat them as well.

NOTE Confidence: 0.873409281333333

00:43:38.290 --> 00:43:40.660 So there is you know so if you have a

NOTE Confidence: 0.873409281333333

00:43:40.729 --> 00:43:43.150 patient where you take out a a biopsy or

NOTE Confidence: 0.873409281333333

00:43:43.150 --> 00:43:45.720 a resection and then they undergo some

NOTE Confidence: 0.873409281333333

00:43:45.720 --> 00:43:47.856 sort of treatment and have recurrence

NOTE Confidence: 0.873409281333333

00:43:47.856 --> 00:43:50.117 there are ways to figure out exactly

NOTE Confidence: 0.873409281333333

00:43:50.117 --> 00:43:52.458 what the underlying processes that are

NOTE Confidence: 0.873409281333333

00:43:52.458 --> 00:43:54.413 contributing to the mutations that.

NOTE Confidence: 0.873409281333333

00:43:54.420 --> 00:43:55.664 Because that recurrence are.

NOTE Confidence: 0.873409281333333

00:43:55.664 --> 00:43:58.349 So that should be of interest to all of

NOTE Confidence: 0.873409281333333

00:43:58.349 --> 00:44:00.113 us who are interested in figuring out

NOTE Confidence: 0.873409281333333

00:44:00.174 --> 00:44:02.109 what's causing recurrence in cancer.

NOTE Confidence: 0.873409281333333

00:44:02.110 --> 00:44:03.630 So there's a clinical as well as the

00:44:03.630 --> 00:44:05.200 sort of more public health side that

00:44:05.200 --> 00:44:06.847 I was talking about with regard to

00:44:06.847 --> 00:44:08.221 these mutations and clinical side of

00:44:08.221 --> 00:44:10.090 how we might be able to apply this.

00:44:10.090 --> 00:44:12.820 And just to give you 2 material

00:44:12.820 --> 00:44:15.121 examples of this, here are two,

00:44:15.121 --> 00:44:18.144 I'm going to show you two sort of

00:44:18.144 --> 00:44:20.029 tree studies of individual patients.

00:44:20.030 --> 00:44:24.430 These were led by Nick Fisk in my lab.

00:44:24.430 --> 00:44:29.094 And here's a patient who was diagnosed

00:44:29.094 --> 00:44:31.879 with stage 3B lung cancer.

00:44:31.880 --> 00:44:34.624 They had an EGFR exon 19 deletion

00:44:34.624 --> 00:44:36.930 and their tumor was resected.

00:44:36.930 --> 00:44:39.540 They were given cisplatin and permatex

00:44:39.540 --> 00:44:41.600 bib and this there's a little pipe

00:44:41.600 --> 00:44:43.903 part and this is a phylogenetic tree

NOTE Confidence: 0.873409281333333

00:44:43.903 --> 00:44:45.833 relating their metastatic tumors to

NOTE Confidence: 0.873409281333333

00:44:45.833 --> 00:44:48.157 the primary tumor and it's been dated.

NOTE Confidence: 0.873409281333333

00:44:48.160 --> 00:44:49.476 We have all these techniques in my

NOTE Confidence: 0.873409281333333

00:44:49.476 --> 00:44:51.067 lab to date that based on the when

NOTE Confidence: 0.873409281333333

00:44:51.067 --> 00:44:52.573 the primary tumor was and how many

NOTE Confidence: 0.873409281333333

00:44:52.573 --> 00:44:53.832 mutations we see etcetera, etcetera.

NOTE Confidence: 0.873409281333333

00:44:53.832 --> 00:44:55.644 These were extracted at a later

NOTE Confidence: 0.873409281333333

00:44:55.644 --> 00:44:56.550 date than this,

NOTE Confidence: 0.873409281333333

00:44:56.550 --> 00:44:58.020 and so that gives us a way

NOTE Confidence: 0.873409281333333

00:44:58.020 --> 00:44:58.950 to calibrate the time.

NOTE Confidence: 0.873409281333333

00:44:58.950 --> 00:45:01.198 And what you see here in these pie

NOTE Confidence: 0.873409281333333

00:45:01.198 --> 00:45:03.487 charts is I've I've made it simpler.

NOTE Confidence: 0.873409281333333

00:45:03.490 --> 00:45:05.380 I just am looking at all other

NOTE Confidence: 0.873409281333333

00:45:05.380 --> 00:45:06.690 kinds of mutagenic sources.

NOTE Confidence: 0.873409281333333

00:45:06.690 --> 00:45:08.178 And one specific source

NOTE Confidence: 0.873409281333333

00:45:08.178 --> 00:45:09.666 that I'm interested in.

NOTE Confidence: 0.873409281333333

00:45:09.670 --> 00:45:10.690 And in this particular case,

NOTE Confidence: 0.873409281333333

00:45:10.690 --> 00:45:12.100 the source I'm interested in

NOTE Confidence: 0.873409281333333

00:45:12.100 --> 00:45:13.510 is the effect of cisplatin,

NOTE Confidence: 0.873409281333333

00:45:13.510 --> 00:45:16.222 which we know has a mutagenic

NOTE Confidence: 0.873409281333333

00:45:16.222 --> 00:45:18.659 effect on tumors even as it.

NOTE Confidence: 0.873409281333333

00:45:18.660 --> 00:45:21.271 The you know applies its own selective

NOTE Confidence: 0.873409281333333

00:45:21.271 --> 00:45:23.266 effect killing tumor cells and what

NOTE Confidence: 0.873409281333333

00:45:23.266 --> 00:45:25.677 you can see here is that the cisplatin

NOTE Confidence: 0.873409281333333

00:45:25.677 --> 00:45:28.329 mutations on this branch so this

NOTE Confidence: 0.873409281333333

00:45:28.329 --> 00:45:29.899 is independently determined right.

NOTE Confidence: 0.873409281333333

00:45:29.899 --> 00:45:31.333 This isn't because cisplatin is here

NOTE Confidence: 0.873409281333333

00:45:31.333 --> 00:45:33.213 we just did the deconvolution and boom

NOTE Confidence: 0.873409281333333

00:45:33.213 --> 00:45:34.881 here are all these despite mutations.

NOTE Confidence: 0.873409281333333

00:45:34.890 --> 00:45:36.810 This white pie piece here almost

NOTE Confidence: 0.873409281333333

00:45:36.810 --> 00:45:39.560 you know a bit less than 1/4 of or

NOTE Confidence: 0.873409281333333

00:45:39.560 --> 00:45:41.889 around 1/5 of the the mutations in

NOTE Confidence: 0.873409281333333

00:45:41.889 --> 00:45:44.463 this tumor are now cisplatin derived

NOTE Confidence: 0.873409281333333

00:45:44.463 --> 00:45:46.053 mutations and we can deconvolve

NOTE Confidence: 0.873409281333333

00:45:46.053 --> 00:45:47.740 that by doing this tree and seeing

NOTE Confidence: 0.873409281333333

00:45:47.786 --> 00:45:49.130 OK on this branch right here.

NOTE Confidence: 0.873409281333333

00:45:49.130 --> 00:45:51.186 That's how many are are that kind of

NOTE Confidence: 0.873409281333333

00:45:51.186 --> 00:45:53.122 mutation and then and then that of

NOTE Confidence: 0.873409281333333

00:45:53.122 --> 00:45:54.808 course that tumor continued to evolve

NOTE Confidence: 0.873409281333333

00:45:54.808 --> 00:45:56.677 and the reason it continued to evolve

NOTE Confidence: 0.873409281333333

00:45:56.677 --> 00:45:59.656 it was the patient was given or alot nib.

NOTE Confidence: 0.873409281333333

00:45:59.660 --> 00:46:01.155 Unfortunately a lot nib wasn't

NOTE Confidence: 0.873409281333333

00:46:01.155 --> 00:46:02.650 very successful in this case

NOTE Confidence: 0.788800622142857

00:46:02.705 --> 00:46:04.576 because they got the EGFR T790M

NOTE Confidence: 0.788800622142857

00:46:04.576 --> 00:46:07.144 resistance mutation on this branch as

NOTE Confidence: 0.788800622142857

00:46:07.144 --> 00:46:12.390 well and the tumor differentiated into

00:46:12.390 --> 00:46:15.394 these metastatic metastatic tumors and

00:46:15.394 --> 00:46:17.542 another metastatic tumor in the pancreas

00:46:17.542 --> 00:46:19.690 and and the point is here just that.

00:46:19.690 --> 00:46:21.415 The proportion of cisplatin was

00:46:21.415 --> 00:46:23.511 discontinued and so the proportion of

00:46:23.511 --> 00:46:25.081 mutations in subsequent branches is

00:46:25.081 --> 00:46:27.268 actually lower out of the total because

00:46:27.268 --> 00:46:29.319 new mutations are being added but they

00:46:29.320 --> 00:46:31.180 aren't cisplatin related mutations.

00:46:31.180 --> 00:46:33.998 So all of this deconvolution and

00:46:33.998 --> 00:46:36.438 understanding of the underlying mutagenic

00:46:36.438 --> 00:46:38.942 causation occurs during this treatment

00:46:38.942 --> 00:46:41.378 process that that that patients receive.

00:46:41.378 --> 00:46:43.580 And we can figure it out.

00:46:43.580 --> 00:46:44.370 There's one more point that

00:46:44.370 --> 00:46:45.370 I just want to make here,

NOTE Confidence: 0.788800622142857

00:46:45.370 --> 00:46:47.746 which is that it turns out,

NOTE Confidence: 0.788800622142857

00:46:47.750 --> 00:46:49.254 and I don't have a plot for this,

NOTE Confidence: 0.788800622142857

00:46:49.260 --> 00:46:51.794 but the T790M mutation is a mutation

NOTE Confidence: 0.788800622142857

00:46:51.794 --> 00:46:54.373 that is very likely to be caused

NOTE Confidence: 0.788800622142857

00:46:54.373 --> 00:46:56.497 just like those other ones I

NOTE Confidence: 0.788800622142857

00:46:56.580 --> 00:46:59.020 showed you by cisplatin mutation.

NOTE Confidence: 0.788800622142857

00:46:59.020 --> 00:47:01.694 So this is a poor ordering clinically

NOTE Confidence: 0.788800622142857

00:47:01.694 --> 00:47:04.095 for these treatments to be given

NOTE Confidence: 0.788800622142857

00:47:04.095 --> 00:47:05.675 because this despite mutation,

NOTE Confidence: 0.788800622142857

00:47:05.680 --> 00:47:07.462 creates a bunch of that genetic

NOTE Confidence: 0.788800622142857

00:47:07.462 --> 00:47:09.169 variation that is exactly what we

NOTE Confidence: 0.788800622142857

00:47:09.169 --> 00:47:10.772 don't want to have if we're going

NOTE Confidence: 0.788800622142857

00:47:10.772 --> 00:47:12.630 to put them on our lot and later.

NOTE Confidence: 0.788800622142857

00:47:12.630 --> 00:47:14.100 And very likely they had that mutation

NOTE Confidence: 0.788800622142857

00:47:14.100 --> 00:47:15.777 right when they were put on their lot nib,

92

00:47:15.780 --> 00:47:18.615 which is why there's a very little

00:47:18.615 --> 00:47:22.146 duration of of of benefit for the patient.

00:47:22.150 --> 00:47:23.690 So this is a great example for

00:47:23.690 --> 00:47:25.589 for in terms of a clinical or

00:47:25.589 --> 00:47:26.769 exogenous source of mutation,

00:47:26.770 --> 00:47:28.072 the cisplatin treatment

00:47:28.072 --> 00:47:29.808 that they were receiving.

00:47:29.810 --> 00:47:31.794 And let me give you another example that's

00:47:31.794 --> 00:47:33.880 about an endogenous change that has an

00:47:33.880 --> 00:47:35.525 interesting effect in a very similar way.

00:47:35.530 --> 00:47:38.146 So here's another lung cancer case.

00:47:38.150 --> 00:47:41.320 This patient was put on her right resection.

00:47:41.320 --> 00:47:43.795 They had a P53 mutation

00:47:43.795 --> 00:47:45.280 already after resection,

00:47:45.280 --> 00:47:47.088 but over a much longer period of time.

00:47:47.090 --> 00:47:48.358 They were never treated

NOTE Confidence: 0.788800622142857

00:47:48.358 --> 00:47:49.626 with cisplatin a much,

NOTE Confidence: 0.788800622142857

00:47:49.630 --> 00:47:51.989 much longer time later they did receive.

NOTE Confidence: 0.788800622142857

00:47:51.990 --> 00:47:55.934 They did get an ESR T790M mutation and

NOTE Confidence: 0.788800622142857

00:47:55.934 --> 00:47:58.662 you can see these plots are solid here

NOTE Confidence: 0.788800622142857

00:47:58.662 --> 00:48:00.692 meaning that the mutational process of

NOTE Confidence: 0.788800622142857

00:48:00.692 --> 00:48:03.265 interest that I wanted to talk about you

NOTE Confidence: 0.788800622142857

00:48:03.265 --> 00:48:05.400 know hasn't happened at all here yet.

NOTE Confidence: 0.788800622142857

00:48:05.400 --> 00:48:07.533 And then you can see unfortunately

NOTE Confidence: 0.788800622142857

00:48:07.533 --> 00:48:09.304 later on they were they were moved

NOTE Confidence: 0.788800622142857

00:48:09.304 --> 00:48:10.667 to Avastin or not unfortunately

NOTE Confidence: 0.788800622142857

00:48:10.667 --> 00:48:12.293 movement they were moved to Avastin.

NOTE Confidence: 0.788800622142857

00:48:12.300 --> 00:48:14.500 It wasn't unfortunate necessarily

NOTE Confidence: 0.788800622142857

00:48:14.500 --> 00:48:16.700 but Erlotinib was discontinued.

NOTE Confidence: 0.788800622142857

00:48:16.700 --> 00:48:19.478 They were given permatex have been

NOTE Confidence: 0.788800622142857

00:48:19.478 --> 00:48:22.319 carboplatin late in latent therapy but.

NOTE Confidence: 0.788800622142857

00:48:22.320 --> 00:48:24.102 But the divergence,

NOTE Confidence: 0.788800622142857

00:48:24.102 --> 00:48:26.478 but they're metastatic tumors

NOTE Confidence: 0.788800622142857

00:48:26.478 --> 00:48:28.576 started genetically diverging about

NOTE Confidence: 0.788800622142857

00:48:28.576 --> 00:48:30.156 two years before their death.

NOTE Confidence: 0.788800622142857

00:48:30.160 --> 00:48:32.437 And the thing that I just want to emphasize

NOTE Confidence: 0.788800622142857

00:48:32.437 --> 00:48:35.258 is here is a continuum B1 mutation S37C,

NOTE Confidence: 0.788800622142857

00:48:35.258 --> 00:48:39.248 which is known to induce.

NOTE Confidence: 0.788800622142857

00:48:39.250 --> 00:48:39.390 Yeah.

NOTE Confidence: 0.790621141724138

00:48:41.970 --> 00:48:43.650 Defects in homologous recombination

NOTE Confidence: 0.790621141724138

00:48:43.650 --> 00:48:46.986 based mutations and you can see in the

NOTE Confidence: 0.790621141724138

00:48:46.986 --> 00:48:48.866 descendant lineages the increased amount

NOTE Confidence: 0.790621141724138

00:48:48.866 --> 00:48:51.815 of that kind of mutation occurring after

NOTE Confidence: 0.790621141724138

00:48:51.815 --> 00:48:54.010 continuing 1B stinging B1 mutation.

NOTE Confidence: 0.790621141724138

00:48:54.010 --> 00:48:56.054 So this is an endogenous process that

NOTE Confidence: 0.790621141724138

00:48:56.054 --> 00:48:58.310 was started by a mutation that we can

NOTE Confidence: 0.790621141724138

00:48:58.310 --> 00:49:00.396 then track again down to the individual

NOTE Confidence: 0.790621141724138

00:49:00.396 --> 00:49:02.520 branch where the mutations are occurring

NOTE Confidence: 0.790621141724138

00:49:02.520 --> 00:49:04.490 and how many of them are caused.

NOTE Confidence: 0.790621141724138

00:49:04.490 --> 00:49:06.884 And then from then on there's a lot of

NOTE Confidence: 0.790621141724138

00:49:06.884 --> 00:49:09.069 cutting and B1 mutation in these tumors,

NOTE Confidence: 0.790621141724138

00:49:09.070 --> 00:49:10.280 but not in the spleen.

NOTE Confidence: 0.790621141724138

00:49:10.280 --> 00:49:10.637 Interesting.

NOTE Confidence: 0.790621141724138

00:49:10.637 --> 00:49:12.422 That's an interesting story here

NOTE Confidence: 0.790621141724138

00:49:12.422 --> 00:49:14.890 is that this could be 1 mutation

NOTE Confidence: 0.790621141724138

00:49:14.890 --> 00:49:17.589 occurred and led to all the metastases

NOTE Confidence: 0.790621141724138

00:49:17.589 --> 00:49:19.774 to all these other tissues.

NOTE Confidence: 0.790621141724138

00:49:19.780 --> 00:49:21.322 The one tissue that had a

NOTE Confidence: 0.790621141724138

00:49:21.322 --> 00:49:22.686 metastasis that was not continuing

NOTE Confidence: 0.790621141724138

00:49:22.686 --> 00:49:24.618 to be 1 mutated was the spleen,

NOTE Confidence: 0.790621141724138

00:49:24.620 --> 00:49:27.260 which is very interesting because Canadian

NOTE Confidence: 0.790621141724138

00:49:27.260 --> 00:49:29.020 B1 mutation causes vascularization.

NOTE Confidence: 0.790621141724138

00:49:29.020 --> 00:49:30.420 The spleen is already quite

NOTE Confidence: 0.790621141724138

00:49:30.420 --> 00:49:30.980 highly vascularized,

NOTE Confidence: 0.790621141724138

00:49:30.980 --> 00:49:32.108 so it may not have been

NOTE Confidence: 0.790621141724138

00:49:32.108 --> 00:49:32.860 needed for the spleen,

NOTE Confidence: 0.790621141724138

00:49:32.860 --> 00:49:35.009 whereas it may have been more important

NOTE Confidence: 0.790621141724138

00:49:35.009 --> 00:49:37.559 to the cancer for the rest of the tumors.

NOTE Confidence: 0.790621141724138

00:49:37.560 --> 00:49:38.128 All right,

NOTE Confidence: 0.790621141724138

00:49:38.128 --> 00:49:40.116 I've sort of gone through all of

NOTE Confidence: 0.790621141724138

00:49:40.116 --> 00:49:41.826 what I wanted to talk about.

NOTE Confidence: 0.790621141724138

00:49:41.830 --> 00:49:43.874 Today, in terms of introducing you to

NOTE Confidence: 0.790621141724138

00:49:43.874 --> 00:49:46.599 this way of actually trying to understand,

NOTE Confidence: 0.790621141724138

00:49:46.600 --> 00:49:47.466 you know,

NOTE Confidence: 0.790621141724138

00:49:47.466 --> 00:49:49.692 why an individual tumor is

NOTE Confidence: 0.790621141724138

00:49:49.692 --> 00:49:51.660 has been made oncogenic.

NOTE Confidence: 0.790621141724138

00:49:51.660 --> 00:49:53.420 I hope that I've at least been able

NOTE Confidence: 0.790621141724138

00:49:53.420 --> 00:49:55.188 to argue that the logic behind what

NOTE Confidence: 0.790621141724138

00:49:55.188 --> 00:49:56.997 we're doing is sound and that the

NOTE Confidence: 0.790621141724138

00:49:56.997 --> 00:49:58.917 process that we're doing is a sound way

NOTE Confidence: 0.790621141724138

00:49:58.920 --> 00:50:00.859 of sort of attributing that cancer effect,

NOTE Confidence: 0.790621141724138

00:50:00.860 --> 00:50:02.165 at least as regards those

NOTE Confidence: 0.790621141724138

00:50:02.165 --> 00:50:02.948 single nucleotide variants,

NOTE Confidence: 0.790621141724138

00:50:02.950 --> 00:50:04.980 which are what we mostly focus on.

NOTE Confidence: 0.790621141724138

00:50:04.980 --> 00:50:06.858 But where we don't, you know,

NOTE Confidence: 0.790621141724138

00:50:06.860 --> 00:50:08.976 we don't know whether that's 10% or 90%

NOTE Confidence: 0.790621141724138

00:50:08.976 --> 00:50:11.832 of the reason why genetically tumors are.

NOTE Confidence: 0.790621141724138

00:50:11.840 --> 00:50:12.788 Because we don't know that yet.

NOTE Confidence: 0.790621141724138

00:50:12.790 --> 00:50:13.802 But regardless,

NOTE Confidence: 0.790621141724138

00:50:13.802 --> 00:50:17.850 if we look at that single nucleotide effect,

NOTE Confidence: 0.790621141724138

00:50:17.850 --> 00:50:19.670 we now can sort of deconvolve that.

NOTE Confidence: 0.790621141724138

00:50:19.670 --> 00:50:21.890 And I'm very curious if anyone

NOTE Confidence: 0.790621141724138

00:50:21.890 --> 00:50:24.009 has thoughts to share with me.

NOTE Confidence: 0.790621141724138

00:50:24.010 --> 00:50:27.349 You know what how this information could

NOTE Confidence: 0.790621141724138

00:50:27.349 --> 00:50:29.550 be used for the benefit of patients,

NOTE Confidence: 0.790621141724138

00:50:29.550 --> 00:50:30.710 for the knowledge of patients,

NOTE Confidence: 0.790621141724138

00:50:30.710 --> 00:50:33.054 but also as I mentioned in the later

NOTE Confidence: 0.790621141724138

00:50:33.054 --> 00:50:35.272 part of my talk in in understanding

NOTE Confidence: 0.790621141724138

00:50:35.272 --> 00:50:37.820 better what our therapies are doing to

NOTE Confidence: 0.790621141724138

00:50:37.890 --> 00:50:39.955 patients over time as well and ways

NOTE Confidence: 0.790621141724138

00:50:39.955 --> 00:50:42.669 that we can ideally order our therapies.

NOTE Confidence: 0.790621141724138

00:50:42.670 --> 00:50:45.614 So that we avoid the evolution of the

NOTE Confidence: 0.790621141724138

00:50:45.614 --> 00:50:47.957 resistance that we're trying to avoid,

NOTE Confidence: 0.790621141724138

00:50:47.960 --> 00:50:50.078 which is so clearly evidenced in

NOTE Confidence: 0.790621141724138

00:50:50.078 --> 00:50:52.583 that one case with the cisplatin

NOTE Confidence: 0.790621141724138

00:50:52.583 --> 00:50:54.119 and relative treatment.

NOTE Confidence: 0.790621141724138

00:50:54.120 --> 00:50:56.076 Thanks very much for your attention.

NOTE Confidence: 0.790621141724138

00:50:56.080 --> 00:50:57.778 I want to thank very much.

NOTE Confidence: 0.790621141724138

00:50:57.780 --> 00:50:59.628 Vincent Cantaro who was a postdoc in my

NOTE Confidence: 0.790621141724138

00:50:59.628 --> 00:51:01.959 lab is now a professor at Emmanuel College,

NOTE Confidence: 0.790621141724138

00:51:01.960 --> 00:51:04.780 remains a a vibrant and

NOTE Confidence: 0.790621141724138

00:51:04.780 --> 00:51:07.036 wonderful collaborator of mine.

NOTE Confidence: 0.790621141724138

00:51:07.040 --> 00:51:08.960 I really like working with him.

NOTE Confidence: 0.790621141724138

00:51:08.960 --> 00:51:10.640 It's been incredibly productive to continue

NOTE Confidence: 0.790621141724138

00:51:10.640 --> 00:51:13.110 to do so and I'm really delighted that.

NOTE Confidence: 0.790621141724138

00:51:13.110 --> 00:51:15.594 He's been able to do so despite a very

NOTE Confidence: 0.790621141724138

00:51:15.594 --> 00:51:18.070 heavy teaching load at Emmanuel College.

NOTE Confidence: 0.790621141724138

00:51:18.070 --> 00:51:20.688 Jeff Mandel over here in the audience,

NOTE Confidence: 0.790621141724138

00:51:20.690 --> 00:51:22.440 graduate student working on the

NOTE Confidence: 0.790621141724138

00:51:22.440 --> 00:51:23.840 cancer effect size calculations

NOTE Confidence: 0.790621141724138

00:51:23.840 --> 00:51:25.807 and the machinery underlying that.

NOTE Confidence: 0.790621141724138

00:51:25.810 --> 00:51:27.602 Nick Fisk who's worked on a lot

NOTE Confidence: 0.790621141724138

00:51:27.602 --> 00:51:28.370 of the tree

NOTE Confidence: 0.866059358

00:51:28.437 --> 00:51:29.927 based analysis in my lab,

NOTE Confidence: 0.866059358

00:51:29.930 --> 00:51:31.848 including those last ones that I mentioned.

NOTE Confidence: 0.866059358

00:51:31.850 --> 00:51:33.500 Everyone else in the town's lab,

NOTE Confidence: 0.866059358

00:51:33.500 --> 00:51:34.460 they're also great.

NOTE Confidence: 0.866059358

00:51:34.460 --> 00:51:36.380 I didn't specifically mention their work.

NOTE Confidence: 0.866059358

00:51:36.380 --> 00:51:38.084 We've got interesting work on epistasis

NOTE Confidence: 0.866059358

00:51:38.084 --> 00:51:40.126 and all kinds of other things that

NOTE Confidence: 0.866059358

00:51:40.126 --> 00:51:41.830 that I really think is outstanding

NOTE Confidence: 0.866059358

00:51:41.830 --> 00:51:43.640 and should be really interesting.

NOTE Confidence: 0.866059358

00:51:43.640 --> 00:51:45.160 But so all the Members,

NOTE Confidence: 0.866059358

00:51:45.160 --> 00:51:46.100 it's a great group.

NOTE Confidence: 0.866059358

00:51:46.100 --> 00:51:48.039 Also I want to thank the NIH, NIDCR,

NOTE Confidence: 0.866059358

00:51:48.039 --> 00:51:49.852 Yale support and head and neck is

NOTE Confidence: 0.866059358

00:51:49.852 --> 00:51:51.508 a great community and I really

NOTE Confidence: 0.866059358

00:51:51.508 --> 00:51:53.907 enjoy being part of it and and it

NOTE Confidence: 0.866059358

00:51:53.907 --> 00:51:55.652 also provides a substantial amount

NOTE Confidence: 0.866059358

00:51:55.652 --> 00:51:58.069 of my funding for cancer work.

NOTE Confidence: 0.866059358

00:51:58.070 --> 00:51:59.498 So thank you very much and I

NOTE Confidence: 0.866059358

00:51:59.498 --> 00:52:00.997 would love to take any questions

NOTE Confidence: 0.866059358

00:52:00.997 --> 00:52:02.497 or hear any thoughts, comments,

NOTE Confidence: 0.866059358

00:52:02.497 --> 00:52:04.819 etcetera from anyone in the audience.

NOTE Confidence: 0.749719333571429

00:52:11.060 --> 00:52:13.076 Jeff, I think you actually have

NOTE Confidence: 0.749719333571429

00:52:13.076 --> 00:52:15.366 an amazing talent to make really

NOTE Confidence: 0.749719333571429

00:52:15.366 --> 00:52:16.978 complex things, explain them in

NOTE Confidence: 0.749719333571429

00:52:16.978 --> 00:52:19.025 a very simple and logical way.

NOTE Confidence: 0.749719333571429

00:52:19.025 --> 00:52:21.560 Do we have a microphone so,

NOTE Confidence: 0.749719333571429

00:52:21.560 --> 00:52:23.604 so that people could actually ask questions?

NOTE Confidence: 0.78279528

00:52:27.170 --> 00:52:28.050 Oh, OK.

NOTE Confidence: 0.66808236

00:52:36.840 --> 00:52:38.966 OK. Amazing work.

NOTE Confidence: 0.66808236

00:52:38.966 --> 00:52:41.701 Jeff I just wanted to ask if you could

NOTE Confidence: 0.66808236

00:52:41.701 --> 00:52:43.516 address how you handle commutations

NOTE Confidence: 0.66808236

00:52:43.516 --> 00:52:46.229 when you look at cancer effect size.

00:52:46.230 --> 00:52:47.808 And you know I'm thinking of

00:52:47.808 --> 00:52:49.536 this finding that we haven't had

00:52:49.536 --> 00:52:51.478 neck cancer where P53 mutation is

00:52:51.478 --> 00:52:53.830 truncating if you don't have CDK into

00:52:53.896 --> 00:52:56.280 a mutated and when you when you have

00:52:56.280 --> 00:52:58.586 a mutation in the DNA binding domain,

00:52:58.590 --> 00:53:00.790 it seems like you need the second mutation.

00:53:00.790 --> 00:53:04.350 So how do you handle that and then I guess

00:53:04.441 --> 00:53:07.350 also when you have P53 mutations or.

00:53:07.350 --> 00:53:10.110 CDK into a mutations or whatever.

00:53:10.110 --> 00:53:13.379 Does that muddy your signatures at all?

00:53:13.380 --> 00:53:18.340 You know to the to the mutational effects of.

00:53:18.340 --> 00:53:20.290 Losing control of of cell cycle

00:53:20.290 --> 00:53:21.265 and DNA repair.

00:53:22.780 --> 00:53:24.748 This is a little more straightforward.

00:53:24.750 --> 00:53:26.870 Does it muddy things when things get changed?

NOTE Confidence: 0.851854498333333

00:53:26.870 --> 00:53:29.014 Yes, it does in the sense that there

NOTE Confidence: 0.851854498333333

00:53:29.014 --> 00:53:31.215 is a temporal difference, right,

NOTE Confidence: 0.851854498333333

00:53:31.215 --> 00:53:32.925 between what was happening before that

NOTE Confidence: 0.851854498333333

00:53:32.925 --> 00:53:34.929 change happened and what happened afterwards.

NOTE Confidence: 0.851854498333333

00:53:34.930 --> 00:53:37.606 And our resolution for understanding those

NOTE Confidence: 0.851854498333333

00:53:37.606 --> 00:53:39.856 temporal differences is somewhat weak, right.

NOTE Confidence: 0.851854498333333

00:53:39.856 --> 00:53:41.720 So generally, if we have a lot of

NOTE Confidence: 0.851854498333333

00:53:41.773 --> 00:53:43.749 samples like we had in those two cases,

NOTE Confidence: 0.851854498333333

00:53:43.750 --> 00:53:45.838 we can sort of piece apart when things

NOTE Confidence: 0.851854498333333

00:53:45.838 --> 00:53:48.164 happened in a nice way and we'll be able

NOTE Confidence: 0.851854498333333

00:53:48.164 --> 00:53:50.068 to understand those sorts of differences.

NOTE Confidence: 0.851854498333333

00:53:50.070 --> 00:53:51.862 But when we're just looking at tumor

NOTE Confidence: 0.851854498333333

00:53:51.862 --> 00:53:53.220 Genesis to resection and we have.

NOTE Confidence: 0.851854498333333

00:53:53.220 --> 00:53:55.306 This association then we have to use

NOTE Confidence: 0.851854498333333

00:53:55.306 --> 00:53:57.628 very large numbers to get sort of

NOTE Confidence: 0.851854498333333

00:53:57.628 --> 00:53:58.996 statistical associations to understand

NOTE Confidence: 0.851854498333333

00:53:58.996 --> 00:54:00.820 that sort of ordering process.

NOTE Confidence: 0.851854498333333

00:54:00.820 --> 00:54:04.204 Which gets me to your second question or

NOTE Confidence: 0.851854498333333

00:54:04.204 --> 00:54:07.796 your first question of of commutation which

NOTE Confidence: 0.851854498333333

00:54:07.800 --> 00:54:09.438 I have a strong opinion on everything.

NOTE Confidence: 0.851854498333333

00:54:09.440 --> 00:54:12.372 So my strong opinion on this is that that

NOTE Confidence: 0.851854498333333

00:54:12.372 --> 00:54:14.902 the general genomics approach towards

NOTE Confidence: 0.851854498333333

00:54:14.902 --> 00:54:18.009 looking at commutation is is flawed in a

NOTE Confidence: 0.851854498333333

00:54:18.009 --> 00:54:20.512 way that is not apparent when you read

NOTE Confidence: 0.851854498333333

00:54:20.512 --> 00:54:22.972 all the papers on it and the argument that.

NOTE Confidence: 0.851854498333333

00:54:22.972 --> 00:54:24.134 I want to make is that when

NOTE Confidence: 0.851854498333333

00:54:24.134 --> 00:54:25.068 you look at commutation,

NOTE Confidence: 0.851854498333333

00:54:25.070 --> 00:54:26.500 you're typically looking at a

NOTE Confidence: 0.851854498333333

00:54:26.500 --> 00:54:27.358 very observational thing,

NOTE Confidence: 0.851854498333333

00:54:27.360 --> 00:54:29.187 which is like how often is this

NOTE Confidence: 0.851854498333333

00:54:29.187 --> 00:54:30.978 one mutated and this one mutated.

00:54:30.980 --> 00:54:33.808 So for the same reasons that I

00:54:33.808 --> 00:54:36.100 outlined in my talk today,

00:54:36.100 --> 00:54:38.392 that there are two reasons why

00:54:38.392 --> 00:54:41.032 you see things mutated, you know,

00:54:41.032 --> 00:54:43.256 underlying mutation or selection.

00:54:43.260 --> 00:54:44.988 There are two reasons why things

00:54:44.988 --> 00:54:45.852 might be commutated.

00:54:45.860 --> 00:54:47.135 They might be commutated because

00:54:47.135 --> 00:54:48.155 when you get one,

00:54:48.160 --> 00:54:49.952 the other one is selected and it

00:54:49.952 --> 00:54:51.667 really creates a great benefit to

00:54:51.667 --> 00:54:53.443 the cell to survive and replicate.

00:54:53.450 --> 00:54:55.700 The other reason might be because

00:54:55.700 --> 00:54:58.544 they both have the same underlying

00:54:58.544 --> 00:54:59.808 mutational process.

NOTE Confidence: 0.851854498333333

00:54:59.810 --> 00:55:01.244 And when you have four orders

NOTE Confidence: 0.851854498333333

00:55:01.244 --> 00:55:02.494 of magnitude of difference in

NOTE Confidence: 0.851854498333333

00:55:02.494 --> 00:55:03.739 mutational process site to site,

NOTE Confidence: 0.851854498333333

00:55:03.740 --> 00:55:05.658 that can be a very big reason

NOTE Confidence: 0.851854498333333

00:55:05.658 --> 00:55:07.350 why you see commutation.

NOTE Confidence: 0.851854498333333

00:55:07.350 --> 00:55:09.121 So commutation is not the signature we

NOTE Confidence: 0.851854498333333

00:55:09.121 --> 00:55:11.854 like it to be to say these things are

NOTE Confidence: 0.851854498333333

00:55:11.854 --> 00:55:13.134 selected together because sometimes

NOTE Confidence: 0.851854498333333

00:55:13.189 --> 00:55:15.142 they may not be even though they're

NOTE Confidence: 0.851854498333333

00:55:15.142 --> 00:55:18.250 strongly come come mutated in a data set.

NOTE Confidence: 0.851854498333333

00:55:18.250 --> 00:55:19.586 So then how do I deal with it?

NOTE Confidence: 0.851854498333333

00:55:19.590 --> 00:55:19.850 Well,

NOTE Confidence: 0.851854498333333

00:55:19.850 --> 00:55:21.670 we can take all of the approaches

NOTE Confidence: 0.851854498333333

00:55:21.670 --> 00:55:23.857 I told you and we're working on

NOTE Confidence: 0.851854498333333

00:55:23.857 --> 00:55:25.467 you know even more sophisticated

NOTE Confidence: 0.851854498333333

107

00:55:25.525 --> 00:55:27.387 approaches now to try to do this.

NOTE Confidence: 0.851854498333333

00:55:27.390 --> 00:55:28.766 I think I have some slides on it,

NOTE Confidence: 0.851854498333333

00:55:28.770 --> 00:55:30.525 so I would love to take the time to.

NOTE Confidence: 0.851854498333333

00:55:30.530 --> 00:55:32.900 Just quickly introduce it since

NOTE Confidence: 0.851854498333333

00:55:32.900 --> 00:55:35.900 they're they're way down here though.

NOTE Confidence: 0.851854498333333

00:55:35.900 --> 00:55:37.470 Ah.

NOTE Confidence: 0.851854498333333

00:55:37.470 --> 00:55:37.666 Yeah.

NOTE Confidence: 0.851854498333333

00:55:37.666 --> 00:55:39.038 So this is the point that you

NOTE Confidence: 0.851854498333333

00:55:39.038 --> 00:55:40.807 that I was just answering to you

NOTE Confidence: 0.851854498333333

00:55:40.807 --> 00:55:42.481 which is mutual exclusivity and Co

NOTE Confidence: 0.851854498333333

00:55:42.481 --> 00:55:43.971 occurrence are patterns that are

NOTE Confidence: 0.851854498333333

00:55:43.971 --> 00:55:45.453 caused by either commutation or

NOTE Confidence: 0.851854498333333

00:55:45.453 --> 00:55:46.868 what I call selective epistasis.

NOTE Confidence: 0.851854498333333

00:55:46.870 --> 00:55:48.781 Again I'm using the terminology from my

NOTE Confidence: 0.851854498333333

00:55:48.781 --> 00:55:49.970 background and evolutionary biology.

NOTE Confidence: 0.851854498333333

00:55:49.970 --> 00:55:51.825 Epistasis meaning 1 gene is having an

00:55:51.825 --> 00:55:53.572 effect on another or the mutation in

00:55:53.572 --> 00:55:55.510 one gene is having an effect on another.

00:55:55.510 --> 00:55:57.135 So typical approaches have not

00:55:57.135 --> 00:55:58.110 acknowledged the possibility

00:55:58.110 --> 00:55:58.760 of commutation

00:55:58.810 --> 00:56:00.784 which is a common underlying mutational bias.

00:56:00.790 --> 00:56:02.288 That's what I just said to you.

00:56:02.290 --> 00:56:03.620 This is a typical slide from I

00:56:03.620 --> 00:56:05.051 don't mean to be you know casting

00:56:05.051 --> 00:56:06.563 aspersion on this as I said this

00:56:06.563 --> 00:56:07.788 is what everyone pretty much.

00:56:07.790 --> 00:56:10.436 Does but but they look for whether

00:56:10.436 --> 00:56:12.065 cancers have sequential mutations

00:56:12.065 --> 00:56:14.609 developed or commutation but we can

00:56:14.609 --> 00:56:17.027 actually take those same analysis same

00:56:17.027 --> 00:56:20.005 the same data and and and deconvolve with

NOTE Confidence: 0.779470155882353

00:56:20.005 --> 00:56:21.865 some fairly sophisticated mathematics

NOTE Confidence: 0.779470155882353

00:56:21.865 --> 00:56:24.840 that Jorge Alfaro Murillo and I did on

NOTE Confidence: 0.779470155882353

00:56:24.840 --> 00:56:26.400 the fluxes mutation rates and scale.

NOTE Confidence: 0.779470155882353

00:56:26.400 --> 00:56:28.402 So selection coefficients for up to five

NOTE Confidence: 0.779470155882353

00:56:28.402 --> 00:56:30.421 genes and look at what the likelihood

NOTE Confidence: 0.779470155882353

00:56:30.421 --> 00:56:32.405 of individual genes are are to get

NOTE Confidence: 0.779470155882353

00:56:32.405 --> 00:56:34.061 mutated what the likelihood Karras is

NOTE Confidence: 0.779470155882353

00:56:34.061 --> 00:56:35.886 going to be muted after P53 etcetera.

NOTE Confidence: 0.779470155882353

00:56:35.886 --> 00:56:37.920 So we can look at all of these.

NOTE Confidence: 0.779470155882353

00:56:37.920 --> 00:56:40.128 Figure out how frequently those happen.

NOTE Confidence: 0.779470155882353

00:56:40.130 --> 00:56:41.594 So this is the flux which is a

NOTE Confidence: 0.779470155882353

00:56:41.594 --> 00:56:42.550 measure of commutation,

NOTE Confidence: 0.779470155882353

00:56:42.550 --> 00:56:44.054 essentially the underlying mutation

NOTE Confidence: 0.779470155882353

00:56:44.054 --> 00:56:46.310 rates and then the scaled selection

NOTE Confidence: 0.779470155882353

00:56:46.363 --> 00:56:48.108 coefficient for the new mutation.

NOTE Confidence: 0.779470155882353

00:56:48.110 --> 00:56:50.126 So these are how likely is P53,

NOTE Confidence: 0.779470155882353

00:56:50.130 --> 00:56:53.188 how likely is KSB mutated after PHP 53

NOTE Confidence: 0.779470155882353

00:56:53.188 --> 00:56:55.740 and then how likely is it how selected

NOTE Confidence: 0.779470155882353

00:56:55.812 --> 00:56:58.004 is it to have KRS after PMI 50?

NOTE Confidence: 0.779470155882353

00:56:58.010 --> 00:56:59.366 P 53 is a separate measure,

NOTE Confidence: 0.779470155882353

00:56:59.370 --> 00:57:01.442 so we can basically take all of those

NOTE Confidence: 0.779470155882353

00:57:01.442 --> 00:57:03.349 and look at all of those different

NOTE Confidence: 0.779470155882353

00:57:03.349 --> 00:57:05.129 things for up to five or six.

NOTE Confidence: 0.779470155882353

00:57:05.130 --> 00:57:06.435 And again there are constraint

NOTE Confidence: 0.779470155882353

00:57:06.435 --> 00:57:08.030 is usually the amount of data.

NOTE Confidence: 0.779470155882353

00:57:08.030 --> 00:57:09.984 We need massive amounts of data

NOTE Confidence: 0.779470155882353

00:57:09.984 --> 00:57:10.680 to understand,

NOTE Confidence: 0.779470155882353

00:57:10.680 --> 00:57:13.256 like 3 way effects or four way effects.

NOTE Confidence: 0.779470155882353

00:57:13.260 --> 00:57:15.724 So you need to have examples of every

NOTE Confidence: 0.779470155882353

00:57:15.724 --> 00:57:17.564 possible combination in that data set

NOTE Confidence: 0.779470155882353

00:57:17.564 --> 00:57:19.316 and that rapidly exhausts our samples.

NOTE Confidence: 0.779470155882353

00:57:19.320 --> 00:57:20.208 But on their hand,

NOTE Confidence: 0.779470155882353

00:57:20.208 --> 00:57:21.540 we're getting a lot more data

NOTE Confidence: 0.779470155882353

00:57:21.589 --> 00:57:23.141 now and so we're able to do this

NOTE Confidence: 0.779470155882353

00:57:23.141 --> 00:57:24.460 with more and more data sets.

NOTE Confidence: 0.779470155882353

00:57:24.460 --> 00:57:26.124 Now this is lung cancer and we were

NOTE Confidence: 0.779470155882353

00:57:26.124 --> 00:57:27.906 able to do it for these five genes.

NOTE Confidence: 0.662026182

00:57:30.240 --> 00:57:32.508 P53 KSDK, 11 RL, RP1B and.

NOTE Confidence: 0.662026182

00:57:32.508 --> 00:57:34.348 And figure out all their

NOTE Confidence: 0.662026182

00:57:34.348 --> 00:57:35.820 relations with each other.

NOTE Confidence: 0.662026182

00:57:35.820 --> 00:57:37.374 This is maybe an easier way to

NOTE Confidence: 0.662026182

00:57:37.374 --> 00:57:38.798 see this instead of a big table,

NOTE Confidence: 0.662026182

00:57:38.800 --> 00:57:40.040 which is just what's the

NOTE Confidence: 0.662026182

00:57:40.040 --> 00:57:41.032 evolutionary trajectory of them.

NOTE Confidence: 0.662026182

00:57:41.040 --> 00:57:42.528 And again, this is all epistatic,

NOTE Confidence: 0.662026182

00:57:42.530 --> 00:57:44.588 like it's all taking into account

NOTE Confidence: 0.662026182

00:57:44.588 --> 00:57:46.170 that commutation factor and the

NOTE Confidence: 0.662026182

00:57:46.170 --> 00:57:48.095 width of the bar is the flux,

NOTE Confidence: 0.662026182

00:57:48.100 --> 00:57:50.011 or how frequently you go from normal

NOTE Confidence: 0.662026182

00:57:50.011 --> 00:57:52.640 to say P53 in this particular case,

NOTE Confidence: 0.662026182

00:57:52.640 --> 00:57:55.514 or LPV one or K Ras or SDK 111.

NOTE Confidence: 0.662026182

00:57:55.514 --> 00:57:58.139 And then you can see that if you KS isn't

NOTE Confidence: 0.662026182

00:57:58.139 --> 00:58:00.477 actually that frequent as a first mutation,

NOTE Confidence: 0.662026182

00:58:00.480 --> 00:58:01.296 but if you do get it,

NOTE Confidence: 0.662026182

00:58:01.300 --> 00:58:04.140 then you're very likely to get LRP 1B.

NOTE Confidence: 0.662026182

00:58:04.140 --> 00:58:05.370 Or SDK 11.

NOTE Confidence: 0.662026182

00:58:05.370 --> 00:58:06.474 If you get P53,

NOTE Confidence: 0.662026182

00:58:06.474 --> 00:58:07.854 you're very likely to then

NOTE Confidence: 0.662026182

00:58:07.854 --> 00:58:09.586 get LRP 1B as well.

NOTE Confidence: 0.662026182

00:58:09.586 --> 00:58:11.770 You're you're you know some probability,

NOTE Confidence: 0.662026182

00:58:11.770 --> 00:58:13.186 but it's not so high of getting curious.

NOTE Confidence: 0.662026182

00:58:13.190 --> 00:58:15.989 After that you're very likely to get a KRS

NOTE Confidence: 0.662026182

00:58:15.989 --> 00:58:18.497 mutation if you have P53 and LRP we want to.

NOTE Confidence: 0.662026182

00:58:18.500 --> 00:58:21.050 One LRP 1B together et cetera.

NOTE Confidence: 0.662026182

00:58:21.050 --> 00:58:22.868 So you can you can look at what the

NOTE Confidence: 0.662026182

00:58:22.868 --> 00:58:24.765 likely trajectory for a given patient is.

NOTE Confidence: 0.662026182

00:58:24.770 --> 00:58:26.210 You could even look at where they are on

NOTE Confidence: 0.662026182

00:58:26.210 --> 00:58:27.686 this trajectory and we haven't done this,

NOTE Confidence: 0.662026182

00:58:27.690 --> 00:58:29.482 but presumably you can figure out what

NOTE Confidence: 0.662026182

00:58:29.482 --> 00:58:31.034 their prognosis was based on where

NOTE Confidence: 0.662026182

00:58:31.034 --> 00:58:32.504 they were on this diagram etcetera.

NOTE Confidence: 0.662026182

00:58:32.510 --> 00:58:35.021 And we have basically a a map of what's

NOTE Confidence: 0.662026182

00:58:35.021 --> 00:58:37.347 actually happening to these these patients.

NOTE Confidence: 0.662026182

00:58:37.350 --> 00:58:39.286 And then down below in the smaller diagrams,

NOTE Confidence: 0.662026182

00:58:39.290 --> 00:58:40.778 I've just divided this up because

NOTE Confidence: 0.662026182

00:58:40.778 --> 00:58:42.490 this is all the fluxes again,

NOTE Confidence: 0.662026182

00:58:42.490 --> 00:58:44.394 but let's divide it up into mutation

NOTE Confidence: 0.662026182

00:58:44.394 --> 00:58:45.527 rates and selection coefficients

NOTE Confidence: 0.662026182

00:58:45.527 --> 00:58:47.396 and what you see is the mutation

NOTE Confidence: 0.662026182

00:58:47.396 --> 00:58:48.710 rates are here are quite.

NOTE Confidence: 0.662026182

00:58:48.710 --> 00:58:50.774 Symmetrical because we haven't

NOTE Confidence: 0.662026182

00:58:50.774 --> 00:58:52.838 accounted for things like.

NOTE Confidence: 0.662026182

00:58:52.840 --> 00:58:53.995 Containing 1B mutation,

NOTE Confidence: 0.662026182

00:58:53.995 --> 00:58:55.920 changing the mutation rate etcetera.

NOTE Confidence: 0.662026182

00:58:55.920 --> 00:58:57.152 In this particular analysis,

NOTE Confidence: 0.662026182

00:58:57.152 --> 00:58:59.000 although in principle we can do

NOTE Confidence: 0.662026182

00:58:59.054 --> 00:59:01.336 that and then on the right are

NOTE Confidence: 0.662026182

00:59:01.336 --> 00:59:03.130 so there's a LRP 1B particularly

NOTE Confidence: 0.662026182

00:59:03.130 --> 00:59:04.630 has a very high mutation rate.

NOTE Confidence: 0.662026182

00:59:04.630 --> 00:59:05.845 So it's relatively high frequency

NOTE Confidence: 0.662026182

00:59:05.845 --> 00:59:07.350 is not that big a deal,

NOTE Confidence: 0.662026182

00:59:07.350 --> 00:59:08.628 although it does seem to have

NOTE Confidence: 0.662026182

00:59:08.628 --> 00:59:09.720 some selective effect as well.

NOTE Confidence: 0.662026182

00:59:09.720 --> 00:59:11.169 And then over here we see the

NOTE Confidence: 0.662026182

00:59:11.169 --> 00:59:12.516 selective effects and you can see

NOTE Confidence: 0.662026182

00:59:12.516 --> 00:59:13.656 there's very strong selection for

NOTE Confidence: 0.662026182

00:59:13.660 --> 00:59:16.198 P53 initially is the major selection

NOTE Confidence: 0.662026182

00:59:16.200 --> 00:59:19.126 and yet that exists after LRP 1B

NOTE Confidence: 0.662026182

00:59:19.126 --> 00:59:23.910 as well but after after P53 or.

NOTE Confidence: 0.662026182

00:59:23.910 --> 00:59:25.054 LRP 1B and P33,

NOTE Confidence: 0.662026182

00:59:25.054 --> 00:59:26.770 then we're very likely to get

NOTE Confidence: 0.662026182

00:59:26.837 --> 00:59:28.440 this Karas mutation, etcetera.

NOTE Confidence: 0.662026182

00:59:28.440 --> 00:59:30.750 So you can really understand what the

NOTE Confidence: 0.662026182

00:59:30.750 --> 00:59:32.520 relative effect of each of these is.

NOTE Confidence: 0.435518738

00:59:33.970 --> 00:59:37.390 Trajectories after the sample size.

NOTE Confidence: 0.435518738

00:59:37.390 --> 00:59:38.398 That's a good question.

NOTE Confidence: 0.860814934615385

00:59:39.310 --> 00:59:40.806 I haven't done the study that I'd like

NOTE Confidence: 0.860814934615385

00:59:40.806 --> 00:59:42.794 to do to answer that, which would be

00:59:42.794 --> 00:59:44.642 like do some very massive analysis.

00:59:44.650 --> 00:59:46.722 It's actually a lot of computation to like

00:59:46.722 --> 00:59:48.948 do that 1000 times subsampling etcetera.

00:59:48.950 --> 00:59:51.480 But what I have done is just do the analysis,

00:59:51.480 --> 00:59:53.320 you know, with one data set and then

00:59:53.320 --> 00:59:55.313 add more data sets and it seems

00:59:55.313 --> 00:59:56.793 quite stable from that perspective.

00:59:56.800 --> 00:59:58.258 That's not really the same because

00:59:58.258 --> 00:59:59.455 we're not subtracting out the

00:59:59.455 --> 01:00:00.694 first data set when we do that.

01:00:00.700 --> 01:00:03.337 But but it's not like it varies all over

01:00:03.337 --> 01:00:05.914 the place and the stability of course is

01:00:05.914 --> 01:00:07.448 proportional to the prevalence, right?

01:00:07.448 --> 01:00:08.560 Of that particular mutation,

01:00:08.560 --> 01:00:09.970 the mutations that are really

01:00:09.970 --> 01:00:11.020 highly prevalence, you know,

NOTE Confidence: 0.860814934615385

01:00:11.020 --> 01:00:13.231 they stay very stable because we've got a lot

NOTE Confidence: 0.860814934615385

01:00:13.231 --> 01:00:15.119 of examples of them with the other genes.

NOTE Confidence: 0.860814934615385

01:00:15.120 --> 01:00:16.688 As soon as you get the lower prevalence,

NOTE Confidence: 0.860814934615385

01:00:16.690 --> 01:00:18.550 it's it's a lot iffier.

NOTE Confidence: 0.860814934615385

01:00:18.550 --> 01:00:18.819 So.

NOTE Confidence: 0.860814934615385

01:00:18.819 --> 01:00:20.971 So really this can only be used right

NOTE Confidence: 0.860814934615385

01:00:20.971 --> 01:00:23.316 now for these for the most prevalent

NOTE Confidence: 0.860814934615385

01:00:23.316 --> 01:00:25.199 kinds of mutations that you see.

NOTE Confidence: 0.860814934615385

01:00:25.200 --> 01:00:27.378 And typically we are for instance

NOTE Confidence: 0.860814934615385

01:00:27.378 --> 01:00:29.715 assembling all the mutations in a given

NOTE Confidence: 0.860814934615385

01:00:29.715 --> 01:00:31.640 gene as one kind of mutation because

NOTE Confidence: 0.860814934615385

01:00:31.701 --> 01:00:33.805 we need that sample size to do that,

NOTE Confidence: 0.860814934615385

01:00:33.810 --> 01:00:35.462 which is something that in my other

NOTE Confidence: 0.860814934615385

01:00:35.462 --> 01:00:37.291 research I usually avoid because I think

NOTE Confidence: 0.860814934615385

01:00:37.291 --> 01:00:38.893 it's really important to understand it.

NOTE Confidence: 0.860814934615385

01:00:38.900 --> 01:00:40.900 Different sites have different effects.

NOTE Confidence: 0.860814934615385

01:00:40.900 --> 01:00:41.170 So

NOTE Confidence: 0.813187945714286

01:00:42.020 --> 01:00:44.348 one thing that that I didn't

NOTE Confidence: 0.813187945714286

01:00:44.348 --> 01:00:46.072 see certain probably this.

NOTE Confidence: 0.813187945714286

01:00:46.072 --> 01:00:49.180 So you can calculate an additional process

NOTE Confidence: 0.813187945714286

01:00:49.252 --> 01:00:51.832 contribution to to the privatization

NOTE Confidence: 0.813187945714286

01:00:51.832 --> 01:00:53.896 in particular individual cases.

NOTE Confidence: 0.813187945714286

01:00:53.900 --> 01:00:55.972 But what happens if you caused the

NOTE Confidence: 0.813187945714286

01:00:55.972 --> 01:00:57.464 cases and obviously you should be able

NOTE Confidence: 0.813187945714286

01:00:57.464 --> 01:00:59.184 to sell it off lung cancers related

NOTE Confidence: 0.813187945714286

01:00:59.184 --> 01:01:01.277 to smoking and those who don't and

NOTE Confidence: 0.813187945714286

01:01:01.277 --> 01:01:02.853 that would be a trial thing to do.

NOTE Confidence: 0.813187945714286

01:01:02.860 --> 01:01:04.684 But could you do the same and create

NOTE Confidence: 0.813187945714286

01:01:04.684 --> 01:01:06.866 a new classification for example for

NOTE Confidence: 0.813187945714286

01:01:06.866 --> 01:01:08.663 initial cancer, breast cancer that

NOTE Confidence: 0.813187945714286

01:01:08.663 --> 01:01:10.910 are going to aging and the other?

NOTE Confidence: 0.813187945714286

01:01:10.910 --> 01:01:12.450 By looking at them separately,

NOTE Confidence: 0.813187945714286

01:01:12.450 --> 01:01:14.898 you might get some idea about

NOTE Confidence: 0.813187945714286

01:01:14.898 --> 01:01:16.122 what's actually causing.

NOTE Confidence: 0.813187945714286

01:01:16.130 --> 01:01:17.460 The.

NOTE Confidence: 0.813187945714286

01:01:17.460 --> 01:01:18.890 The Unknown edition signature.

NOTE Confidence: 0.88755983

01:01:21.850 --> 01:01:23.326 Yeah, I definitely think you

NOTE Confidence: 0.88755983

01:01:23.326 --> 01:01:25.458 could cluster them. I think you

NOTE Confidence: 0.877339712142857

01:01:25.470 --> 01:01:28.175 know the you're reducing the

NOTE Confidence: 0.877339712142857

01:01:28.175 --> 01:01:31.314 dimensionality of the data when you

NOTE Confidence: 0.877339712142857

01:01:31.314 --> 01:01:33.845 go from the raw data back to the

NOTE Confidence: 0.877339712142857

01:01:33.845 --> 01:01:36.227 processes and so you have a reduced

NOTE Confidence: 0.877339712142857

01:01:36.227 --> 01:01:38.207 dimensionality of that raw data.

NOTE Confidence: 0.877339712142857

01:01:38.210 --> 01:01:39.634 And then you're and then if you were

NOTE Confidence: 0.877339712142857

01:01:39.634 --> 01:01:40.828 to cluster on the basis of this,

NOTE Confidence: 0.877339712142857

01:01:40.830 --> 01:01:42.096 you would be taking that reduced

NOTE Confidence: 0.877339712142857

01:01:42.096 --> 01:01:43.171 dimensionality data and trying to

NOTE Confidence: 0.877339712142857

01:01:43.171 --> 01:01:44.146 say does that predict something.

NOTE Confidence: 0.877339712142857

01:01:44.150 --> 01:01:46.094 So I I think from a machine learning

NOTE Confidence: 0.877339712142857

01:01:46.094 --> 01:01:47.712 standpoint you might want to just go

NOTE Confidence: 0.877339712142857

01:01:47.712 --> 01:01:49.409 back to that broad data in some way,

NOTE Confidence: 0.877339712142857

01:01:49.410 --> 01:01:50.598 but there might be some way

NOTE Confidence: 0.877339712142857

01:01:50.598 --> 01:01:51.390 of thinking about it.

NOTE Confidence: 0.877339712142857

01:01:51.390 --> 01:01:52.182 That I say that,

NOTE Confidence: 0.877339712142857

01:01:52.182 --> 01:01:53.370 but then I also think there's

NOTE Confidence: 0.877339712142857

01:01:53.420 --> 01:01:54.370 a second part of that,

NOTE Confidence: 0.877339712142857

01:01:54.370 --> 01:01:56.854 which is that I do think you do better

NOTE Confidence: 0.877339712142857

01:01:56.854 --> 01:01:59.050 looking at actual biological processes,

NOTE Confidence: 0.877339712142857

01:01:59.050 --> 01:02:00.796 even if it involves some reduction

NOTE Confidence: 0.877339712142857

01:02:00.796 --> 01:02:01.669 of the data,

NOTE Confidence: 0.877339712142857

01:02:01.670 --> 01:02:03.026 because it simplifies the data in

NOTE Confidence: 0.877339712142857

01:02:03.026 --> 01:02:04.844 a way that means you don't go off

NOTE Confidence: 0.877339712142857

01:02:04.844 --> 01:02:06.146 on these random tangents of all

NOTE Confidence: 0.877339712142857

01:02:06.192 --> 01:02:07.728 the noisy stuff you're looking at.

NOTE Confidence: 0.877339712142857

01:02:07.730 --> 01:02:08.990 So, so there's, there's a,

NOTE Confidence: 0.877339712142857

01:02:08.990 --> 01:02:10.775 I guess there's a tension I think

NOTE Confidence: 0.877339712142857

01:02:10.775 --> 01:02:12.786 you should be wary of in doing that,

NOTE Confidence: 0.877339712142857

01:02:12.790 --> 01:02:14.286 but I don't see any reason you couldn't

NOTE Confidence: 0.877339712142857

01:02:14.286 --> 01:02:15.781 do that and and it would probably

NOTE Confidence: 0.877339712142857

01:02:15.781 --> 01:02:17.140 be highly predictive in some cases.

NOTE Confidence: 0.877339712142857

01:02:17.140 --> 01:02:18.605 You're probably going to see

NOTE Confidence: 0.877339712142857

01:02:18.605 --> 01:02:20.070 most skin cancers very easily,

NOTE Confidence: 0.877339712142857

01:02:20.070 --> 01:02:21.390 you know, predictive that way.

NOTE Confidence: 0.877339712142857

01:02:21.390 --> 01:02:24.558 Because they're just UV all over the place.

NOTE Confidence: 0.877339712142857

01:02:24.560 --> 01:02:26.440 Some other cancers are probably

NOTE Confidence: 0.877339712142857

01:02:26.440 --> 01:02:28.761 quite hard to distinguish one from

NOTE Confidence: 0.877339712142857

01:02:28.761 --> 01:02:31.017 the other just by the mutational

NOTE Confidence: 0.877339712142857

01:02:31.017 --> 01:02:32.950 processes that underlie their cause,

NOTE Confidence: 0.877339712142857

01:02:32.950 --> 01:02:35.694 and so I could imagine doing that.

NOTE Confidence: 0.877339712142857

01:02:35.700 --> 01:02:36.870 We haven't done anything like that.

NOTE Confidence: 0.7372582

01:02:40.800 --> 01:02:41.710 Any other comments?

NOTE Confidence: 0.634007122857143

01:02:43.180 --> 01:02:45.098 To ask questions, then the audience on

NOTE Confidence: 0.763244012

01:02:46.070 --> 01:02:47.980 there, there was, I thought,

NOTE Confidence: 0.763244012

01:02:47.980 --> 01:02:51.340 Q&amp;A, but I there it is.

NOTE Confidence: 0.763244012

01:02:51.340 --> 01:02:52.950 We have time. Ohh. Yeah.

NOTE Confidence: 0.763244012

01:02:52.950 --> 01:02:53.940 We've got some questions here,

NOTE Confidence: 0.763244012

01:02:53.940 --> 01:02:55.140 but maybe one more for you and then

NOTE Confidence: 0.763244012

01:02:55.140 --> 01:02:56.320 I'll go to the online questions.

NOTE Confidence: 0.763244012

01:02:56.320 --> 01:02:57.840 Yes, OK. Thank you.

NOTE Confidence: 0.763244012

01:02:57.840 --> 01:02:59.999 Thanks, Jeff. Fantastic work.

NOTE Confidence: 0.77099422

01:03:02.760 --> 01:03:05.818 I think your methodology is on the right

NOTE Confidence: 0.77099422

01:03:05.818 --> 01:03:09.482 track and nothing to worry about at all.

NOTE Confidence: 0.77099422

123

01:03:09.482 --> 01:03:11.926 The opposite is true.

NOTE Confidence: 0.77099422

01:03:11.930 --> 01:03:14.895 My only concern is availability

NOTE Confidence: 0.77099422

01:03:14.895 --> 01:03:17.860 of data in the future,

NOTE Confidence: 0.77099422

01:03:17.860 --> 01:03:21.400 especially for new types of cancers.

NOTE Confidence: 0.77099422

01:03:21.400 --> 01:03:23.716 Are we asking the right questions?

NOTE Confidence: 0.77099422

01:03:23.720 --> 01:03:27.176 Are we collecting the right data?

NOTE Confidence: 0.77099422

01:03:27.180 --> 01:03:34.140 Be meaning human as humans. And.

NOTE Confidence: 0.77099422

01:03:34.140 --> 01:03:36.546 I'd like us humans to ensure

NOTE Confidence: 0.77099422

01:03:36.546 --> 01:03:39.120 that this data is available,

NOTE Confidence: 0.77099422

01:03:39.120 --> 01:03:44.419 it's it's open source and it's reliable

NOTE Confidence: 0.77099422

01:03:44.420 --> 01:03:47.318 and what are your thoughts on that?

NOTE Confidence: 0.77099422

01:03:47.320 --> 01:03:48.880 Yeah, so that's a great question.

NOTE Confidence: 0.77099422

01:03:48.880 --> 01:03:50.941 I mean I think that the volume of data

NOTE Confidence: 0.77099422

01:03:50.941 --> 01:03:52.793 sets on like tumor Genesis for section

NOTE Confidence: 0.77099422

01:03:52.793 --> 01:03:55.083 kind of data is going to increase very

NOTE Confidence: 0.77099422

01:03:55.083 --> 01:03:57.347 well on its own like we don't need

NOTE Confidence: 0.77099422

01:03:57.347 --> 01:03:59.800 to pay attention to that question.

NOTE Confidence: 0.77099422

01:03:59.800 --> 01:04:01.834 The the datasets that I think I would like

NOTE Confidence: 0.77099422

01:04:01.834 --> 01:04:04.119 to see more of are these multi sample data.

NOTE Confidence: 0.77099422

01:04:04.120 --> 01:04:05.748 That's from individual patients.

NOTE Confidence: 0.77099422

01:04:05.748 --> 01:04:06.969 Back in 2016,

NOTE Confidence: 0.77099422

01:04:06.970 --> 01:04:08.906 I was lucky to be funded by Gilead

NOTE Confidence: 0.77099422

01:04:08.906 --> 01:04:10.470 to actually sequence these large

NOTE Confidence: 0.77099422

01:04:10.470 --> 01:04:12.180 numbers of metastatic and primary

NOTE Confidence: 0.77099422

01:04:12.180 --> 01:04:14.127 tumors and they were really there.

NOTE Confidence: 0.77099422

01:04:14.130 --> 01:04:15.815 The potential of those data

NOTE Confidence: 0.77099422

01:04:15.815 --> 01:04:17.163 sets is really high,

NOTE Confidence: 0.77099422

01:04:17.170 --> 01:04:18.595 especially if they have a

NOTE Confidence: 0.77099422

01:04:18.595 --> 01:04:19.450 clinical annotations alongside.

NOTE Confidence: 0.77099422

01:04:19.450 --> 01:04:21.304 So you can map it to to understand what

NOTE Confidence: 0.77099422

01:04:21.304 --> 01:04:23.006 was happening for the patient at the

NOTE Confidence: 0.77099422

01:04:23.006 --> 01:04:25.270 same time as what was happening genetically.

NOTE Confidence: 0.77099422

01:04:25.270 --> 01:04:27.100 That data set though was

NOTE Confidence: 0.77099422

01:04:27.100 --> 01:04:28.907 heterogeneous by cancer type, right?

NOTE Confidence: 0.77099422

01:04:28.907 --> 01:04:31.266 And I haven't seen similar sized data

NOTE Confidence: 0.77099422

01:04:31.266 --> 01:04:34.149 sets on individual cancer types gathered.

NOTE Confidence: 0.77099422

01:04:34.150 --> 01:04:35.640 And it's not, you know,

NOTE Confidence: 0.77099422

01:04:35.640 --> 01:04:37.656 it's a lot of money like it's a couple

NOTE Confidence: 0.77099422

01:04:37.656 --> 01:04:38.940 $1,000,000 to do that sequencing,

NOTE Confidence: 0.77099422

01:04:38.940 --> 01:04:41.268 but you could do that for

NOTE Confidence: 0.77099422

01:04:41.268 --> 01:04:42.820 every cancer type for.

NOTE Confidence: 0.77099422

01:04:42.820 --> 01:04:43.380 You know,

NOTE Confidence: 0.77099422

01:04:43.380 --> 01:04:45.060 $30 million or something like that.

NOTE Confidence: 0.77099422

01:04:45.060 --> 01:04:46.740 And I think that would be so worth

NOTE Confidence: 0.77099422

01:04:46.740 --> 01:04:48.592 it because we would learn so much

NOTE Confidence: 0.77099422

01:04:48.592 --> 01:04:49.972 about the evolutionary trajectory of

NOTE Confidence: 0.77099422

01:04:50.020 --> 01:04:51.756 each of these cancer types by looking

NOTE Confidence: 0.77099422

01:04:51.756 --> 01:04:53.340 at multi sample data like that.

NOTE Confidence: 0.77099422

01:04:53.340 --> 01:04:55.020 But I haven't managed to sort

NOTE Confidence: 0.77099422

01:04:55.020 --> 01:04:56.877 of put together the argument to

NOTE Confidence: 0.77099422

01:04:56.877 --> 01:04:58.392 get funding to do that.

NOTE Confidence: 0.77099422

01:04:58.400 --> 01:05:02.930 I encourage you to elevate that, you know.

NOTE Confidence: 0.77099422

01:05:02.930 --> 01:05:07.790 Definitely. Yeah. Thanks. Like.

NOTE Confidence: 0.77099422

01:05:07.790 --> 01:05:11.956 Just. Comments. I'm sorry.

NOTE Confidence: 0.77099422

01:05:11.956 --> 01:05:13.644 OK. First, I enjoy your talk.

NOTE Confidence: 0.77099422

01:05:13.644 --> 01:05:14.178 Thank you.

NOTE Confidence: 0.868388082

01:05:14.220 --> 01:05:17.110 But I'm not so sure.

NOTE Confidence: 0.868388082

01:05:17.110 --> 01:05:19.590 Given the tumor heterogeneity.

NOTE Confidence: 0.868388082

01:05:19.590 --> 01:05:22.074 Your math, just the tumor cell.

NOTE Confidence: 0.868388082

01:05:22.074 --> 01:05:23.629 We don't even talk about

NOTE Confidence: 0.868388082

01:05:23.629 --> 01:05:25.090 the microenvironment.

NOTE Confidence: 0.868388082

01:05:25.090 --> 01:05:29.200 Math sequence will really be useful.

NOTE Confidence: 0.868388082

01:05:29.200 --> 01:05:32.090 With all the other tools.

NOTE Confidence: 0.868388082

01:05:32.090 --> 01:05:34.010 You know, otherwise you're going to.

NOTE Confidence: 0.868388082

01:05:34.010 --> 01:05:34.836 For instance,

NOTE Confidence: 0.868388082

01:05:34.836 --> 01:05:37.314 you just mentioned about the cluster.

NOTE Confidence: 0.868388082

01:05:37.320 --> 01:05:38.580 Approach.

NOTE Confidence: 0.868388082

01:05:38.580 --> 01:05:41.380 You can have a mutation in different

NOTE Confidence: 0.868388082

01:05:41.380 --> 01:05:43.750 tumor cells within the tumor mass.

NOTE Confidence: 0.87285068

01:05:45.800 --> 01:05:47.450 When you do the analysis,

NOTE Confidence: 0.87285068

01:05:47.450 --> 01:05:48.630 you put them all together.

NOTE Confidence: 0.8229486875

01:05:50.740 --> 01:05:51.748 Does that make sense?

NOTE Confidence: 0.839879968947368

01:05:53.490 --> 01:05:54.633 I think I might need to talk to you

NOTE Confidence: 0.839879968947368

01:05:54.633 --> 01:05:56.004 at more length to sort of fully

NOTE Confidence: 0.839879968947368

01:05:56.004 --> 01:05:57.624 understand your question, but but I

NOTE Confidence: 0.839879968947368

01:05:57.624 --> 01:06:00.460 guess what I would comment is just that.

NOTE Confidence: 0.839879968947368

01:06:00.460 --> 01:06:02.370 And I say this is the kind of data we need.

NOTE Confidence: 0.839879968947368

01:06:02.370 --> 01:06:04.186 I'm mostly talking about for the kind of

NOTE Confidence: 0.839879968947368

01:06:04.186 --> 01:06:05.984 work that I'm talking about rather than

NOTE Confidence: 0.839879968947368

01:06:05.984 --> 01:06:07.662 for everything to solve cancer, of course.

NOTE Confidence: 0.839879968947368

01:06:07.662 --> 01:06:09.230 So, but but in order to understand the

NOTE Confidence: 0.839879968947368

01:06:09.278 --> 01:06:10.762 underlying selective coefficients and

NOTE Confidence: 0.839879968947368

01:06:10.762 --> 01:06:12.246 understand the mutational processes,

NOTE Confidence: 0.839879968947368

01:06:12.250 --> 01:06:14.596 I do think large amounts of.

NOTE Confidence: 0.839879968947368

01:06:14.600 --> 01:06:16.450 Tumor resection data which will

NOTE Confidence: 0.839879968947368

01:06:16.450 --> 01:06:17.560 be gathered anyway,

NOTE Confidence: 0.839879968947368

01:06:17.560 --> 01:06:19.528 but also more of this multi sample data

NOTE Confidence: 0.839879968947368

01:06:19.528 --> 01:06:21.432 so that we can understand dynamically

NOTE Confidence: 0.839879968947368

01:06:21.432 --> 01:06:23.840 over time what's happening which we can't.

NOTE Confidence: 0.839879968947368

01:06:23.840 --> 01:06:25.973 We can do, I said in a probabilistic way,

NOTE Confidence: 0.839879968947368

01:06:25.980 --> 01:06:27.948 but never in a very satisfying way with

NOTE Confidence: 0.839879968947368

01:06:27.948 --> 01:06:29.987 just the tumor genesis resection data.

NOTE Confidence: 0.782779761851852

01:06:31.650 --> 01:06:34.471 It makes the noise that the tumor

NOTE Confidence: 0.782779761851852

129

01:06:34.471 --> 01:06:36.635 cellularity differences we bring in and

NOTE Confidence: 0.782779761851852

01:06:36.635 --> 01:06:39.234 I think it also remains you are gorgeous

NOTE Confidence: 0.782779761851852

01:06:39.234 --> 01:06:41.994 question about the copy number changes.

NOTE Confidence: 0.782779761851852

01:06:42.000 --> 01:06:42.870 So how do you adjust,

NOTE Confidence: 0.782779761851852

01:06:42.870 --> 01:06:45.174 what is that you know if it has

NOTE Confidence: 0.782779761851852

01:06:45.174 --> 01:06:47.867 17 copies of imitation it has that

NOTE Confidence: 0.782779761851852

01:06:47.867 --> 01:06:49.907 signature that will be amplified.

NOTE Confidence: 0.782779761851852

01:06:49.910 --> 01:06:51.595 And it's not necessarily black

NOTE Confidence: 0.782779761851852

01:06:51.595 --> 01:06:53.827 would be the actual sometimes

NOTE Confidence: 0.782779761851852

01:06:53.827 --> 01:06:55.335 higher prevalence of contribution

NOTE Confidence: 0.782779761851852

01:06:55.335 --> 01:06:57.770 of the particular audition process,

NOTE Confidence: 0.782779761851852

01:06:57.770 --> 01:07:00.506 but it's just that the gene.

NOTE Confidence: 0.782779761851852

01:07:00.510 --> 01:07:04.290 I see these questions about the the.

NOTE Confidence: 0.782779761851852

01:07:04.290 --> 01:07:06.775 So the adjacent normal tissues

NOTE Confidence: 0.782779761851852

01:07:06.775 --> 01:07:08.763 requires mutations and they

NOTE Confidence: 0.782779761851852

01:07:08.763 --> 01:07:10.820 actually introduce noise, right?

01:07:12.750 --> 01:07:16.692 Yes. So both of those are sources of noise

01:07:16.700 --> 01:07:20.580 in the sense that on average as we look at,

01:07:20.580 --> 01:07:22.748 so the say talk about a gene amplification

01:07:22.748 --> 01:07:24.279 for instance is a great example.

01:07:24.280 --> 01:07:25.840 When you get a gene amplification,

01:07:25.840 --> 01:07:28.192 you know the the mutation itself may not

01:07:28.192 --> 01:07:30.145 be contributing the cancer effect size that

01:07:30.145 --> 01:07:32.458 we analyze when we get this kind of data.

01:07:32.460 --> 01:07:34.420 But what is true is that those

01:07:34.420 --> 01:07:37.003 mutations and the amount of copy number

01:07:37.003 --> 01:07:39.103 amplification that they typically have

01:07:39.103 --> 01:07:41.223 contributes this amount because we're just

01:07:41.223 --> 01:07:43.706 looking at whether or not we see these.

01:07:43.706 --> 01:07:45.398 The patients and whatever other processes

01:07:45.398 --> 01:07:47.264 are going on, we're averaging over. So.

01:07:47.264 --> 01:07:49.459 So the cancer effect size is still I

NOTE Confidence: 0.916899381111111

01:07:49.459 --> 01:07:51.379 would say it's still the measure of how

NOTE Confidence: 0.916899381111111

01:07:51.439 --> 01:07:53.539 much that mutation is contributing to it.

NOTE Confidence: 0.916899381111111

01:07:53.540 --> 01:07:55.304 But the means by which it contributes

NOTE Confidence: 0.916899381111111

01:07:55.304 --> 01:07:57.248 we don't really know from this analysis.

NOTE Confidence: 0.916899381111111

01:07:57.250 --> 01:07:59.322 It's a it's just that wider question of

NOTE Confidence: 0.916899381111111

01:07:59.322 --> 01:08:01.238 how much is this variant contributing

NOTE Confidence: 0.916899381111111

01:08:01.238 --> 01:08:03.248 and and if it needs amplification

NOTE Confidence: 0.916899381111111

01:08:03.303 --> 01:08:04.598 as part of that process,

NOTE Confidence: 0.916899381111111

01:08:04.600 --> 01:08:06.544 well then we need to do a more

NOTE Confidence: 0.916899381111111

01:08:06.544 --> 01:08:08.234 detailed analysis that looks both at

NOTE Confidence: 0.916899381111111

01:08:08.234 --> 01:08:09.962 amplification and the and the mutation

NOTE Confidence: 0.916899381111111

01:08:10.019 --> 01:08:12.206 and then we'll be able to say like how

NOTE Confidence: 0.916899381111111

01:08:12.206 --> 01:08:13.596 important that mutation is in terms of.

NOTE Confidence: 0.916899381111111

01:08:13.600 --> 01:08:15.424 Cancer affect how important the amplification

NOTE Confidence: 0.916899381111111

01:08:15.424 --> 01:08:17.536 vacation is in terms of cancer effect

01:08:17.536 --> 01:08:18.946 compared to the mutation itself.

01:08:18.950 --> 01:08:19.766 That's not something we've

01:08:19.766 --> 01:08:20.786 been able to do yet,

01:08:20.790 --> 01:08:23.136 but it's something on our agenda.

01:08:23.140 --> 01:08:24.616 It's very difficult but I think

01:08:24.616 --> 01:08:26.090 it's achievable but very difficult.

01:08:28.550 --> 01:08:30.188 I think I better quickly ask,

01:08:30.190 --> 01:08:32.422 I feel sorry for the people

01:08:32.422 --> 01:08:33.910 who ask questions online.

01:08:33.910 --> 01:08:35.926 The one question is,

01:08:35.926 --> 01:08:38.950 is mutation a biochemical reaction to

01:08:39.044 --> 01:08:42.084 TR GRC a substitute of T or G or C?

01:08:42.090 --> 01:08:43.450 The mutations I'm talking about

01:08:43.450 --> 01:08:45.144 in this entire study were all

01:08:45.144 --> 01:08:46.288 single new type mutations.

01:08:46.290 --> 01:08:48.530 In the context of A3,

NOTE Confidence: 0.73867273

01:08:48.530 --> 01:08:50.288 what I meant by trinucleotide context

NOTE Confidence: 0.73867273

01:08:50.288 --> 01:08:52.507 is the 3 mutations in the central one.

NOTE Confidence: 0.73867273

01:08:52.510 --> 01:08:54.385 How was that mutated to

NOTE Confidence: 0.73867273

01:08:54.385 --> 01:08:55.510 another single nucleotide?

NOTE Confidence: 0.73867273

01:08:55.510 --> 01:08:57.050 There are ways to look at doublets

NOTE Confidence: 0.73867273

01:08:57.050 --> 01:08:58.248 there are ways to look at.

NOTE Confidence: 0.73867273

01:08:58.250 --> 01:09:00.212 Some other more complicated indels which

NOTE Confidence: 0.73867273

01:09:00.212 --> 01:09:02.960 we have in the lab almost implemented,

NOTE Confidence: 0.73867273

01:09:02.960 --> 01:09:06.032 but other mutation types we don't

NOTE Confidence: 0.73867273

01:09:06.032 --> 01:09:08.394 have actually looked at Yuval

NOTE Confidence: 0.73867273

01:09:08.394 --> 01:09:10.278 Kluger's question I think thank you.

NOTE Confidence: 0.73867273

01:09:10.280 --> 01:09:13.312 You have echoed that for for me on low.

NOTE Confidence: 0.73867273

01:09:13.312 --> 01:09:15.064 So I believe I answered that.

NOTE Confidence: 0.8935888325

01:09:17.920 --> 01:09:20.482 That you know basically it's true

NOTE Confidence: 0.8935888325

01:09:20.482 --> 01:09:23.350 that we don't know the specific,

134

NOTE Confidence: 0.8935888325

01:09:23.350 --> 01:09:25.022 you know when we talk about this mutation

NOTE Confidence: 0.8935888325

01:09:25.022 --> 01:09:26.680 and how much is cancer effect sizes,

NOTE Confidence: 0.8935888325

01:09:26.680 --> 01:09:28.542 that's in the context of everything that

NOTE Confidence: 0.8935888325

01:09:28.542 --> 01:09:30.235 happens to that mutation in cancers

NOTE Confidence: 0.8935888325

01:09:30.235 --> 01:09:31.915 and it's the average across that.

NOTE Confidence: 0.811270495

01:09:33.970 --> 01:09:35.740 But Tim Robinson has a question,

NOTE Confidence: 0.811270495

01:09:35.740 --> 01:09:37.987 which is, can the spectrum of mutations

NOTE Confidence: 0.811270495

01:09:37.987 --> 01:09:40.416 tell us about the chance that the

NOTE Confidence: 0.811270495

01:09:40.416 --> 01:09:42.166 tumor will respond to treatment?

NOTE Confidence: 0.91101238

01:09:44.610 --> 01:09:47.170 It may well, so for instance you know this,

NOTE Confidence: 0.91101238

01:09:47.170 --> 01:09:49.126 the fact that there were cisplatin

NOTE Confidence: 0.91101238

01:09:49.126 --> 01:09:51.540 mutations is going to tell you that it's

NOTE Confidence: 0.91101238

01:09:51.540 --> 01:09:53.505 likely to have an EGFR T790M resistant

NOTE Confidence: 0.91101238

01:09:53.505 --> 01:09:55.335 mutation sort of sitting there waiting

NOTE Confidence: 0.91101238

01:09:55.335 --> 01:09:57.270 to come out when you give it a lot.

NOTE Confidence: 0.91101238

01:09:57.270 --> 01:09:58.754 So in a sense that spectrum could

NOTE Confidence: 0.91101238

01:09:58.754 --> 01:10:00.387 tell us about the chance that a

NOTE Confidence: 0.91101238

01:10:00.387 --> 01:10:01.567 tumor was bound to treatment.

NOTE Confidence: 0.91101238

01:10:01.570 --> 01:10:03.826 But in general if I could I would

NOTE Confidence: 0.91101238

01:10:03.826 --> 01:10:05.753 rather look at look for EGFR

NOTE Confidence: 0.91101238

01:10:05.753 --> 01:10:08.768 T790M itself directly for example,

NOTE Confidence: 0.91101238

01:10:08.770 --> 01:10:11.034 if the vast majority of mutations are in

NOTE Confidence: 0.91101238

01:10:11.034 --> 01:10:13.306 Melanoma and Melanoma are B rap 600 and the.

NOTE Confidence: 0.91101238

01:10:13.310 --> 01:10:15.045 The vast memory of cancer

NOTE Confidence: 0.91101238

01:10:15.045 --> 01:10:16.780 causation by mutation is there.

NOTE Confidence: 0.91101238

01:10:16.780 --> 01:10:19.044 Does that inform the chance that the tumor

NOTE Confidence: 0.91101238

01:10:19.044 --> 01:10:21.266 will respond to directed therapy to be wrap?

NOTE Confidence: 0.47006366

01:10:23.510 --> 01:10:26.538 Umm, I think the, you know, the number

NOTE Confidence: 0.47006366

01:10:26.538 --> 01:10:28.369 of mutations I don't think does at all.

NOTE Confidence: 0.47006366

01:10:28.370 --> 01:10:30.104 I think that what's important to

NOTE Confidence: 0.47006366

01:10:30.104 --> 01:10:32.040 understand about Viraf E7 and E and

NOTE Confidence: 0.47006366

01:10:32.040 --> 01:10:33.691 it's cancer effect size, which by the

NOTE Confidence: 0.47006366

01:10:33.691 --> 01:10:35.590 way is a very high cancer effect size,

NOTE Confidence: 0.47006366

01:10:35.590 --> 01:10:38.182 is that if you can get a therapy that

NOTE Confidence: 0.47006366

01:10:38.182 --> 01:10:40.367 treats the rap fee 600 effectively,

NOTE Confidence: 0.47006366

01:10:40.370 --> 01:10:42.449 it will be a very effective therapy.

NOTE Confidence: 0.47006366

01:10:42.450 --> 01:10:44.025 And there's a good example of that.

NOTE Confidence: 0.47006366

01:10:44.030 --> 01:10:45.958 And there's a caveat to that example also,

NOTE Confidence: 0.47006366

01:10:45.960 --> 01:10:48.216 which is that the raffish under 600 E,

NOTE Confidence: 0.47006366

01:10:48.220 --> 01:10:49.828 as many people know, there's vemurafenib,

NOTE Confidence: 0.47006366

01:10:49.830 --> 01:10:51.522 which is a very effective therapy

NOTE Confidence: 0.47006366

01:10:51.522 --> 01:10:52.368 for skin cancer.

NOTE Confidence: 0.47006366

01:10:52.370 --> 01:10:53.640 The only problem is there's.

NOTE Confidence: 0.47006366

01:10:53.640 --> 01:10:55.750 Very rapid evolution of resistance.

NOTE Confidence: 0.47006366

01:10:55.750 --> 01:10:57.892 Nothing about cancer effect tells you

NOTE Confidence: 0.47006366

01:10:57.892 --> 01:10:59.990 how quickly resistance will be evolved,

NOTE Confidence: 0.47006366

01:10:59.990 --> 01:11:02.475 and in that case this also interplays

NOTE Confidence: 0.47006366

01:11:02.475 --> 01:11:04.200 with CNV's because at least one

NOTE Confidence: 0.47006366

01:11:04.200 --> 01:11:05.907 of the explanations for why that

NOTE Confidence: 0.47006366

01:11:05.907 --> 01:11:07.895 rapid rises occurs is that you get

NOTE Confidence: 0.47006366

01:11:07.895 --> 01:11:09.624 amplification of the variant BRAF

NOTE Confidence: 0.47006366

01:11:09.624 --> 01:11:12.009 V600E that basically overwhelms the

NOTE Confidence: 0.47006366

01:11:12.009 --> 01:11:14.491 treatment of vemurafenib and means that

NOTE Confidence: 0.47006366

01:11:14.491 --> 01:11:16.780 you and that's a very fast process.

NOTE Confidence: 0.47006366

01:11:16.780 --> 01:11:18.831 Amplification of a gene in a genome

NOTE Confidence: 0.47006366

01:11:18.831 --> 01:11:21.689 is not hard to do as a high mutation

NOTE Confidence: 0.47006366

01:11:21.689 --> 01:11:22.961 rate happens very quickly.

NOTE Confidence: 0.47006366

01:11:22.970 --> 01:11:23.564 Some cells have.

NOTE Confidence: 0.47006366

01:11:23.564 --> 01:11:24.950 More of it somehow is less than

NOTE Confidence: 0.47006366

01:11:24.992 --> 01:11:26.227 those ones with more selected.

NOTE Confidence: 0.47006366

01:11:26.230 --> 01:11:27.958 It's very easy to select on that basis.

NOTE Confidence: 0.47006366

01:11:27.960 --> 01:11:30.232 So so it I think it informs you

NOTE Confidence: 0.47006366

01:11:30.232 --> 01:11:32.121 about how likely a treatment is

NOTE Confidence: 0.47006366

01:11:32.121 --> 01:11:34.871 to have a big effect at the moment

NOTE Confidence: 0.47006366

01:11:34.871 --> 01:11:36.487 you apply the treatment.

NOTE Confidence: 0.47006366

01:11:36.490 --> 01:11:37.795 How quickly you evolve resistance

NOTE Confidence: 0.47006366

01:11:37.795 --> 01:11:38.578 is another question.

NOTE Confidence: 0.5114488

01:11:40.830 --> 01:11:44.804 Umm. And already moustaki the

NOTE Confidence: 0.5114488

01:11:44.804 --> 01:11:47.068 sources of mutations smoking,

NOTE Confidence: 0.5114488

01:11:47.070 --> 01:11:50.208 UV infection affect the normal non

NOTE Confidence: 0.5114488

01:11:50.208 --> 01:11:51.917 transform tissues. Yes they do.

NOTE Confidence: 0.5114488

01:11:51.917 --> 01:11:53.450 Can you use your approach to calculate

NOTE Confidence: 0.5114488

01:11:53.496 --> 01:11:55.302 the cancer effect mutations on the tumor

NOTE Confidence: 0.5114488

01:11:55.302 --> 01:11:56.849 micro movement have on tumorigenesis.

NOTE Confidence: 0.5114488

01:11:56.850 --> 01:11:58.626 One might argue a lot of these mutation

NOTE Confidence: 0.5114488

01:11:58.626 --> 01:12:00.094 sources act on the environment reducing

NOTE Confidence: 0.5114488

01:12:00.094 --> 01:12:02.169 the fitness of a normal cell allowing the.

NOTE Confidence: 0.5114488

139

01:12:02.170 --> 01:12:03.328 This is a really interesting question.

NOTE Confidence: 0.5114488

01:12:03.330 --> 01:12:04.590 We are working on this.

NOTE Confidence: 0.5114488

01:12:04.590 --> 01:12:06.529 So the the bottom line is that

NOTE Confidence: 0.5114488

01:12:06.529 --> 01:12:08.721 and I'll be very quick with this

NOTE Confidence: 0.5114488

01:12:08.721 --> 01:12:11.114 answer that once we are able to

NOTE Confidence: 0.5114488

01:12:11.114 --> 01:12:13.039 figure out these cancer effects.

NOTE Confidence: 0.5114488

01:12:13.040 --> 01:12:15.088 Then we can ask it to the extent

NOTE Confidence: 0.5114488

01:12:15.088 --> 01:12:17.020 that we have annotated data on

NOTE Confidence: 0.5114488

01:12:17.020 --> 01:12:19.054 this tumor was exposed to this

NOTE Confidence: 0.5114488

01:12:19.120 --> 01:12:21.360 given treat this given environment,

NOTE Confidence: 0.5114488

01:12:21.360 --> 01:12:23.810 we can ask how does that environment

NOTE Confidence: 0.5114488

01:12:23.810 --> 01:12:25.380 affect the cancer effect.

NOTE Confidence: 0.5114488

01:12:25.380 --> 01:12:26.796 So we can ask if you're,

NOTE Confidence: 0.5114488

01:12:26.800 --> 01:12:27.900 if you have different ages,

NOTE Confidence: 0.5114488

01:12:27.900 --> 01:12:29.756 not just what mutations are caused by aging,

NOTE Confidence: 0.5114488

01:12:29.760 --> 01:12:31.902 but how much does the cancer effect of a

NOTE Confidence: 0.5114488

01:12:31.902 --> 01:12:33.859 given mutation change as someone ages.

NOTE Confidence: 0.5114488

01:12:33.860 --> 01:12:35.484 So there's ways to do that with

NOTE Confidence: 0.5114488

01:12:35.484 --> 01:12:36.940 the kind of data we have.

NOTE Confidence: 0.5114488

01:12:36.940 --> 01:12:39.590 Again it requires bigger sample

NOTE Confidence: 0.5114488

01:12:39.590 --> 01:12:40.704 sizes in general,

NOTE Confidence: 0.5114488

01:12:40.704 --> 01:12:43.040 but we're looking at that right now with.

NOTE Confidence: 0.5114488

01:12:43.040 --> 01:12:43.916 Regard to smoking,

NOTE Confidence: 0.5114488

01:12:43.916 --> 01:12:45.960 because smoking of course can have a

NOTE Confidence: 0.5114488

01:12:46.021 --> 01:12:48.307 direct effect of mutating individual genes,

NOTE Confidence: 0.5114488

01:12:48.310 --> 01:12:50.732 but it can also have a physiological

NOTE Confidence: 0.5114488

01:12:50.732 --> 01:12:52.866 effect of degrading the normal cells

NOTE Confidence: 0.5114488

01:12:52.866 --> 01:12:54.924 in general in the lung ecosystem.

NOTE Confidence: 0.5114488

01:12:54.930 --> 01:12:56.729 And because you have degraded normal cells,

NOTE Confidence: 0.5114488

01:12:56.730 --> 01:12:57.955 that could increase your chance

NOTE Confidence: 0.5114488

01:12:57.955 --> 01:12:58.690 of getting cancer.

NOTE Confidence: 0.5114488

141

01:12:58.690 --> 01:13:00.196 Or it could mean that certain

NOTE Confidence: 0.5114488

01:13:00.196 --> 01:13:01.764 mutations are more likely to be

NOTE Confidence: 0.5114488

01:13:01.764 --> 01:13:03.300 able to make cancer proliferate and

NOTE Confidence: 0.5114488

01:13:03.300 --> 01:13:04.819 survive better than other mutations.

NOTE Confidence: 0.5114488

01:13:04.820 --> 01:13:05.132 So.

NOTE Confidence: 0.5114488

01:13:05.132 --> 01:13:07.316 So the Physiology could be very important,

NOTE Confidence: 0.5114488

01:13:07.320 --> 01:13:09.008 and there are ways to get at that.

NOTE Confidence: 0.5114488

01:13:09.010 --> 01:13:10.410 But you need to know this first,

NOTE Confidence: 0.5114488

01:13:10.410 --> 01:13:12.419 and then you can ask the question

NOTE Confidence: 0.5114488

01:13:12.419 --> 01:13:13.670 about Physiology affecting things.

NOTE Confidence: 0.5114488

01:13:13.670 --> 01:13:15.994 And I think I'm out of time.